**南 开 大 学**

本科生毕业论文（设计）

中文题目：论桑德尔对罗尔斯“自我观”的建构与批判

外文题目：Sandel's Construction and Criticism of Rawls' "Self View"

学 号：\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

姓 名：\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

年 级：\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

专 业：\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

系 别：\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

学 院：\_\_ 哲学院 \_\_\_

指导教师：\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

完成日期：\_\_\_2020.4.27\_\_\_\_

**摘 要**

《正义论》的发表引起了西方学界的巨大反响。其中，桑德尔以其对罗尔斯自我观的建构和批判而受到广泛关注。国内外学界一般聚焦于桑德尔对罗尔斯正义论的批判，注重研究桑德尔的批判的局限性，而没有充分重视桑德尔对罗尔斯的批判是以其对罗尔斯主体的阐释性建构作为前提的。桑德尔独特性恰恰就在于他是在建构主体中批判主体，因此对建构过程的关注就显得尤为重要。桑德尔从罗尔斯的道义论立场出发，指出其主体所面临的“阿基米德点”问题，并通过阐释其解决方案——原初状态，最终为罗尔斯找到恰当的主体解释，即多样化主体以及预先个体化且优先于目的的占有主体。本文从桑德尔的建构过程入手，试图为其梳理出一条清晰的论证线索，并在此基础上为探究桑德尔建构和批判的局限性，最终发现桑德尔的建构过程并不严密，得到的主体观念也与罗尔斯存在偏差。

关键词：自我;占有主体;反思的平衡;道德人（格）

**Abstract**

The publication of "A Theory of Justice" has caused a huge response from Western academic circles. Among them, Michael J. Sandel received extensive attention for his construction and criticism of John Rawls' self-view. Scholars generally focus on Sandel's criticism of John Rawls' theory of justice, studying the limitations of Sandel's criticism, but do not pay enough attention to the premise that Sandel's criticism of John Rawls is based on his interpretative construction of Rawls' subject. Sandel's uniqueness is precisely that he criticizes the subject in the process of constructing the subject, so the attention to the construction process is particularly important.From the standpoint of Rawls' deontology, Sandel pointed out the problem of "Archimedes' point" that his subject faced, and by explaining his solution-the original state, he finally found an appropriate explanation for Rawls' subject, that is, the essential plurality of the human subject and the notion of a subject of possession, individuated in advance and given prior to its ends. This paper starts with Sandel's construction process, trying to sort out a clear clue for it, and on this basis to explore the limitations of Sandel's construction and criticism, finally found that Sandel's construction process is not strict, and the obtained subject concept is also deviated from Rawls.

**Keyword**: self; a subject of possession; Reflective equilibrium; Moral person (personality)

目 录

[引 言 1](#_Toc39073622)

[一、罗尔斯的“道义论的自由主义” 2](#_Toc39073623)

[（一）正义的首要性 2](#_Toc39073624)

[（二）自我优先于目的 5](#_Toc39073625)

[（三）正当优先于善 6](#_Toc39073626)

[二、“阿基米德点”与原初状态 7](#_Toc39073627)

[（一）“阿基米德点”问题 7](#_Toc39073628)

[（二）原初状态的解决方案 8](#_Toc39073629)

[（三）正义的环境假设的两种解释方式 10](#_Toc39073630)

[（四）反思的平衡的检验方法 13](#_Toc39073631)

[三、反思的平衡与主体 14](#_Toc39073632)

[（一）反思的平衡的逆向应用 14](#_Toc39073633)

[（二）重建主体——多样性主体 16](#_Toc39073634)

[（三）重建主体——预先个体化且优先于目的的“占有主体” 17](#_Toc39073635)

[四、桑德尔对罗尔斯自我观的批判及其局限性 20](#_Toc39073636)

[（一）桑德尔对罗尔斯自我观的批判 20](#_Toc39073637)

[（二）建构方法的问题 21](#_Toc39073638)

[（三）建构主体的问题 23](#_Toc39073639)

[结语 25](#_Toc39073640)

引 言

罗尔斯《正义论》的发表在西方学界引起了巨大的反响，可以说凭借一己之力“扭转了欧美哲学界的研究方向”，使得长期受到冷落、轻视的规范性政治哲学和道德哲学重新回到哲学以及社会科学研究的中心地位[[1]](#footnote-1)，正是在这种意义上，罗尔斯的《正义论》标志着规范政治哲学的当代“复兴”。在当代西方政治哲学领域，罗尔斯的正义理论是奠基性的，罗尔斯之后的正义理论在某种意义上要么是对罗尔斯正义论的辩护和修正，要么是在与罗尔斯正义论的“对峙”中得到发展的。20世纪80年代，社群主义政治哲学便是在对以罗尔斯为首的自由主义政治哲学的批判和质疑中崛起的。从总体上说，社群主义认为，新自由主义错误的根源就在于将原子式的个人作为其社会政治理论的出发点，设定了“自我”相对于社会的优先性理解，因而社群主义者主张，应基于公共利益的社会性立场和视角对“自我”或者主体进行重新界定，以此矫正自由主义政治哲学之不足。可以说，对“自我”的批判直接触及了罗尔斯为代表的自由主义理论的根基，是关系其理论有效性的根本。而在社群主义政治哲学中，直接对罗尔斯自我观进行集中批判的便是桑德尔，凡涉及罗尔斯的自我观的讨论，就必然涉及到桑德尔。

迈克尔·J.桑德尔（以下简称桑德尔）是社群主义批判的代表性人物，奠定其地位的是《自由主义与正义的局限》一书，这本书也被视为批评《正义论》最有力的著作之一[[2]](#footnote-2)。在本书中，桑德尔从对罗尔斯的主体即“自我”观念的建构与批判开始，揭示出罗尔斯的差异原则、共同体观念、“善”理论与其主体理论前提存在的矛盾，基于此，桑德尔指出，罗尔斯要么放弃自由主义的主体概念，要么放弃正义理论的有效性和协调性。

虽然桑德尔对罗尔斯的整个批判看起来非常有力与深刻，但是，其果真如桑德尔所宣称的那样“损害”了罗尔斯正义理论的根基吗？从国内外学界的相关研究看，学者们大都直接评估桑德尔对罗尔斯的批判，并认为由于桑德尔的批判暗含着与罗尔斯相同的出发点，使其失去了批判力度；很少有学者对桑德尔的“建构主体”这个前提性的问题展开研究和批判。因为桑德尔所有的批判都建立在他为罗尔斯建构的主体理论基础之上，所以决定桑德尔整个批判的有效性的关键就在于桑德尔式的“罗尔斯主体”是否符合罗尔斯本人的论述。倘若不符合，便是瞄错了方向，打错了靶子，桑德尔的整个批判也就失去了其意义。因此，本文尝试考察如下问题：桑德尔是如何挖掘出《正义论》中隐含的主体观念的？又如何建构起罗尔斯的主体理论？桑德尔的整个批判过程是否存在偏差，他的重建是否符合罗尔斯本人的思想？在此基础之上，以期揭示桑德尔对罗尔斯主体理论的批判和重建的局限性。

# 一、罗尔斯的“道义论的自由主义”

桑德尔首先将罗尔斯的理论界定为“道义论”性质的，因为在罗尔斯的自由主义中，“正义、公平和个人权利概念占核心地位”，并且以康德的义务论为哲学基础，旗帜鲜明的反对功利主义，主张权利优先于善。桑德尔说：“我所关注的自由主义”——这种以罗尔斯为代表的、在当时政治哲学中占据核心地位的自由主义——“最好应当被称作‘道义论的自由主义’”[[3]](#footnote-3)。

简单来说，道义论的两个主张是：首先，道义论从个体出发，认为每个人都是一个特殊的个体与他人相区分，都有神圣不可侵犯的权利，而且每一个人都有自身的目的、善观念以及利益，这些目的、善观念以及利益都是其自主选择的结果；其次，对于社会以及共同体来说，它们是由这样的个人组成，并且其本身并不预设任何特殊善观念作为目的。社会的安排、制度既不建立在特殊善观念基础之上，也不为某些善观念、目的而服务，而是基于个人权利。因此，那些为了社会善的目的而侵犯个人权利的安排是不可以接受的。

在桑德尔看来，罗尔斯的道义论的自由主义的核心观点是：正义的首要性、自我优先于目的、正当优先于善。

## （一）正义的首要性

桑德尔认为道义论是罗尔斯的核心立场，而正义的首要性正是道义论的应有之义。

罗尔斯在《正义论》开篇就将正义视为社会制度和人类活动的首要德性，他说：“正义是社会制度的首要德性，正像真理是思想体系的首要德性一样……同样，某些法律和制度，不管它们如何有效率和安排有序，只要它们不正义，就必须加以改造或废除。”[[4]](#footnote-4)，即正义具有相对于其他价值的绝对优先性。而桑德尔将之称为“正义的首要性”，认为正义的首要性是相对于社会中各种流行价值而言的首要性，正义是“诸价值的价值”，超越诸价值冲突，为调和价值冲突提供标准。

具体而言，正义的首要性包括两个层面的理解：

首先，正义的首要性是道德意义上的首要性。罗尔斯指出：“每个人都拥有一种基于正义的不可侵犯性，这种不可侵犯性即使以整个社会的福利之名也不能逾越。” [[5]](#footnote-5)桑德尔指出，这称之为道德意义的首要性。无论何种情况，正义的要求都超过任何利益、价值的考量，不能因为普遍的善和公共利益牺牲正义。体现在个人，就是个人权利不可侵犯性，任何以社会公益的名义都不可逾越。

桑德尔认为，道德意义的首要性并不足以将罗尔斯的道义论的自由主义与其他各类自由主义区分开，因为很多自由主义家都强调个人权利的神圣不可侵犯，洛克同样认为人生来就具有基于自然法平等的权利，这种权利比任何联邦所能僭越的权利都要强大。但“道义论的自由主义”要求不止于此，这就引申出正义首要性的第二个层面，认识论层面的首要性。

所谓认识论上的优先性，就是正义的独立获得，即正义原则并非从社会流行的价值、善观念中产生的，而是独立于这些价值、善观念产生，否则就与各特殊善观念无异。认识论意义的优先性可以在康德那里找到源头，“善和恶的概念必须不先于道德的法则（哪怕这法则表面看来似乎必须由善恶概念提供基础），而只（正如这里也发生的那样）在这法则之后并通过它来得到规定。”[[6]](#footnote-6)对于道义论而言，在道德法则之前不能有任何特殊的善观念，道德法则必须独立于特殊善观念而得到的。

在桑德尔看来，道义论的自由主义需要同时满足两者，特别是满足第二个层面，因为认识论意义的优先性是确保正义的首要性、道德意义的优先性的必要条件。很多自由主义者只满足第一个要求，这并不能确保正义的首要性。例如密尔的自由主义。他虽然也主张个人权利的不可侵犯、将正义视为最神圣、最有约束力的部分，但他是把正义建立在社会功利前提下。显然，密尔认为正义具有首要性地位，只不过是因为正义原则比其他道德原则能带来更大的社会功利。然而，桑德尔的理解是，这就抹杀了正义与其他道德原则差别，而且将正义建立在并不可靠的功利主义基础之上，这样的基础最终会通过普遍功利损害正义的首要性，侵犯个人权利。所以正义的首要性，就必须同时满足这两个层面的要求，特别是满足认识意义层面的要求。

那么桑德尔眼中的罗尔斯接下来的问题就在于，如何才能满足正义的认识意义上的优先性，即如何才能独立的、不涉及任何特殊善观念的产生正义原则？这样的正义原则的基础为何？

桑德尔引用康德的观点为道义论做出解答，认为道德法则（正义原则）的基础在于拥有自律意志的实践理性主体，通过先验主体的选择得到道德法则。康德首先从认识论角度出发探究主体。“一个人甚至对于他自身，也就是按照他通过**内部感觉**对自己所具有的知识，也不可以妄自宣称知道自己自在地本身是怎么样的”[[7]](#footnote-7)，通过内省得到的只是感觉而不是关于主体的认识，而杂多的感觉、知觉背后的某个东西才是主体。这样的主体先于并独立于知觉经验，并且负责将杂多的表象统一起来，最终形成单一的意识，否则自我将会是一团杂乱的感觉。正如康德所说“只是由于我能在一个**意识**中理解这些表象的杂多，我才把它们全都称为我的表象；因为否则我就会拥有一个如此驳杂不－的自己，就像我拥有我所意识到的那些表象一样了”[[8]](#footnote-8)，这便是认识论意义上的主体，而这最终会引向实践意义的主体。康德提出，我们既是“经验的客体”，即被经验影响和支配的存在者，同时也是自律的主体。正如康德所说，“他（理性存在者）具有两种立场……第一，就他属于感性世界而言，他服从自然规律（他律），第二，就他属于理知世界而言，他服从独立于自然的、并非经验性的、而只是建基于理性的那些法则”[[9]](#footnote-9)。这意味着，我之中存在着一个超越感性世界的主体，它独立并先于任何经验性内容（如自然环境、善观念），不受其限制，能够依据理性自由的选择道德原则。简而言之，康德最终确立了一个先于且独立于经验、遵循理性的自律主体为道德法则奠基。

桑德尔认为，康德最终将正义奠基在先验主体上，罗尔斯虽然不同意先验主体的做法，但同样是将正义建立在具有选择能力的主体之上。这便引出了道义论的核心概念——自我优先于目的。

## （二）自我优先于目的

康德最终诉诸先验主体为正义原则的独立性奠基，而这启示着罗尔斯道义论对自我的理解：一个能够先于并且独立于欲望、目的，具有与生俱来的选择能力的自我，简而言之，自我优先于目的。

罗尔斯在批驳目的论时明确指出：“从根本上揭示我们的本性的，不是我们的目标，而是这样一些原则……要由这些原则来调节人们借以形成这些目标的背景条件，和人们追求这些目标的方式。因为，自我优先于由它来肯定的目的，甚至一种支配性目的也是由自我在大量的可能性中选择的。……因此，我们应当把目的论学说提出的正当和善之间的关系翻转过来，把正当看作是优先的。”[[10]](#footnote-10)与康德一致，罗尔斯认为，自我优先并独立于其目的，具有选择能力，其目的受到正当原则的规约。

桑德尔认为，这体现出罗尔斯的自我具有道德意义和认识意义两个层面的优先性。首先，自我的优先性是道德上的优先，自我不同于其目的、欲望，而具有独立的价值。自我是在各种目的、欲望背后的，需要我们去尊重的存在者。另外，从认识论意义上，自我要独立于目的、欲望以及外部环境，这是一个与环境划分开、纯粹独立的自我。桑德尔认为，罗尔斯实际上区分出了“我之所是”与“我之所有”，前者是主体的本性，而后者是主体拥有的属性，这些属性对于我来说不是完全必要的，我拥有的属性的变化并不会改变自我本身，这就形成一个完全独立的自我。

可以看到，罗尔斯除了认为自我优先于目的，还论及了正当与善的关系，这正是理解罗尔斯正义理论的另一关键。罗尔斯为我们提供了一把理解正当与善的“钥匙”，罗尔斯认为目的论对于正当与善的理解是错误的，而正确理解正当与善关系的“钥匙”便是自我优先于目的。那么这把“钥匙”又是如何打开正当与善关系之“锁”呢?

## （三）正当优先于善

“正当”与“善”是伦理学两个基本概念，正当从最简单的意义上可以理解为对与错，而善则可以理解为好与不好，围绕这两个概念产生的核心问题就是在道德判断中何者优先的问题。目的论认为，善具有优先性，只要符合特定的善观念、目的的要求，就具有正当性。比如：功利主义将效用视为最重要的的善，一切正当性判断都基于功利的衡量。与之相对的是道义论，其主张正当优先于善，不预设任何善观念的价值，同时也不是所有的善观念都有价值，只有满足正当原则、具有正当性的善观念才有价值。

显然，罗尔斯主张正当优先于善。“在公平的正义中，正当的概念是优先于善的概念的。一个正义的社会体系确定了一个范围，个人必须在这一范围内确定他们的目标。”[[11]](#footnote-11)可以说，罗尔斯的正义理论就是基于正当的要求而建立的，正义原则就是正当性标准，人们可以自由的追求符合正义原则的目的或者欲望。桑德尔认为，与正义首要性一样，罗尔斯的观点也可以理解为道德要求的优先和认识论上的优先。前者意味着正当原则要高于欲望、社会福利的满足，后者要求正当独立于各种特殊的善观念而获得，不以善为基础。这近乎是正义首要性的翻版。

在桑德尔看来，正义的首要性、自我优先于目的以及正当优先于善三者在罗尔斯理论中是相互贯通的。“与正义之首要性紧密相关的一个更为普遍的观念，是正当对善的优先性”[[12]](#footnote-12)，正当的优先性相对于正义的首要性是更一般、更普遍性的概念，正义的首要性基于正当的优先性，遵循着正当优先于善的逻辑。同样，正当的优先性最终通过自我的优先性得到证明。在罗尔斯那里，要正确理解正当与善之间的关系，就必须从自我优先于目的出发。一方面，自我优先于目的体现出来的是一个不受欲望、环境的特殊性、偶然性影响的绝对独立的主体，这样的主体在康德看来就是理性存在自身，只有这样彻底独立的主体才能为普遍的道德原则、正义原则提供正当性证明。另一方面，自我优先于目的意味着在自我选择之前不存在任何目的，因此不存在任何先在的具有价值的善观念，一切善观念都是在自我选择中得到其价值。

概括地说，桑德尔抓住了罗尔斯正义理论中最为根本的命题，即自我优先于目的，并认为正是自我优先于目的确立了正当优先于善，为独立的和优先的正义原则提供正当性证明，使得正义的首要性得以确立。按照桑德尔的思路，最根本的问题就是：罗尔斯正义理论中，优先于和独立于目的的自我是什么样的主体？这个问题触及到了桑德尔对罗尔斯的阐释性建构，桑德尔对罗尔斯正义理论的批判也正是从建构罗尔斯的主体理论开始。

# 二、“阿基米德点”与原初状态

## （一）“阿基米德点”问题

桑德尔发现，罗尔斯虽然赞同道义论的各种主张，但是在主体理论上与康德存在较大分歧。首先，罗尔斯反对先验主体，因为罗尔斯认为先验主体是过于武断的，“对罗尔斯来说，康德的观念模糊而武断，因为，一种抽象的未能具体化的主体如果不武断的话，又如何产生决定性的正义原则？” [[13]](#footnote-13)并且，罗尔斯认为“自我优先于目的”并不绝对导向先验主体，“它否认，在完全缺乏一种经验基础的情况下，惟有先验的和独立的自我才能成为超验的或本体的主体”[[14]](#footnote-14)。

因此，罗尔斯需要在保证道义论前提下，提出一种非先验的、与人类境况联系的主体理论。桑德尔将之称为“阿基米德点”问题。桑德尔指出：“一个完全与其经验既定特征相分离的自我，只不过是一种抽象的意识”，相反则是“一个让位于极其抽象的主体的彻底情境化的主体”[[15]](#footnote-15)，桑德尔的意思是说，我们需要找到一个合适的主体理论，既不能距离经验太远近乎先验，又不能距离经验太近。同样，正义原则也因此面临着“阿基米德点”的问题。如果它来自于社会流行的价值和善观念，那么正义就失去了绝对的首要性；如果把经验完全排除在正义原则来源之外，那么正义就基于同样可疑的先验假设。无论是哪种情况，都不能令人满意，因此，也需要在两者之间找到一个恰当的“阿基米德点”。桑德尔认为，为了解决“阿基米德点”的问题，罗尔斯提出了原初状态理论，试图在原初状态中找到合适的主体。

## （二）原初状态的解决方案

原初状态是罗尔斯解决“阿基米德点”问题的方案，那么问题首先在于：原初状态能否为罗尔斯提供一个合适的点去界定主体？

罗尔斯对正义原则的证成，可分为原初状态假设和契约论证明两部分，前者是前提假设，后者是证成过程，即通过原初状态主体的一致选择为正义原则辩护。因此，要保证正义的首要性，就必须确保主体在原初状态假设中选择正义原则时不受特殊善观念的影响。所以在这个意义上，我们可以将原初状态假设理解为关于主体的理论，其目的是确保主体的选择不受特殊善观念的影响。

具体而言，桑德尔认为，原初状态可以分为两个部分：“一种是指各方所不知道的东西，另一种是指各方确乎知道的东西” [[16]](#footnote-16)，前者是罗尔斯的无知之幕假设。罗尔斯认为在原初状态中人们的认识受到无知之幕限制，人们不知道他的财富、阶级和社会地位，不知道智力、能力等各类自然天赋才能，也不知道自己的善观念、生活追求和心理偏好，同时人们也不知道所处社会的境况、文明情况。后者指原初状态各方具有“弱的善观念”，并且知道“善的弱观念”的价值。

上述两部分就构成了原初状态对于主体的认识层面的限制。桑德尔认为，罗尔斯的设计可以用来回应“阿基米德点”问题，即是说，以非先验的方式满足道义论的要求。首先，无知之幕的假设使得主体在选择正义原则的时候不会受到自己的境遇、目的、欲望的影响，不会受到自然和社会偶然环境的影响，这使主体免于特殊性的影响。其次，弱的善理论。“为了建立这些正当原则却需要依赖某种善概念，因为，我们需要假定各方在原初状态中的动机。由于不允许这些假定危及正当概念的优先地位，用于为正义原则论证的善理论被限制到最为基本的范围”[[17]](#footnote-17)。罗尔斯承认，建立正义原则需要依赖某些善的概念（依赖这些善提供合理选择的标准），但为不让这些善损害正义首要性、正当优先性，便区分出了弱的善理论。弱的善理论将善限制到一些具有普遍特性的首要善（包括权利、权力、自由、机会、收入和财富），即所有“有理性的人”都会意欲的东西，无论一个人持有什么样的生活计划。尽管仅凭普遍的首要善不足以对特殊的目的做出判断和衡量，但正因此，首要善可以凭借其普遍特性满足道义论要求，成为选择正义原则的前提条件。

总而言之，桑德尔指出，无知之幕和弱的善理论两个假设以如下方式解决了“阿基米德点”问题：首先，无知之幕防止我们基于个人利益和偏好选择特殊的原则，排除了特殊经验。其次，弱的善理论实现了正当优先于善。正当（正义原则）具有对善（强理论）的优先性，先于其产生，约束着善（强理论）。同时，正当（正义原则）来自于善（弱理论），而善（弱理论）体现了人类普遍的意欲从而免于特殊性的攻击，也使得正当原则的得出摆脱超验与武断的困境。通过这样的设计，我们既保证了正当的独立性和优先性，维护了道义论立场，也避免陷入先验假设。

那么这样的设计是否存在问题呢？桑德尔讨论了两个相关的批评意见。一个反驳针对的是首要善，罗尔斯认为无论何种生活方式，首要善具有普遍的价值。但反驳者认为诸如权利、自由、机会、收入和财富这些价值，只是“西方自由的布尔乔亚（资产阶级）生活计划的偶然偏好”[[18]](#footnote-18)，只是一些流行价值，与强意义的善无异，无法为正当的独立性提供辩护。第二个反驳针对的是无知之幕。反驳者认为无知之幕过于厚重、过于抽象，以至于脱离人类环境，不能做出任何有决定性的选择，无法得到期望的原则。

但是桑德尔并没有沿着这两个反驳进行下去，因为他关注的问题的关键是，“它（原初状态）是否成功地贯彻了其‘在经验理论的范围之内’重建康德式道德主张和政治主张的意图呢？……更具体地说，原初状态的表述能否在强意义上给出并支持罗尔斯企图开出的正义之首要性主张？”[[19]](#footnote-19)桑德尔认为，罗尔斯的原初状态从整体上看，更像是一种经验主义解释，无法支持道义论，而要探究这个问题，我们必须对原初状态的其他条件进行更深入地考察。

## （三）正义的环境假设的两种解释方式

对于原初状态的假设很重要的一部分就是正义的环境假设。罗尔斯假设存在着一种正义的环境，而正义的环境是使正义成为必要的环境。在具体阐述正义环境之前，罗尔斯也从整体对正义的环境进行了阐述。罗尔斯认为，人类社会同时具有利益一致和利益冲突两种特征，人类进行合作是因为社会合作总是比单独生活要过得好，而利益冲突是因为人们存在不同的追求和目的，对于合作所产生的利益不会轻易达成一致，这就出现了利益冲突。因此，正义之所以会被需要，是因为人们在合作的过程中会产生利益冲突，而正义的环境应当最先保证人们会选择合作的同时会产生利益冲突。

对于正义环境的解释，罗尔斯毫不掩饰的采用休谟的经验主义解释，他说：“休谟对它们（正义的环境）的解释是特别明晰的，我前面的概述对休谟特别详细的讨论并没有增加什么实质性的东西。”[[20]](#footnote-20)罗尔斯认为，正义的环境分为客观、主观两部分。首先，是使人类合作既可能又必要的客观环境，具体而言包括两点，首先是人们之间各方面的相对平等，没有一个人可以压倒其他人；其次，资源的中等匮乏的环境。两者都使得社会合作不成为多余。正义的主观环境涉及到合作的主体，一方面各方存在相近或互补的需求和利益；另一方面，各方有各自的善观念和目的，持有不同的生活计划，而且都认为自己的善观念、目的是值得追求的。前者为合作提供条件，后者为相互冲突提供可能性。在主观环境中，罗尔斯还特别强调了“互无利益关涉”的假设，在原初状态中，合作各方只关心自身的利益，追求自身利益的实现，而不管他人的利益如何。

桑德尔指出，这样一种经验主义解释与道义论存在着深刻的分歧。因为经验主义使得正义依赖于某些条件，无法保证正义的绝对首要性。而使得经验主义解释与道义论共存的方法是证明如下事实，即正义的环境会出现在所有社会，而且对于这些社会来说，正义总是优先于其他美德而被考虑。但是，事实上经验主义解释无法提供充足的证明。桑德尔的看法是，在经验主义解释中“正义是社会制度的首要美德，并非像真理之于理论那样绝对，而是**有条件的**”[[21]](#footnote-21)。首先，正义只会出现在合作的主体发生分歧与冲突条件下，而且只是在存在大量分歧的社会里是首要的德性。但是，并非所有社会都是存在大量分歧的社会，存在某些具有一致的价值和目的的共同体，在这些共同体中只存在较低程度的正义。同时，经验主义解释倾向于将正义看作是一种补救性美德，以道德堕落为条件，当社会道德堕落时可以用正义加以修整。这就意味着，存在着一个与正义相对立的美德，“正如休谟的解释所主张的，正义的补救性质蕴含着另一系列至少是具有同样重要性的美德。”[[22]](#footnote-22)

因此，桑德尔认为，正义环境的经验主义的解释与正义首要性存在着不可弥合的冲突，我们必须抛弃经验主义解释，选择道义论解释。

在这里不难看出，桑德尔实际上是从哲学谱系的角度重新对罗尔斯正义理论的立场给出了定位，这的确有助于我们更深刻的理解罗尔斯的理论。哲学史上存在两种对立的立场：一种是休谟式的经验主义理解，另一种是康德的先验道义论的解释。前者认为，正义是人类约定俗成的产物，完全依赖于人类所处的状态和条件。对于康德来说，正义完全脱离人类经验环境，建立在道德形而上学之上。正义的环境并不是现实的人类社会状况，而是一种抽象的目的王国，在这个王国里面，正义和理性给定的道德原则才会发挥作用。对于康德来说，正义的环境实际上是一个抽象的领域，当人们从特殊的处境中抽象出来（去除个人差异），以一种普遍化的观点去意愿、去行事，便进入了目的王国之中，按照普遍法则行事。

罗尔斯继承并试图超越休谟的经验主义和康德的先验道义论。罗尔斯虽然认同道义论，但并不同意康德的目的王国和先验主体，他认为这些观念由于一些“先验的评价标准和极端空洞的自我观念”[[23]](#footnote-23)而与人类无涉。最终，罗尔斯选择将正义理论奠基在普遍性和对于人类具有可能性的假设之上，用罗尔斯自己的话说，“**必须允许道德哲学如其所愿地应用可能的假定和普遍的事实**”[[24]](#footnote-24)，这是对经验的有限制的应用。在桑德尔看来，罗尔斯立场的本质就是排除人际的特殊差异：“罗尔斯企图通过采用一种对正义环境的经验主义解释但又要排除实际存在的个人之间的偶然性差异的方式来重新规定目的王国概念”[[25]](#footnote-25)。这也正是原初状态各假设所追求的，无知之幕防止人们受到目的、欲望以及环境等等特殊经验的影响，实现了与康德近似的效果，即“从有理性的存在者之间的个人差异中抽象出来，也从他们私自的目的的全部内容中抽象出来”[[26]](#footnote-26)。首要善也排除了人们受到特殊性、偶然性的影响，因为人们拥有的仅仅是一般的善观念，并不知道具体的目的。可以看到，罗尔斯并不要求排除所有经验（而这正是康德所要求的），罗尔斯只是要排除体现人际差别的经验，保留着反映普遍性的经验内容，而正是这一点使得罗尔斯的主体和正义原则保留着与真实世界的联系。桑德尔认为，罗尔斯的尝试最终会陷入一串矛盾之中、顾此失彼。

笔者认为，桑德尔实际上也是据此展开正义的环境的“道义论”解释，将原初状态理解为可能的假设，其中一切假设都是与人类相关联的但又不必要符合现实情况。而且，原初状态中的假设是排除掉偶然性、特殊性的普遍性假设。

桑德尔解释道，罗尔斯可能会如此回答经验主义的反驳：反驳者误解了正义的环境，正义的环境只不过是假设的条件，是原初状态中主体实施意图的条件，并非普通人类生活的实际条件。所以我们并不能用人类现实情况来反驳，而应该判断正义环境的假设在多大程度上能够产生与我们正义信念相应的正义原则。因此，正义环境的描述并不需要是符合经验现实的，无论在什么意义下都只是虚构的环境，并不能直接等同于现实情况，比如原初状态中的互无利益关涉动机假设并不能让我们断言现实世界中人们也是这样的。

但是这样的理解会产生一个新的问题：“如果原初状态的描述性前提在任何直接的意义上并不接受经验的检验，那么，它将接受何种类型的检验呢？”[[27]](#footnote-27)

## （四）反思的平衡的检验方法

与假设相伴而生的就是对假设的检验，既然基于经验的检验已经被证明无效，那么我们应当如何检验原初状态的假设呢？针对什么是原初状态最可取的描述的问题，罗尔斯提出了反思平衡的方法，简而言之，就是从两端出发，其中一端是普遍享有和弱意义上的条件，看能否产生一些有意义并且符合我们深思熟虑后的正义信念的原则。如果能，那么就说明前提是合适的，但是如果不能的话，我们或者修改前提假设，或者修改我们对正义的判断。最终经过这样的几个来回，我们最终会达到一种对原初状态的描述：这种描述既符合原初状态的前提条件，又符合我们深思熟虑并且调整过后的正义判断。而这种平衡实际上是原初状态所产生的正义原则和我们对于正义的判断相符合（也可以理解为原初状态产生的正义原则和由我们正义判断产生的正义原则这两个相契合）。

桑德尔指出，反思的平衡检验方法具有循环论证嫌疑，需要加以调整，他说：“使反思的平衡方法避免成为一种循环论证的，是对每一个方面作出独立判断的判断标准的有效性，尽管该标准是临时的，但我们可以根据它来调节和修正其他的判断”[[28]](#footnote-28)。桑德尔认为，既然反思的平衡从两端同时出发，那么就需要分别对这两端的前提进行检验和证明，而这需要一种独立的标准。对于原初状态的假设一端，桑德尔认为应当检验假设的真实可信性和合理性；对于正义信念一端，应当检验由原初状态前提产生的正义原则的可欲求性，依据的标准便是我们的正义直觉、正义信念。但是判断原初状态假设的真实可信性和合理性并不能依靠我们的直觉，因为直觉可能会产生各种奇异的、甚至脱离人类的实际状态假设，而桑德尔认为，原初状态假设不能脱离人类的情况。

但是桑德尔发现，罗尔斯似乎并不特别明确什么应当作为描述性的标准。罗尔斯曾指出：“我们要证明一种特殊的对最初状态的描述，就要展示它联合了这些共同分享的预设。我们的论证要从广泛接受但却很弱的前提开始，达到比较具体的结论”[[29]](#footnote-29)，桑德尔总结为“普遍享有”和“弱的可取”两个标准，前者认为原初状态的假设应当被各方所广泛接受、共同分享，后者认为原初状态的假设是一个相对较弱的假设。但这两个标准都不够明晰，难以令人满意，需要加以改造。

在前述关于“阿基米德点”的问题中，对于为主体找到一个恰当的定位来说，原初状态是罗尔斯提出的解决方案。因此，对原初状态的理解就显得至关重要。经过讨论，我们也得到一些富有成效的结论：原初状态不是经验性的事实，而应当视为一种假设；原初状态假设独特之处就在于，它不是任意提出的，不是先验的，而是基于可能的假设以及普遍的要求，与人类世界存在关联且排除掉特殊性内容，实现人际间的普遍性。原初状态假设不能用经验事实加以反驳，正确的检验方式是，借助拥有临时的却独立的标准的反思的平衡。

但是，我们仍然遗留一些问题：尚未找到合适的描述性的标准；在原初状态理解之下，罗尔斯的主体是何种主体。对此，桑德尔进行了大胆的尝试，逆向应用反思的平衡方法，一方面确立了罗尔斯的主体理论，另一方面罗尔斯的主体恰好又可以提供描述性的标准。

# 三、反思的平衡与主体

## （一）反思的平衡的逆向应用

为了解决遗留的问题，桑德尔认为，我们必须与罗尔斯保持一定的距离，他说：“如果我们不用罗尔斯的那种语言，至少不用与其整体观念相一致的术语，这些问题都能得到解答。因此，要想找到答案的话，可能先要与其文本保持一定的距离，以便使文本的意义显现出来”[[30]](#footnote-30)，“我们的出发点严格限制在文本之内” [[31]](#footnote-31)，他接着说：“我们可以把罗尔斯的理论作为一个整体来看待，特别是其解决原初状态的某些存在疑问的特征之能力，这是我们用任何其他方法都无法解决的”[[32]](#footnote-32)。以上论述表明：桑德尔要采用一种整体的视角分析罗尔斯的理论体系，而且要采取类似“阿基米德点”的做法，既远离文本，又保证出发点基于文本。

桑德尔着眼于反思的平衡方法，并提出一些具有创造性的做法。“从一个方向并透过原初状态之镜来看，我们可以见出两个正义原则；而从另一个方向来看，我们将看见我们自己形象的反射。如果对称地运用反思平衡的方法，那么，原初状态就必须不只是创造一种道德理论，而且还要创造一种哲学人类学。”[[33]](#footnote-33)这就是说，桑德尔将反思的平衡视为一种完全对称性的推理方法，可以借此得到罗尔斯的主体理论。

如果说罗尔斯选择的方向是从原初状态的假设到正义原则，将原初状态的假设视为既定的去论证正义原则。那么桑德尔就要进行逆向应用，从相反的方向出发，对称的应用反思的平衡，前一个方向可以透过原初状态看到正义原则，那么桑德尔的方向就是从正义原则出发看到原初状态的假设。桑德尔进一步认为，原初状态的假设可以反映出罗尔斯的主体假设。这就意味着，如果将符合我们正义判断的正义原则作为临时给定的前提，就推导出适合正义原则的原初状态以及原初状态主体。

如此可以达到两点有意义的成果：首先，如果这个逆应用是正确的，那么就意味着原初状态假设独立的判断标准是“道德主体的本性”，即对自我的构成性理解。只要符合主体的本性，假设就是合理的，这可以避免假设的任意性问题。其次，找到了一种挖掘罗尔斯主体理论的方法，而这将帮助我们理解罗尔斯正义理论的根基——主体，从而可以利用主体来检验整个正义理论。

但这仅提供了大致的思路，还需要更为具体的规定。桑德尔指出了出发点和即将探讨的问题：既然从体现人们正义信念、正义判断的正义原则出发，那么就可以合理的假定主体是具有正义能力、以正义作为首要美德的主体，那么这样的主体应当是怎么样的主体，对于这样的主体，什么必须为真呢？同时，桑德尔认为主体的描述“在某种意义上它是经验的，但又不‘仅仅’是经验的”[[34]](#footnote-34)，对主体的描述不是分析命题，不能理解为从一个概念推导出主体的规定，但主体又必须优先于任何特殊的经验、是必然的。并且，这种主体理论并不是单纯的描述性的，并不是单纯对外在于我们的对象进行描述，而且也是建构性的，是我们通过反思参与其中的。“我们对它们的了解是使它们为真的一部分，也是使我们成为我们所是的能反思的、能自我理解的生物之一部分”[[35]](#footnote-35)，我们对主体理论的了解，即对自我的反思，体现了我们作为一个能反思的生物的特征，也使得主体理论为真。

最后，桑德尔认为，对主体的解释是一种哲学人类学解释，“就其作为哲学的解释而言，是指它是通过反思而非经验的普遍化所达到的，就其作为人类学的解释而论，它又关系到人类主体在多种可能形式的认同中的本性”[[36]](#footnote-36)，前者指获得的方式，通过自我反思获得的主体理论，而不是将表现人类本性的经验普遍化形成人类的主体理论；后者指我们讨论的目标，达到一种关于人类主体本性的认识。

在此基础之上，桑德尔对罗尔斯的主体进行了阐释性的“建构”。

## （二）重建主体——多样性主体

罗尔斯本人是在反驳功利主义的时候提的多样性的，罗尔斯认为功利主义把个人选择的原则拓展到社会是没有道理的，功利主义的做法实际上就是无视人们之间的善观念、目的的差异，把社会视为单一的大个的人，强行实现人类的统一。因此，正确的做法应该是尊重人类的多样性本性。罗尔斯说到：“如果我们承认调节任何事物的正确原则都依赖于那一事物的性质，承认存在着目标互异的众多个人是人类社会的一个基本特征，我们就不会期望社会选择的原则是功利主义的” [[37]](#footnote-37)，因为调节事物的原则应该基于事物本身的性质而给予，人类具有多样性是一个既定事实，那么就应当基于人类（整体意义的人类）的多样性本性调节人类本身。

但是，桑德尔式并非基于罗尔斯的上述观念而构想主体的多样性本性。如上文所说，可以合理假定主体是具有正义能力、以正义为首要美德的主体。桑德尔认为，由于正义出现的必要条件是互异人类的合作出现分歧，那么只要正义出现，就必然存在“数量繁多”的主体。因此，桑德尔解释到：“对罗尔斯来说，任何具有正义能力的生物之首要特征，是其数量的繁多”[[38]](#footnote-38)，即多样性。

既然正义的前提是人类的社会合作，那么统一性也应当具有其地位。问题是：统一性是否也应当成为道德主体的本性？桑德尔认为，将统一性也视为罗尔斯的社会合作的正义性前提，是对罗尔斯的误解，统一性并不具有和多样性同等的优先性。首先，统一性蕴含着人类主体多样性的必要前提，因此人类主体的统一性派生于多样性，无法和多样性共同成为人类主体的本性。进一步说，在社会利益冲突和一致关系问题上，合作的主体发生冲突是由于人类的多样性本性，是必然的。而利益一致则因为人类碰巧具有类似的需求和利益，是不必然的。同时，桑德尔也从知识论的角度进行论证，认为自我是独特的个体是先于经验而给定的，对应着多样性。个体间目的的契合是一个经验的问题，无法预先知道，因为每个人的特殊目的本身就是一个经验的问题，这与统一性相对应。

因此，按照桑德尔对罗尔斯的阐释，在罗尔斯那里，作为一个拥有正义能力的主体的第一个（提及的先后顺序）本性就是多样性：“每一个体的人都是一个道德主体，而且每一道德主体都是一个个体的人”[[39]](#footnote-39)。这是自我与他人的关系问题。还有另外一个平行问题，即自我与目的的关系问题。

## （三）重建主体——预先个体化且优先于目的的“占有主体”

前文已经指出，罗尔斯面临着“阿基米德点”问题，要在非先验的前提下，为优先于目的主体找到合适的位置。对此，桑德尔的解读认为，正是罗尔斯的独特立场使其需要同时解决两个问题：自我如何与其目的区分开来，以及自我如何与其目的相联系[[40]](#footnote-40)。因为主体既不是完全先验的，与目的存在联系，也不是完全情境化的，需要与目的区分开。

桑德尔认为，回答这两个问题需要“占有主体”这个观念，而这又是隐藏在罗尔斯的“互无利益关涉”假设当中的。桑德尔指出：“在互无利益关涉的假设中，我们发现了罗尔斯主体观念的关键，发现了我们必将成为的、将正义作为首要美德的主体图像。”[[41]](#footnote-41)首先，桑德尔认为，不能将互无利益关涉的假设理解为心理学假设或动机假设，这会让自我受到经验的规定而失去其独立性，使正义丧失首要性。相反，应该将其理解为主体的本性假设，这是对自我的本性的理解，而不仅仅是对充斥着偶然性的自我欲望、偏好和目的的理解。因此，互无利益关涉的假设预示着主体从本性上与自己的利益、目的相关联，即一切利益必须是某一主体的利益。而这就蕴含着一个主体观念——占有主体。

何谓占有？在桑德尔看来，“占有”指主体拥有特征、善观念、欲望或者目的。占有的独特性就在于它的双重性（正好满足罗尔斯需要解决的自我与目的区分/关联的问题）：一方面我与占有对象相关（它是我的而不是你的），另一方面我又与占有对象相别（它是我的而不是我）。这也就是“我”与“我的”的区分，占有实际上就是通过区分出“我的”把“我”从偶然性当中拯救出来，保留了自我的尊严和完整性。因此，“占有主体”就是既与占有对象相区分（“我的”不是“我”），但同时又以某种方式与占有对象相关联的主体。

但是，“占有”不是稳定的，会以两种不同的方式丧失。桑德尔认为，一方面当“我”与占有对象的距离逐渐消失的时候，它逐渐成为“我”，而不是“我的”。因此“我”便逐渐失去了对它的占有，甚至反而被它所占有，难以与之相分别。另一方面，当“我”放弃了占有对象时，比如不再抱有某个目的，我也失去了对它的占有。而这两种占有的丧失实质上是对自我的剥夺。前者是对自我的边界和独立性的剥夺，各种欲望和目的冲击着“我”，使“我”无法分清自我的边界终止于何处，我的属性的范围又开始于何处。后者是对自我的连续性的剥夺，丧失的目的和欲望曾经属于我，构成了我的生活计划，代表着我的连续性。

为此，桑德尔进一步提出，有两种不同的修缮方式可以修复自我的剥夺，前一种是认知维度的“发现”，因为“我”丧失了对于自我边界的认知，所以需要通过一种反思性的自我认知，来探究自我的本性，从而在自我与目的之间创造出一个空间。而另一种是唯意志论维度的“选择”，因为“我”失去了我的目的或者欲望，所以需要发挥意志选择的作用，跨越主体和对象之间的空间，将两者联系起来。

经过上述分析，可以发现，“占有主体”本身实际上包含着两种可能的主体理论：第一种是“目的优先于自我”的自我，自我被先在的目的所规定，因此自我需要通过认知维度进行自我理解。第二种是“自我优先于目的”的自我，自我是先在的，其边界是预先给定的，因此问题就不是自我认识，而是唯意志维度的目的选择。最关键的是，“占有主体”概念本身并不否认任何一种可能的合理性，正如桑德尔所说：“作为占有主体的自我概念，并不能单独实现这一（主体）图像[[42]](#footnote-42)”，必须依靠某种前提限制——“先在个体化的主体[[43]](#footnote-43)”——将自我的界限先在的固定下来，从而才能形成完整的主体理论图像。于是，顺着桑德尔的思路，最终他为罗尔斯找到了“预先个体化且优先于其目的而给定的占有主体的概念” [[44]](#footnote-44)。

不仅如此，桑德尔还指出，从反思的平衡逆应用的角度也可以得到同样的结论。假定的符合正义信念、判断的正义原则并没有赋予主体以某种目标，而是强调主体自身的选择能力，得到的便也是唯意志论维度的选择主体。

最后，桑德尔指出，该主体并没有预设任何特殊的善观念。原因在于，为主体做辩护的互无利益关涉假设并不是动机假设，而是关于“主体本性的假设”，前者预设着善观念，而后者是对自我本性的判断，先于善观念，所以任何善观念对于该主体都是可能的，无论是追求个人利益还是重视共同体价值。桑德尔肯定地说：“所有的利益价值和善观念都对罗尔斯式的自我开放，只要这些利益、价值和善观念能被算作是预先个体化且优先于其目的给定的主体的利益。”[[45]](#footnote-45)

至此，桑德尔对于罗尔斯的主体的建构已经完成。我们可以看到，桑德尔从相反的方向运用反思的平衡，从符合正义信念、判断的正义原则出发进行推导，得到了多样性主体以及预先个体化且优先于目的的占有主体。正是在此基础上，桑德尔才对罗尔斯的自我观及其正义理论展开批判的。对我们来说，问题在于：桑德尔对罗尔斯的建构性阐释是否成功，桑德尔对罗尔斯的批判是否具有局限性呢？

# 四、桑德尔对罗尔斯自我观的批判及其局限性

## （一）桑德尔对罗尔斯自我观的批判

通过前面的分析，桑德尔从道义论的立场出发从罗尔斯的理论中阐释性地建构出一套主体理论。简而言之，桑德尔认为，罗尔斯的自我是一个脱离偶然性、摆脱外部环境影响的“无拘的自我[[46]](#footnote-46)”，任何特征都是外在于自我的，并且试图通过这样的自我来为正义原则奠基。

桑德尔的靶子就是这种被他建构出来的“罗尔斯的自我”。他多次指出，罗尔斯的做法最终导致了自我的消解：“罗尔斯试图通过使自我摆脱世界来假定自我的自律。他的批评者说，他为了保留自我，却最终消解了自我”[[47]](#footnote-47)；“只剩下一个缺乏在经验上可确认的特征乃至消解在抽象之中的自我”[[48]](#footnote-48)；罗尔斯保留的自我是非现实的，是摆脱任何经验特征、幽灵般的存在。不仅如此，桑德尔还指出，罗尔斯的自我忽视了我们之间的“共同性”，仅仅将共同体的价值视为主体的可能目标。“如果说，功利主义没有认真对待我们的差异性，那么，公平正义就没有认真对待我们的**共同性**。当公平正义把自我的界限视之为优先的，并将之一劳永逸地固定下来时，它也就把我们的共同性降格为善的一个方面，进而又把善降格为纯粹的偶然性，成为一种与‘道德立场无关’的任意需求和欲望的产物”[[49]](#footnote-49)。在桑德尔看来，共同性不应当仅仅被视为偶然的善观念，而应当被视为自我的构成部分，与此相应，共同体也是主体的构成性部分。

诚如有学者指出：“自由主义者夸大了我们与（我们的）社会关系保持距离并抽象自我的能力，从而夸大了我们个人选择的能力和价值”[[50]](#footnote-50)，如果自我真如罗尔斯所设想的那样，那么“在反省自己的时候，我们就应该能够透过我们的特殊目的看到一个无拘的自我”，但“我们的自我洞察表明，自我并非无拘的”[[51]](#footnote-51)，而是由厚重的特征所限制的。因此，桑德尔认为自我并不是完全独立的，优先于目的的，而是由我们的目的、善观念以及所属的共同体所构成，这是有其洞见的。

那么我们应当如何评价桑德尔的理论工作呢？学界一般从相似性入手，指出桑德尔的自我与罗尔斯的自我存在内在的一致性，从而弱化了桑德尔的批判力度。桑德尔虽然反对优先于目的的占有主体，主张构成性的自我，但其自我观中仍然暗含着“罗尔斯似的占有的自我”，因为“桑德尔并不同意自己所主张的就是一个彻底情境化的自我，**为目的所建构的自我能够重新建构**”[[52]](#footnote-52)。也就是说，桑德尔的自我并不是完全被目的、共同体所界定，还具有选择的维度，“桑德尔承认个人可以重新检查他的目的——就算承认目的构成了他的自我”[[53]](#footnote-53)。这一点我们在分析桑德尔的“占有主体”时其实已经初现端倪，我们发现桑德尔的“占有主体”概念同时包含着两种可能的主体理论，其中“目的先于自我”的“占有主体”，就是桑德尔所隐含的自我观。

通过对比桑德尔建构的主体与罗尔斯本人观念不同之处可以看出，桑德尔不是简单的批判主体，而是在建构主体中批判主体。因此，对桑德尔的评价就包含着两个不同的问题：桑德尔“建构的罗尔斯主体”有何问题或局限？桑德尔对“建构的罗尔斯主体”的批判有何问题或局限？学界往往关注后一个问题，但在我们看来，前者是后者成立的逻辑前提。

## （二）建构方法的问题

判断一种阐释性的建构成功与否，考察其方法是否正确是第一要务。桑德尔通过逆向应用反思的平衡来建构罗尔斯的主体是需要我们进行深思的。我们认为，这个方法的运用会使桑德尔陷入一种他未意识到的矛盾，而且桑德尔本人也没有自始至终运用这个方法。

如前文所述，罗尔斯的反思的平衡方法实际上是检验“两个”正义原则是否一致的方法，即一方是由原初状态假设推导得出的正义原则，另一方是与正义信念、判断相符的正义原则。如果二者不一致，要么就要调整原初状态假设，再次推导正义原则；要么调整我们的深思熟虑的正义判断，使之符合推导产生的正义原则。从这个解释中可以看到，从原初状态假设到正义原则的推导是在先的，是后续一致性检验的前提条件，没有前面的第一次推导，就没有后续的再次调整。因此，反思的平衡不是完全对称的，而是有方向的（第一次推导的方向）。而且这个方向需要得到辩护，否则就没有理由采取这个方向而不采取另一个方向。从罗尔斯的正义原则推导过程来看，原初状态假设到正义原则的推导方向，是基于契约论传统以及具有选择能力的“道德人”。这点也得到罗尔斯本人的承认，他明确指出：“该书（正义论）力求明确地从洛克、卢梭和康德所代表的社会契约论中开出一种正义论”[[54]](#footnote-54)。何怀宏教授也对康德的契约论思想进行过概括：“道德原则是理性选择的目标，它不仅是为所有人接受的、普遍有效的原则，而且是公开的原则。而这种普遍的道德立法又是在人作为自由和平等和理性存在物的条件下被一致同意的，是一种意志的法律”[[55]](#footnote-55)。这表明，无论是正义原则产生过程，还是“道德人”的前提基础，都可以发现罗尔斯的理论与康德契约论思想具有很大的关联性，都是基于理性存在者的选择产生正义原则。

但是，桑德尔却将逆向应用视为理所当然，没有为其提供任何辩护，这显然是武断的。即便我们暂且将逆向应用视为得到辩护的，在其应用过程中也会得到一个自相矛盾的结论。逆向应用反思的平衡是从与深思熟虑的正义判断相对应的正义原则推导出的原初状态假设以及主体。在此，实际上是根据我们对于正义的认识来建构出原初状态以及处于原初状态的主体，而这个过程无疑蕴含着一个目的优先于自我的立场、一个认识意义上的占有主体，而这正与得出的结论——自我优先于目的的占有主体直接相矛盾的。

即便假定这个方法在应用过程中不会产生问题，桑德尔也没有像他所说的那样自始至终遵循着这个方法来论证主体。这直接体现在关于“占有主体”的证明中。前文说过，桑德尔是从互无利益关涉的假设出发的，并发现其中蕴含有占有主体概念，在此基础上对占有主体展开讨论。而这实际上是挖掘文本内容背后的前提，不是通过正义原则进行推导。在得出结论后，桑德尔又从拥有正义感，以正义为首要美德的主体对占有主体进行推导，但这会令人怀疑，因为这更多是验证意义上的推导，已经预设了推导的结果。

因此，无论从方法合法性、方法本身还是推论中对方法的坚持上，反思的平衡“逆应用”都是存在问题的。

## （三）建构主体的问题

桑德尔建构了两个主体概念——多样性主体和预先个体化且优先于目的的占有主体。事实上，桑德尔所建构的主体是抽象的形而上学主体，其非但与罗尔斯的自我观不一致，反而正是罗尔斯想极力避免和反对的。

罗尔斯论证到：“契约论假设了各方是道德人和理性人”[[56]](#footnote-56)，“它（两个正义原则）平等地把每一个人看作一个道德人来对待”[[57]](#footnote-57)。显然，罗尔斯的个人概念不是“主体”（罗尔斯很少使用主体），而是“道德人（格）”。进而，罗尔斯认为道德人（格）具有两种特征：第一种是善观念的能力，即制定合理的生活计划的能力；第二种正义感的能力，按某种正当原则调节自己行为的欲望。罗尔斯后来也用rational和reasonable形容道德人，善观念的能力可以理解为理性（rational），正义感的能力理解为合理性（reasonable）[[58]](#footnote-58)。正如弗雷曼所说：“自由而平等的人具有两个道德能力。第一，讲究‘合理的'（to be reasonable）能力，支持正义的道德能力，理解、利用合作条款并且与他人合作的能力；第二，讲究‘理性的’（to be rational）能力，拥有‘理性的’善观念，形成、修正、理性地追求融贯价值的能力，它为有关给予生命和生命的追求以意义的观念奠定了基础”[[59]](#footnote-59)。可以看到，理性（rational）代表经济学意义的工具理性，即如何能够以最有效的手段实现目的；而合理性（reasonable）是道德理性或价值理性，即是否符合正义、合作要求[[60]](#footnote-60)。对于罗尔斯来说，这两者对于道德人都是必要的。正义原则依靠理性（rational）选择，但国内外学者普遍认为，仅凭理性并不足以选择正义原则，因此合理性就成为证成正义原则必不可少的环节。因此，对罗尔斯的主体来说，理性与合理性都必不可少，甚至合理性更为重要。张国清教授指出，“罗尔斯认为，‘reasonable’和‘rational’是一种词典编纂式排序，reasonable始终排在rational之前，‘合理性预设着理性，且理性从属于合理性’”[[61]](#footnote-61)，人际社会合作需要合理性作为价值前提，确保合作得以形成。

桑德尔在建构主体时却只关注体现“理性”的占有主体，忽视了具有正义能力的主体。虽然桑德尔提到了将正义视为首要美德的主体，但也只是存在于前提假设之中，显然，由此推得的“占有主体”才是桑德尔关注的重点。他只将“占有主体”认定为“在未失之于先验的情况下换回道义论伦理所要求的概念”[[62]](#footnote-62)，后面也只是基于该主体观念对罗尔斯正义理论其他部分展开批判，几乎不再提具有正义能力的主体。桑德尔如此做的原因就在于：自我是正义首要性和正当的优先性的基础。因此，反映自我优先于目的的“占有主体”便是主体的本质特征，而将正义视为首要德性的主体被前者所解释，是前者的必然结果。因此，桑德尔实际上将正义主体还原为占有主体，而这是对罗尔斯的片面理解，甚至与其立场正相反。

此外，正如桑德尔以“哲学人类学”作为其目标，他实际上建构出的是具有哲学意义的形而上学主体。这也得到桑德尔本人的承认，“它最终导致我们成为了这样一个主体：一个被剥夺了经验的自我特性的主体，这同样类似于罗尔斯决心避免的康德式的超验的或抽象的主体”[[63]](#footnote-63)。但是，罗尔斯并无意建构抽象的形而上学主体。在其晚期著作《政治自由主义》中，罗尔斯对于个人概念的质疑做出回应与澄清。如果说《正义论》仍然诉诸康德式的个人观念，那么《政治自由主义》可以说是完全舍弃了任何可疑的形而上学成分。他重新建构了一种政治自由主义以区别于完备性的自由主义，后者将其理论建立在一些普遍意义的道德、哲学学说之上，规定着一种对个人、正义、政治等等的普遍性理解。在理性多元论和文化多元主义时代背景下，理性多元论是一种社会事实，即现代民主社会中存在着诸多不可调和的、相互反对着的宗教、道德观念，完备性的自由主义已经失去了其吸引力而且无力解决人们之间的差异。因此需要一种政治自由主义。罗尔斯认为，政治自由主义拒绝参与到哲学、宗教以及道德层面的争论之中，也不以其为依据，而是将政治作为一种独立的中立性的领域，最终试图达到“重叠共识”来解决政治实践问题。与之相应，罗尔斯提出“政治的个人概念”，它基于“民主社会的公共政治文化”而独立于任何道德、宗教学说。政治的个人概念将个人限制在自由而平等的公民范围之内，作为公民在进行公共活动的时候独立于自己的特殊观念和目的，无论我们在拥有何种宗教、道德观念。这意味着，罗尔斯的政治自由主义完全排除一种对自我的哲学解释，而仅仅是从所谓的“民主社会的公共政治文化”寻找其根源。

因此，桑德尔用形而上学意义的主体理解罗尔斯是不合适的。当然，罗尔斯虽然反对形而上学，但是，他在《正义论》中对“道德人（格）”的解释没有完全摆脱形而上学的痕迹。

结语

桑德尔从确定罗尔斯的道义论立场出发，发现了罗尔斯需要解决的“阿基米德点”问题，于是从“阿基米德点”问题的解决方法——原初状态的假设——切入，讨论原初状态自身的性质、证明以及对于主体的限制与约束，最终在寻找原初状态假设的合法性证明、检验方法找到了反思的平衡方法。桑德尔在此通过逆向应用反思的平衡方法，以将正义视为首要美德的主体为出发点，对罗尔斯的主体进行建构。

本文着重探讨“桑德尔‘建构的罗尔斯主体’有何问题或局限”，最后发现桑德尔建构方法的合法性存在问题，并且还隐藏着一个目的优先于自我的立场，与罗尔斯的立场完全相反；建构的结果也与罗尔斯存在准确性和性质的问题。可以说，桑德尔对罗尔斯主体的建构并非其设想的那么成功。但是我们也不能完全否认桑德尔对罗尔斯主体理论建构的贡献，正是桑德尔以及诸多社群主义者的深刻批判，才有《政治自由主义》的产生。从这个意义上来讲，桑德尔对罗尔斯主体理论的批判也击到了《正义论》的软肋上。

**参考文献**

**著作：**

1. 费雷曼.罗尔斯[M].张国清译.北京：华夏出版社，2013年.
2. 何怀宏.正义理论导引——以罗尔斯为中心[M].北京：北京师范大学出版社，2015年.
3. 康德.道德形而上学奠基[M].杨云飞译.北京:人民出版社,2013年.
4. 康德.实践理性批判[M].邓晓芒译.北京：人民出版社,2003年.
5. 康德.纯粹理性批判[M].邓晓芒译.北京：人民出版社，2004年.
6. 迈克尔·桑德尔.自由主义与正义的局限[M].万俊人等译.南京：译林出版社，2011年.
7. 威尔·金里卡.当代政治哲学[M].刘莘译.上海：上海三联书店，2004年.
8. 约翰·罗尔斯.正义论（修订版）[M].何怀宏、何包钢、廖申白译.北京：中国社会科学出版社，2009.
9. 约翰·罗尔斯.政治自由主义[M].万俊人译.南京：译林出版社，2011年.
10. Catherine Galko Campbell.Persons Identity, and Political Theory A Defense of Rawlsian Political Identity[M].New York Dordrecht Heidelberg London:Springer,2014.

**论文：**

1. 程广云.从理想原则到现实关怀——罗尔斯理性主义解析[J].哲学动态,2005(05):34-39.
2. 龚群.当代社群主义对罗尔斯自由主义的批评[J].中国人民大学学报,2010,24(01):9-15.
3. 龚群.桑德尔对自由主义自我观的批评[J].中山大学学报(社会科学版),2011,51(03):149-154.
4. 韩震,李伟.桑德尔对罗尔斯“自我”概念的批判[J].国外社会科学,1998(02):28-33.
5. 李伟.桑德尔政治哲学研究[D].苏州大学,2006.
6. 万俊人.罗尔斯问题[J].求是学刊,2007(01):14-23.
7. 汪建达,吴新民.论桑德尔对罗尔斯新社会契约论的批判——兼与姚大志先生商榷[J].浙江大学学报(人文社会科学版),2002(06):140-143.
8. 徐友渔.当代西方政治哲学的若干理论[J].国外社会科学,2001(05):2-9.
9. 姚大志.桑德尔的人类学重构及其还原[J].社会科学研究,2012(01):111-116.
10. 应奇,周建华.社群主义的自我观──兼评社群主义与新自由主义之争[J].长春市委党校学报,2001(03):59-63.
11. 张国清,汪远旺.正义感是人的首要美德——罗尔斯双重理性假设之考察[J].吉林大学社会科学学报,2016,56(04):152-159+192.
1. 万俊人:《罗尔斯问题》，求是学刊，2007年第1期 [↑](#footnote-ref-1)
2. 李伟.《桑德尔的政治哲学研究》：[博士学位论文].苏州大学，2006年 [↑](#footnote-ref-2)
3. 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，2011年第1版，第12页 [↑](#footnote-ref-3)
4. 罗尔斯：《正义论（修订版）》，何怀宏、何包钢、廖申白译，北京：中国社会科学出版社，2009年第1版，第3页 [↑](#footnote-ref-4)
5. 罗尔斯：《正义论（修订版）》，何怀宏、何包钢、廖申白译，北京：中国社会科学出版社，2009年第1版，第3页 [↑](#footnote-ref-5)
6. 康德：《实践理性批判》，邓晓芒译，北京：人民出版社，2003年第1版，第85页 [↑](#footnote-ref-6)
7. 康德：《道德形而上学奠基》，杨云飞译，北京：人民出版社，2013年第1版，第97页 [↑](#footnote-ref-7)
8. 康德：《纯粹理性批判》，邓晓芒译，北京：人民出版社，2004年第1版，第90页 [↑](#footnote-ref-8)
9. 康德：《道德形而上学奠基》，杨云飞译，北京：人民出版社，2013年第1版，第98页 [↑](#footnote-ref-9)
10. 罗尔斯：《正义论（修订版）》，何怀宏、何包钢、廖申白译，北京：中国社会科学出版社，2009年第1版，第443页 [↑](#footnote-ref-10)
11. 罗尔斯：《正义论（修订版）》，何怀宏、何包钢、廖申白译，北京：中国社会科学出版社，2009年第1版，第25页 [↑](#footnote-ref-11)
12. 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，2011年第1版，第31页 [↑](#footnote-ref-12)
13. 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，2011年第1版，第26页 [↑](#footnote-ref-13)
14. 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，2011年第1版，第26页 [↑](#footnote-ref-14)
15. 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，2011年第1版，第35页 [↑](#footnote-ref-15)
16. 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，2011年第1版，第38页 [↑](#footnote-ref-16)
17. 罗尔斯：《正义论（修订版）》，何怀宏、何包钢、廖申白译，北京：中国社会科学出版社，2009年第1版，第311页 [↑](#footnote-ref-17)
18. 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，2011年第1版，第41页 [↑](#footnote-ref-18)
19. 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，2011年第1版，第42页 [↑](#footnote-ref-19)
20. 罗尔斯：《正义论（修订版）》，何怀宏、何包钢、廖申白译，北京：中国社会科学出版社，2009年第1版，第98页 [↑](#footnote-ref-20)
21. 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，2011年第1版，第46页 [↑](#footnote-ref-21)
22. 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，2011年第1版，第46页 [↑](#footnote-ref-22)
23. 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，2011年第1版，第52页 [↑](#footnote-ref-23)
24. 罗尔斯：《正义论（修订版）》，何怀宏、何包钢、廖申白译，北京：中国社会科学出版社，2009年第1版，第40页 [↑](#footnote-ref-24)
25. 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，2011年第1版，第53页 [↑](#footnote-ref-25)
26. 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，2011年第1版，第53页 [↑](#footnote-ref-26)
27. 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，2011年第1版，第58页 [↑](#footnote-ref-27)
28. 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，2011年第1版，第58页 [↑](#footnote-ref-28)
29. 罗尔斯：《正义论（修订版）》，何怀宏、何包钢、廖申白译，北京：中国社会科学出版社，2009年第1版，第14页 [↑](#footnote-ref-29)
30. 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，2011年第1版，第62页 [↑](#footnote-ref-30)
31. 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，2011年第1版，第63页 [↑](#footnote-ref-31)
32. 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，2011年第1版，第62页 [↑](#footnote-ref-32)
33. 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，2011年第1版，第63页 [↑](#footnote-ref-33)
34. 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，2011年第1版，第65页 [↑](#footnote-ref-34)
35. 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，2011年第1版，第65页 [↑](#footnote-ref-35)
36. 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，2011年第1版，第66页 [↑](#footnote-ref-36)
37. 罗尔斯：《正义论（修订版）》，何怀宏、何包钢、廖申白译，北京：中国社会科学出版社，2009年第1版，第23页 [↑](#footnote-ref-37)
38. 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，2011年第1版，第66页 [↑](#footnote-ref-38)
39. 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，2011年第1版，第68页 [↑](#footnote-ref-39)
40. 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，2011年第1版，第70页 [↑](#footnote-ref-40)
41. 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，2011年第1版，第71页 [↑](#footnote-ref-41)
42. 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，2011年第1版，第71页 [↑](#footnote-ref-42)
43. 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，2011年第1版，第71页 [↑](#footnote-ref-43)
44. 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，2011年第1版，第76页 [↑](#footnote-ref-44)
45. 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，2011年第1版，第78页 [↑](#footnote-ref-45)
46. 威尔·金里卡：《当代政治哲学》，刘莘译，上海：上海三联书店，2004年第1版，第411页 [↑](#footnote-ref-46)
47. 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，2011年第1版，第114页 [↑](#footnote-ref-47)
48. 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，2011年第1版，第165页 [↑](#footnote-ref-48)
49. 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，2011年第1版，第196页 [↑](#footnote-ref-49)
50. Catherine Galko Campbell. Persons Identity, and Political Theory A Defense of Rawlsian Political Identity. New York Dordrecht Heidelberg London: Springer,2014.21 [↑](#footnote-ref-50)
51. 威尔·金里卡：《当代政治哲学》，刘莘译，上海：上海三联书店，2004年第1版，第411页 [↑](#footnote-ref-51)
52. 龚群：《桑德尔对自由主义自我观的批评》，中山大学学报（社会科学版），2011年第3期 [↑](#footnote-ref-52)
53. 威尔·金里卡：《当代政治哲学》，刘莘译，上海：上海三联书店，2004年第1版，第415页 [↑](#footnote-ref-53)
54. 约翰·罗尔斯：《政治自由主义》，万俊人译，南京：译林出版社，2011年第1版，第22页 [↑](#footnote-ref-54)
55. 何怀宏：《正义理论导引——以罗尔斯为中心》，北京：北京师范大学出版社，2015年第1版，第7页 [↑](#footnote-ref-55)
56. 罗尔斯：《正义论（修订版）》，何怀宏、何包钢、廖申白译，北京：中国社会科学出版社，2009年第1版，第258页 [↑](#footnote-ref-56)
57. 罗尔斯：《正义论（修订版）》，何怀宏、何包钢、廖申白译，北京：中国社会科学出版社，2009年第1版，第58页 [↑](#footnote-ref-57)
58. 依据张国清《正义感是人的首要美德——罗尔斯双重理性假设之考察》中统一的译法 [↑](#footnote-ref-58)
59. 费雷曼：《罗尔斯》，张国清译，北京：华夏出版社，2013年第1版，第57页 [↑](#footnote-ref-59)
60. 张国清、汪远旺.《正义感是人的首要美德——罗尔斯双重理性假设之考察》.吉林大学社会科学学报，第56卷 第4期 [↑](#footnote-ref-60)
61. 张国清、汪远旺.《正义感是人的首要美德——罗尔斯双重理性假设之考察》.吉林大学社会科学学报，第56卷 第4期 [↑](#footnote-ref-61)
62. 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，2011年第1版，第76页 [↑](#footnote-ref-62)
63. 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，2011年第1版，第115页 [↑](#footnote-ref-63)