



首页 → 学术文章 → 应用伦理学基础理论

## 斯蒂芬·达沃尔：自利的契约论和非自利的契约论\*

自利的契约论和非自利的契约论\*

[美] 斯蒂芬·达沃尔 (Stephen Darwall)

陈真 (南京师范大学哲学系) 译

内容提要：契约论的基本思想是：道德应该仿效人们相互间所同意的互惠或地位平等的人之间的合作，或者预设于其中。一个行动是对是错必须取决于该行动是否符合或违反某些原则，这些原则是平等的人们之间达成的或者可能达成的合乎需要的契约的对象。自利的契约论认为，各方的平等状况仅仅是事实上的，各方对原则的选择也是以理性地自利为基础的。非自利的契约论则认为，合理的互惠或道德上平等的人们之间的公平性的一种理想。

关键词：契约论，自利的契约论，非自利的契约论

人们称之为“自利契约论的” (contractarian) 或“非自利契约论的” (contractualist) 规范伦理学理论源于这样一种观念——即道德应该仿效相互间所同意的互惠或地位平等的人之间的合作，或者预设于其中。这一思想认为，正是在契约的概念中包含了道德的思想，反过来也是一样。如此，一个行动是对是错必须取决于该行动是否符合或违反某些原则，这些原则是平等的人们之间达成的或者可能达成的合乎需要的契约的对象。这一总的思想可以通过不同的方式来发展，取决于一个人怎样表征相关的契约、契约的各方和他们的平等状况、以及契约得以达成的条件。一个总的区别就是自利的契约论 (contractarianism) 和非自利的契约论 (contractualism) 之间的区别，按照前者，各方的平等状况仅仅是事实上的 (*de facto*)，各方对原则的选择也是以理性地自利 (*rational self-interest*) 为基础的。而后的出发点则是合理的互惠或道德上平等的人们之间的公平性的一种理想。一个道德原则的有效性不管怎样都取决于某种选择或契约的想法似乎有点古怪。但只有极少的情况下，自利的契约论者或非自利的契约论者主张道德的标准取决于实际的契约。[1] 更多的情况下，他们主张道德原则是那些在某种理想的、虚拟的条件下人们可能理性地或合理地选择或同意的原则。

### 自利的契约论

自利契约论的经典表述源自托马斯·霍布斯的《利维坦》。[2] 霍布斯先考虑这样一种情景，一个行动者独立于他人进行思考，出发点是他自己的欲求或利益。他认为，每一个人将他自己所欲求的看成是好的，因此给了他理由以实现他所欲求的。但是，如果我们所有的人都一起追求我们各自的欲求和利益，那么会发生什么样的情况呢？虽然，在给定他人的行动的条件下，每个人追求他自己的利益其结果也许会使他们的实际利益得到最大的提升，但这并不意味着所有的人追求他们自己的利益，而不是按照某种非利己的原则行动，其实际的结果都会导致所有的人的（甚至任何人的）利益得到最大的提升。当情况并非如此，当对个人利益的集体追求导致各方更糟糕的情况时，我们就面临着所谓集体行动的问题。

这种情况在所谓“囚徒悖论”的游戏理论的例子中得到了很好的说明。在“囚徒悖论”中，两个人因怀疑抢劫被捕入狱。地方检察官分别告诉他们，她因缺少足够的证据而不能判他们的强盗罪，但她可以很容易地判他们非法进入他人住宅罪，有期徒刑一年。她分别给每人提出一个交易的条件：如果他们中一人坦白认罪，而同伙拒绝认罪，则认罪者无罪释放，同伙将判有期徒刑二十年。如果两人都认罪，则各判有期徒刑五年。

假定每个人所在乎的仅仅是尽可能少的有期徒刑，那么这一情景就会有如下的结构：如果甲认罪而乙拒不认罪，则甲会得到第一好的结果，而乙则会得到第四好的（最糟的）结果。反过来也是一样，如果乙认罪而甲不认罪，则结果就会反过来：甲得到第四好的结果，而乙得到第一好的结果。如果两人都认罪，则他们获得他们第三好的结果。如果两人都不认罪，两人则获得他们第二好的结果。

各人究竟应该怎样行动？让我们先从甲的立场进行推理。乙的行动将独立于甲，他或认罪或不认罪。因此，甲似乎应该认罪，因为无论乙做什么，甲如果认罪，他（甲）的情况都会更好。如果乙认罪，甲通过认罪可以得到第三而不是第四好的结果。如果乙不认罪，则甲通过认罪将得到最好的而不是第二好的结果。因此，甲应该认罪。不管乙做什么，甲通过认罪其情况都会更好。

但乙的情景和甲的情景一模一样。因而任何支持甲认罪的理由也同时适用于乙。因此，如果甲认罪其处境最好，那么对乙来说情况也同样如此。现在可以清楚地看到为什么这一问题叫做集体行动的问题。当甲和乙单独采取个人行动的时候，行动者都最有可能获得最好的结果，然而将他们的行动放在一起，其结果对每个人来说都更糟。给定另一个人的行动，如果每个人都采取对自己最有利的行动，则两人都会认罪。但其结果各人只获得第三好的结果，而如果他们都不认罪，他们本来可以取得第二好的结果。虽然牢房的语境听起来很奇怪，不认罪实际上是甲和乙合作的策略。如果甲和乙能够为了相互的利益而合作，他们就不会认罪，从而获得第二好的结果，而不是他们分别追求自己利益所获得的第三好的结果。

当人们放弃追求他们各自的利益并遵守规则或履行职责的时候，他们就合作。集体遵循规则比每个人独自追求自己的利益更好地提升了每个人的利益。显然，合作对生活中许多有价值的东西是必须的，或许在复杂的现代社会中尤其如此，因为在复杂的现代社会中我们不能假设共同文化或宗教传统所形成的真正的共同的利益将体现在重要交往活动的所有领域。道德可以被看作是特别广泛、无所不在的合作形式。因而道德上对错的原则可以理解为规则，具体的要求，许可，等等，它们构成最广泛的可能合作的基础，这一合作的方案涉及的不仅仅是这一或那一群体、社区或政治组织，而且包括所有的合格的人类的或理性的行动者。[3]

因此，按照自利的契约论，一个行动是正确还是错误是由这种最广泛意义上的合作的规则所决定的。举例来说，假定有这样一条规则：不帮助需要帮助的他人是错误的，只要所涉及的牺牲不要太大（比如，不要超过 $\alpha$ 的水平并且/或者牺牲和需要之比不要超过 $\beta$ 的水平）。由于我们都以各种各样的方式彼此依赖（这就是前面提到的事实上的平等），可以证明存在着水平 $\alpha$ 和 $\beta$ ，这样大家都遵守推导出来的规则将会比大家都独自追求自己的利益更好地提升每个人的利益。如果情况确实如此，自利契约论就会认为不遵守这条规则是错误的。

因此，初略地说，自利的契约论认为做什么样的事情才是正确的取决于大家在他们的思考和道德实践中接受什么样的规则，受什么样规则的指导才符合每个人的利益。对于一个给定的情景，大家都遵守上述规则比每一个人各自分别追求各自的利益能更好地提升每个人的利益，然而，如果各种各样不同的规则都具有这样的特征，我们又该如何呢？这里正是契约或合约的概念进入自利契约论画面的地方。将“无契约”的点看成一个基准点，在这个点上所有的人都认为自己除了他们自己的利益和价值标准外不受任何事情的约束，自利契约论者将我们实际上受哪些道德规则约束的问题看成是对从这一基准点出发的理性谈判的问题的解答。在这一解答中，我们都有更感兴趣赞同某些互利的原则，根据这些原则以避免“无契约”的点，但对究竟实际上赞同哪些原则，我们的兴趣各有不同。[4] 作为谈判结果的原则如何有利于不同的谈判者取决于如果没有契约谁的损失最大。比如，让我们考虑我们接受何种互助的原则。如果那些有着较少资源而又有更多弱点的人当不能达成契约时比那些有更多资源而较少弱点的人损失更大，那么，和每一个人都和那些具有较少资源的人一样脆弱的情形相比，理性的谈判也许会导致较少负担的互助原则。

因此，这就是自利契约论评价道德问题的基本框架。为了找出具体情况下我们的道德义务，我们必须考虑从“无契约”的点（即每个人都努力增加自己的个人利益并满足其价值标准的点）出发的谈判中会产生什么样的关于处理问题的原则的契约。不管怎样，重要的是要充分理解，在按照可能同意的原则采取实际行动时，契约各方不仅仅是提升他们的利益，他们还进行合作。合作有利于每一个人，但正如“囚徒悖论”所表明的，这种情况的实现要求个人限制提升他们自己的利益。每一个人都愿意选择他人承担更多必要牺牲的方案，但也准备根据大家理性上能够同意的合作原则履行自己的职责。

### 非自利的契约论

非自利的契约论有着类似的结构。它也将正确行为的原则看成是一个理性契约的对象。然而，尽管自利的契约论将道德原则看成是理性自利谈判的结果，非自利的契约论却将相关的契约看成是受辖于某种平等待人的道德理想，这种道德的理想确实会和按照自利契约论的方式就合作的基本条件进行谈判的思想不一致。从非自利的契约论的观点看，自利的契约论的问题是：它似乎假定个人对他们所控制的资源在没有达成合作规则的情况下事实上有道德的权利要求。与此不同，非自利的契约论者则认为在那种情况下通过谈判所达成的规则并没有道德的力量。然而，为什么要假定人们有这样的道德的权利要求？从道德的观点看，除非我们假定某种自然权利的背景理论，否则，这似乎完全是主观任意的。如果自利契约论自身的道德力量已经依赖上述假设，那么它不可能辩护这一假设而不陷于循环论证。[5]

非自利的契约论充满生气的思想可以从伊曼努尔·康德的绝对命令的“目的王国”的表述中找到。康德认为任何服从道德律的人都必须看作是能够“制定道德律”的人。[6] 只有这样，道德律才能作为一个自由的道德行动者共同体的共同法，而这些行动者只服从他们自己制定的法规。这是卢梭思想的不同说法，按照卢梭的思想，合乎法理的政治共同体是这样的一个团体，在这个团体中，每一个人“和所有的人行动一致，然而只服从自己”。[7] 根据卢梭的看法，一个自由平等的人的共同体是可能的仅当他们的法律表达了他所称之为“总意志”的东西，即每一个作为自由和平等的成员的意志。同样，康德将道德律理解为由每一个道德行动者所“制定”的，如果每一个人作为“目的王国”的自由和平等的成员将会对之“立法”。这里我们看到了和自利的契约论的一个主要的区别。正确性的道德原则不是个人从他们不同的个人的立场出发，经过基于个人利益的谈判，所规定的和试图获得承认的规则。毋宁说，它们是个人作为一个自由的和平等的人从一个共同的立场出发所规定（和同意）的规则。

然而，更具体地说这一立场应该如何理解呢？康德提供了某些启示，他说只要我们不仅“对理性者的私人目的的全部内容进行抽象，而且对他们个人的不同之处进行抽象”，我们就可以设想出一种“在共同的客观法规辖下的理性者的有组织的联合体”。[8] 这一想法导致了约翰·罗尔斯（John Rawls）的思想，即正义的原则就是站在对导致个人或社会相区别的任何特征一无所知的“无知之幕后”背后的“原初立场”上可以理性选择的原则。[9] 特别是，选择的各方必须对他们个人的财产、能力、天赋、性别、种族、社会经济地位、以及他们自己的利益或个人的价值观念一无所知。罗尔斯试图证明，虽然如此，按照这样的方式理解，我们都可以假定选择的

各方能够有兴趣自主地选择和追求他们的利益（不管这些利益最终表现出来的结果如何）。他们尊重实现这些利益所必须的“首先想望的东西”：自由，机会，财富，以及自尊的“社会基础”。

这样，按照罗尔斯的思想，正义是由从原初的立场上，即作为和其他人一样自由和平等的人，所理性选择的原则所决定的。罗尔斯确实假定这一理性的选择在无知之幕后限定的范围内是自利的，但这并不能以任何方式将道德的推理还原为个人的利益。假定选择的各方不是受个人利益所驱动，而是受对其他的个人的关心所驱动。由于无知之幕后使他们无法知道有关信息从而使他们能够修改原则以符合任何特殊个人的利益，因此，从功能上讲，假定各方是自利的原则所驱动和假定他们是其他个人的受托者，两者之间没有任何区别。原初的立场实际上就是一个，即任意一个，自由和平等的个人的立场。

罗尔斯将他的思想作为一种正义——即他所称之为“作为公平的正义”——的理论提出来的。最近他又强调它应该作为一种政治的而不是一种更为一般的道德理论来理解。[10]但在他的早期的著作中，他认为它也可以看作是一种道德的理论：“作为公平的正确性”。[11]

第二种非自利的契约论的探讨方法起因于对一个平等的人提出权利要求意味着什么的考虑。当一个人对另一个人以这样的方式提出对某件事情的要求时，她试图给此人一个理由，该理由建立在作为同等地位的人之所需的基础上。她仿佛在说，“这是一个合理的你所能授予我的权利要求，正像当你处于我的地位时，当你认为你可以对我提出同样合理的要求时，你所能看到的那样。”这样的要求隐含互相承认的平等的人所能接受的，至少无法合理拒斥的行为规则的思想。T·M·斯坎伦发展了这样一种非自利的契约论的探讨方法，他设想了一种共同体，其成员希望能够通过其他成员无法合理反驳的原则对自己的行为进行辩护，而这些其他的成员也有着同样的目的。[12]这样，道德对错的原则就可以被看作是建构互相负责的平等人的共同体的规范。

为了应用这一标准，我们必须对什么是合理的做出判断。我们如何才能做到这一点呢？除了替别人设身处地，除了看看我们是否可以将某一个要求或者针对某个规则的反对意见看作是当我们处于别人的情景中可以合理的提出的要求或意见外，似乎别无其他办法。这是一个复杂的判断。它不是简单地预测我们将会提出同样的要求或同样的反对意见。我们也许认为我们会这样做，但是这样做可能是不合理的。为了做出一个必不可少的判断，我们和任何人一样，似乎必须尽量不偏不倚地站在他人的立场上看问题，看看我们是否会赞同对另一个相互承认的平等的人可以合理提出的要求或反对意见。

#### 注释

\* 本文译自Stephen Darwall为他所编辑的*Contractarianism/Contractualism*（《自利的契约论和非自利的契约论》）（UK Oxford: Blackwell, 2003）一书所写的导言（见该书第1~8页），标题为译者所加。关于Contractarianism和Contractualism的译法，译者曾和Thomas Scanlon，霍桂恒和甘绍平交换过意见，特别是Scanlon的看法，使译者最终选定了现在的译法，特此表示谢意。由于该导言的最后一个自然段主要是对该书所收文选编排的说明，和全文主题无关，故略去未译，原文中除书名以外的斜体一律以黑体字表示，特此说明。——译者。

[1] 吉尔伯特·哈曼（Gilbert Harman）是一个例外。参见他的“Moral Relativism Defended”（“道德相对论的辩护”），*The Philosophical Review* 84（1975）：3~22，和本书选自哈曼的文章。

[2] 最初发表于1651年。关于霍布斯的伦理哲学，参见我的*Philosophical Ethics*（《哲学伦理学》）（Boulder, CO: Westview Press, 1998）。

[3] 霍布斯的观点认为大规模群体之间的合作没有政治权威是不可能的，因为制裁的威胁对使每一个人确信其他人在合作中将会履行他们的职责是必要的。然而，他的意思并不是说，制裁的威胁也给了每一个人遵守互利规则的理由。如果人们为了避免受到制裁而放弃在其他情况下会提升他们个人的利益的为的话，他们不是在合作，而仅仅是将制裁的因素考虑进提升他们利益的过程中。霍布斯的观点认为，当合作的方案按照我们所描述的方式是互利的时候，人们就有理由合作，也就是说，有理由放弃他们自己的个人利益。然而，他认为合作的合理性取决于是否能保证他人也会合作。制裁所要做的就是给每一个人合理的确信，确信他人将会履行他们的职责。因此，霍布斯认为重要的是每一个人都有理由实际遵守对他们自己不利的规则（如同“囚徒悖论”中所表明的那样），条件是他们能够确信他人遵守规则并且这一事实的存在并不以任何方式依赖于政治权威。我们需要政治机构是为了确信通过确保这一事实所假定的理由存在的条件，给予这一事实以实际的效果。

[4] 这一观点最著名的表述见大卫·哥梯尔（David Gauthier）的著作。由于霍布斯认为为了使道德规范发挥实际的效果，绝对的政治权威是必须的，因此在他的理论中没有类似的谈判的作用。正是君权的意志决定了哪些互利规则的更为详尽的说明具有道德的力量。

[5] 如果我们考虑个人怎样从自利的实用的推理过渡到自利契约论的推理，这一问题就会来自另一个方面。如果每一个行动者按照她自己独自的利益推理，那么，在没有建立合作规则的情景中，她会认为通过和他人谈判达成协议，受某种规则的约束是合乎理性的。但是，这如何能够给她理由实际遵守这些规则呢？她达成契约遵守这些规则的基于利益的理由并不能给她实际遵守它们的理由，因为合作规则的全部要点和功能要求它们限制她追求她自己的利益。为了使她能够按照这些规则所要求的那样进行推理，她必须已经接受合作的道德的理由。也许接受自利契约论的合作的道德理由甚至是符合她的利益的，然而，尽管这有可能给她理由想接受道德的理由，但她不可能因为这些理由而接受道德的理由。

[6] Immanuel Kant, *Groundwork of the Metaphysics of Morals*（《道德形而上学的基础》）（最初发表于1785年，Ak., p. 431（in the standard [German] edition of Kant's works, published by the Prussian [later German] Academie; the canonical reference is Ak., p. 431; p. 82 in this volume). 关于康德的伦理哲学，参见我的*Philosophical Ethics*（Boulder, CO: Westview Press, 1998）。

[7] Jean-Jacques Rousseau, *The Social Contract and Other Later Political Writings* (《社会契约和其他晚期的政治著作》), Victor Gourevitch, ed. (Cambridge: Cambridge University Press, 1997), p. 49 (p. 63 in this volume).

[8] Kant, *Groundwork*, Ak., p. 433 (p. 84 in this volume).

[9] John Rawls, *A Theory of Justice* (《正义论》) (Cambridge, MA: Belknap Press of Harvard University Press, 1971), see esp. pp. 136~42 (pp. 183~8 in this volume).

[10] John Rawls, *Political Liberalism* (《政治自由主义》) (New York: Columbia University Press, 1993).

[11] Rawls, *A Theory of Justice*, p. 111.

[12] 除了本书所收入的文章外, 参见T. M. Scanlon, *What We Owe to Each Other* (《我们相互间的责任》) (Cambridge, MA: Belknap Press of Harvard University Press, 1998), 尤其是第147至257页。

(本译文发表在《世界哲学》2005年第4期。译者现在将“囚徒悖论”译为“囚徒困境”。)

中国社会科学院应用伦理研究中心

北京建国门内大街5号 邮政编码: 100732 电话与传真: 0086-10-85195511

电子信箱: cassethics@yahoo.com.cn