

Globethics Repository

The logo for Globethics, featuring the word "Globethics" in white, sans-serif font centered within a solid blue rectangular background.

评罗尔斯的规则功利主义惩罚思想[Comments on the rule utilitarianism from Rawls]

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	Article
Authors	王, 立峰
Publisher	国家行政学院
Rights	With permission of the license/copyright holder
Download date	2026-07-10 19:02:42
Link to Item	http://hdl.handle.net/20.500.12424/179195

王立峰：评罗尔斯的规则功利主义惩罚思想

摘要：在人类历史的大多数时间里，对法律惩罚正当性的证明源自两种理论分野，一是主张正义和应得的报应主义，一是主张社会利益最大化的功利主义。在《规则的两个概念》一文中，罗尔斯提出了一种关于法律惩罚的混合理论，他首先对惩罚的制度层面和实践层面进行区分，在此基础上分别从两个层面对惩罚进行证明，认为惩罚的制度层面可以用功利主义进行证明，实践层面可以用报应主义进行证明。这样他就调和报应主义与功利主义在惩罚证明上的冲突。但是，由于规则功利主义自身的缺陷，也由于罗尔斯本人对传统报应主义、功利主义认识的局限以及对法官和立法者之间界限的划分的理想化，这种调和的努力被证明是失败的。

关键词：惩罚、报应主义、规则功利主义、功利主义

Abstract: In the long history of humankind, Utilitarian and retributive theories of punishment provide different reasons for justifying legal punishment. Retributivists claim justice and desert, utilitarians claim maximization of social utility. In “ Two concepts of Rules ”, Rawls put forward a mixed theory to reconcile the utilitarian and retributive theories. He distinguishes between the justification of an institution and the justification of practice. Rawls points out that utilitarianism justify the punishment as an institution and retributivism justify punishment as a particular case. But because of the weakness of rule utilitarianism, the limits of Rawls ’ understandings of traditional utilitarian and retributive theories, idealization of division between legislator and judge, the reconciliation failed.

Keywords: punishment, retributivism, rule utilitarianism, utilitarianism.

一、引言

长久以来，法律惩罚的正当性根据何在？换句话说，国家惩罚一个罪犯的依据是什么，我们为什么惩罚？这个问题令法学家和哲学家们冥思苦想而仍困惑不解。在1955年“规则的两个概念”一文中，罗尔斯写道：惩罚问题一直是一个令人困惑的道德问题。关于惩罚的困惑不在于人们在惩罚是否具有正当性上持有不同意见……很少有人完全拒绝惩罚……困难在于如何证明惩罚的正当性：道德哲学家们为此进行了各种各样的争论，提出了各种各样的理论，但没有一种理论获得普遍接受，没有一种理论能够远离嫌恶。[i]

罗尔斯为什么对惩罚的正当性问题感到如此困惑呢？

在人类历史的大多数时间里，对法律惩罚正当性的证明源自两种理论分野，一是主张正义和应得的报应主义，一是主张社会利益最大化的功利主义。两种理论看起来水火不容、誓不两立。功利主义把报应主义描述成一种非理性的直觉反应，是复仇情感的发泄，因此否认它的道德地位。早在古希腊时期，在这场争论伊始，柏拉图就指出：

在惩罚行为不当者的时候，除非是像野兽一样盲目复仇，人们并不关心这个行为不当者过去所犯的错误的，或者依据它过去所犯错误对它施加惩罚。一个理性的人不能根据行为不当者过去所犯罪行来施加惩罚，毕竟覆水难收，他所要考虑的是未来，是如何防止行为不当者再次犯罪，或者通过施加惩罚的场面，防止其他人犯同样的错误。[ii]

报应主义则指出，功利主义因为忽视了正义原则，不管功利主义的惩罚对社会多么有益，那么结果是导致各种各样不公正的惩罚。维斯特马克在《道德理念的起源与发展》一书中对功利主义进行了批评，他说：“那些试图实现功利主义所要追求的所有惩罚效果的人，不仅在功利主义理论的反对者看来，而且可能在功利主义理论的坚决拥护者看来，他们比惩罚的罪犯更具有罪恶性。”[iii]

正是在报应主义与功利主义各执一端的争论里，包括罗尔斯在内的很多学者试图化解这种困惑。他们把建构一个规范性的惩罚理论视为自己的使命。这种规范性理论能够包纳一切可能性，但这种规范性理论是否可能呢？在功利主义与报应主义思想之外，一些学者提出了自己的见解，认为，报应主义与功利主义各有一定道理，一个关于惩罚的正当性的理论应是两种理论的混合版，即把二者综合起来，在两种理论之间搭建一座桥梁，把强调过去与强调现在、强调正义与强调共同体的善的主张加以调和。显然，这些理论尝试是一种中间路线，它们把两种理论之间的巨大差异看作是劳动分工的不同。这种混合理论的代表人物是罗尔斯。罗尔斯对法律惩罚问题的认识主要体现在《规则的两个概念》一文中。

二、罗尔斯在《规则的两个概念》中的认识

《规则的两个概念》一文发表于1955年第1期的《哲学评论》。这篇文章一经问世，就引起了学界的极大关注。这不仅因为罗尔斯在这篇文章中提出了一种关于法律惩罚的混合理论，而且他的证明思路也大大不同于以往学者们的研究进路，他的思路是首先对惩罚的制度层面和实践层面进行区分，在此基础上分别从两个层面对惩罚进行证明。

什么是惩罚（制度层面）？罗尔斯认为，一个人不论什么时候被依法剥夺了作为一个公民所享有的一些正常权利，那他就是接受了惩罚；当然惩罚的前提是他触犯了法治，这种触犯要依据法律的正当程序、通过审判来确认，而对权利的剥夺要由国家公认的法律权威机构来实施；同时，法治要明确说明什么是犯法行为，对这种犯法行为应该适用什么刑罚；法院要严格解释法律，法律的公布实施在时间上要先于犯法行为的发生。[iv]

显然，罗尔斯所说的法律惩罚外延很广泛，泛指由国家根据法律正当程序而施加于罪犯身上的所有可能的处罚，相当于刑事惩罚。这种惩罚不包括学校对学生、家长对子女实施的具有教育意义的惩罚。

罗尔斯对传统的惩罚理论是如何认识的呢？对于报应主义，罗尔斯认为，在报应主义者看来，“惩罚的正当性根据在于恶行应当受到惩罚。一个做了坏事的人受到相应的惩罚，这在道德上是恰当的。一个刑事罪犯在犯罪后就应当给予惩罚，刑罚的轻重取决于他犯罪行为的邪恶程度。”[v]

那什么是功利主义的刑罚观呢？罗尔斯认为，所谓功利主义的刑罚观是指：“过去的就让它过去了。对于当下的决定而言，只有未来的结果才是重要的。惩罚正当性的证明与它所能带来的可能结果有关，结果是惩罚可以作为维持社会秩序的一种机制。...如果惩罚能够有效促进社会的最大利益，那它就是正当的，反之，它就是不正当的。” [vi]

需要注意的是，罗尔斯所讲的报应主义强调犯罪行为的恶性，并根据恶性决定刑罚的轻重；而一般报应主义关注罪犯的犯罪行为给受害人所造成的伤害，并根据伤害、罪犯的责任和应得性决定一种比例惩罚。

显然，在罗尔斯看来，报应主义与功利主义之间存在冲突。报应主义强调只有犯罪才应该得到惩罚，而功利主义认为，是否惩罚则取决于惩罚能否促进社会的利益。两者之间的冲突突出体现在如何对待无辜者上。根据功利主义，如果惩罚一个无辜者会带来社会的稳定，那无辜者就可以成为国家利用的手段，国家可以惩罚无辜者；在报应主义看来，无辜者绝对不能受到惩罚，因为他没有犯罪，也就不应该得到惩罚。

那么，如何解决报应主义与功利主义之间的冲突呢？罗尔斯认为，在惩罚的正当性证明问题上，两种理论事实上回答了不同的问题，确切说，回答了不同层面的问题。在“规则的两个概念”一文中，罗尔斯写道：“必须在所应用和实施的规则体系与该规则体系指导下的具体行为之间做出区分。对于规则而言，功利主义的考虑是合适的；对于具体规则在具体个案中的适用而言，报应主义的考虑是合适的。” [vii]也就是说，功利主义是从惩罚制度本身出发来证明惩罚的正当性；而报应主义则从具体的惩罚个案出发来证明惩罚的正当性，比如说，对一个具体犯罪行为应该根据罪行的轻重而施加相应程度的刑罚。

罗尔斯对此作了进一步的解释，认为立法者与法官在惩罚正当性的证明中扮演了不同的角色。也就是说，制定惩罚制度的立法者，考虑的是共同体的善；把惩罚制度的规则适用到具体个案中的法官，依据法律规定和犯罪人的罪行做出惩罚，并宣告无辜者无罪。他说：

法官和立法者站在不同的立场，从不同的角度看待惩罚：一个关注的是过去，一个关注的是将来。法官本人对惩罚的证明似乎像报应主义的观点；立法者本身对惩罚的证明似乎像功利主义的观点。两种观点都有一个情境.....那么，一旦人们发现这些观点是承担着不同职责、在不同的办公室里的人们，在不同的情境下谈论刑法（刑法是由规则体系组成）时，最初的迷惑就会得到澄清.....，把两种观点适用于不同的情境，是个由来已久的策略，根据这个策略，就会对两种观点加以调和。 [viii]

可见，在罗尔斯看来，法官和立法者站在不同的立场上，用不同的眼光审视问题：一个关心的是过去，另一个关心的是未来。法官所做的，就其本身来讲，可以用报应主义的观点来证明是正当的；立法者所做的，就其本身来讲，可以用功利主义的观点来证明是正当的。

就这样，依据解决不同层面的问题，适用于不同的场域，或者说，在不同的时间段上，罗尔斯把报应主义与功利主义两种截然不同的惩罚理论进行了调和。显然，根据罗尔斯的观点，比之具体个案的法律适用，法律是主要的；比之法官的作用，立法者的作用是主要的。相应地，由于把报

应主义放在次要的位置上，这种混合理论从根本上说是功利主义的。

罗尔斯的观点值得商榷。

首先，罗尔斯对报应主义所做的解释是站不住脚的。他认为，报应主义并不必然关心作为制度的惩罚的正当性问题，报应主义关心的是作为个案的具体惩罚的正当性。他写道：“作为一种制度，法律机器的根本目的在于建立和保持道德恶行与其造成伤害之间的相当性，一个报应主义者会这样主张吗？当然不。”^[ix]但是，报应主义者却确实在关注作为制度的惩罚的正当性问题，并非对此视而不见。事实上，报应主义对惩罚的证明是在总体上进行的，而非仅在具体个案的层面上进行。如果作为制度的惩罚缺乏道德基础的话，报应主义怎么会断言惩罚的轻重应该与受害人给受害人造成的伤害成比例呢？也就是说，为什么还会主张罪刑相当呢？依据报应主义，惩罚之所以得到证明，不仅在于一个罪行应该根据其罪行的大小得到相应的惩罚，而且，作为制度的惩罚的正当性在于它所具有的道德基础，也就是说作为惩罚的制度在逻辑上与社会道德原则是一致的，而这些社会道德原则并没有把威慑或者未来效果作为惩罚制度的首要原则。这样看来，罗尔斯断言报应主义者忽视了对制度层面惩罚的证明问题，是有问题的。

其次，罗尔斯对功利主义的认识也是有问题的。他说：“在惩罚只能施加于法律的触犯者这点上，功利主义者是持赞同态度的。”^[x]其言下之意，报应主义者认为，一个受到惩罚的人，必须要有罪过，功利主义者对报应主义者的这个观点也是持相同立场的。另外，在他看来，功利主义者会认为，不管一个官员是否相信惩罚会带来良好社会效果，他都不应该拥有实施惩罚的自由裁量权力。^[xi]当然，这是规则功利主义的观点而不是行为功利主义的观点。在一个奉行规则功利主义的社会，罗尔斯的主张应该会如愿以偿的。这样一来，那种认为功利主义会在特定情形下惩罚无辜者的反对意见就不攻自破了。因为按照罗尔斯的逻辑，功利主义不能为了社会福利的最大化而侵犯无辜者的权利。进一步说，在罗尔斯看来，从长远看，侵犯无辜者的权利并不会带来好的社会效果。如他所说：“如果一个人把惩罚制度的实际运作描述成惩罚无辜者的制度，那这种制度就蕴含着巨大的风险，显然，这种制度不会促成有益目的的实现。”^[xii]而事实上，功利主义者对惩罚不会作出如罗尔斯这般的证明。

第三，罗尔斯把法官与立法者之间界限的划分过于理想化了。事实上，法官在量刑时有相当的自由裁量权力，特别是在英美法系国家，法官甚至可以造法，其自由裁量权力是相当大的。这么看，立法者与法官之间的界限是很模糊的，而不像罗尔斯想像的那样，界限清晰，法官负责向后看，立法者负责向前看。司法实践表明，法官在适用惩罚时不仅向后看，也向前看。^[xiii]需要注意的是，当法官给犯罪人处刑时，法官一般拥有相当的裁量自由和回旋余地。这样法官和立法者之间的界限就是模糊的。

综上，罗尔斯试图在报应主义与功利主义之间作出一种调和，他做了，但是这种调和并不成功。但是，罗尔斯对制度与实践的划分非常重要，对于惩罚问题的研究富有启发意义，其后的哈特以及芬等学者都顺延了他的这种论证思路。^[xiv]

三、对规则功利主义惩罚观的进一步分析

罗尔斯在1971年出版《正义论》，标明他是康德道义论的继承者。他在《正义论》一书的开篇之

初就指出，正义是社会制度的首要价值，某些法律和制度，不管它们如何有效率和有条理，只要它们不正义，就必须加以改造或废除。[xv]可见，正当对善具有优先性。这个规则同样适用于惩罚制度的设置，即正义对效率和福利的优先。但在《规则的两个概念》一文中，在对刑罚问题的认识上，他是一个功利主义者，确切说，是规则功利主义者。

如前所述，罗尔斯试图对功利主义和报应主义进行调和，证明二者在惩罚理论中都能找到自己的立足点。罗尔斯的调和建基于作为制度的惩罚与作为实践的惩罚。他认为，惩罚制度，包括对一个具体犯罪如何适用刑罚、量刑的规则，可以用功利主义证明是正当的；对一个具体犯罪如何量刑的规则可以用报应主义证明是正当的。例如，法官对一个有罪的盗窃犯处刑，他只需要严格依据法律规定来确定具体刑罚，而根本不考虑这个刑罚是否具有威慑犯罪的效果。法官判处刑罚的根据在于被告人进行了犯罪行为。另一方面，立法者考虑的是，法律规定的刑罚能否取得威慑盗窃犯罪的效果。那么，按照罗尔斯的观点，既然功利主义理论适用于惩罚制度而不是具体的处刑，那么对功利主义导致可能惩罚无辜者的批评就会不攻自破。因为如果一个惩罚制度允许法官惩罚无辜者，那么这个制度就不是功利主义，不是为了谋求社会的利益。因为对无辜者的惩罚会极大伤害社会和法律的权威，无益于社会。

行为功利主义认为，当且仅当一个行为与其它行为相比产生最好的结果时，这个行为才是可证的。这种原则适用于每一个行为。规则功利主义则要求，每一个行为都应遵循最佳规则。这个最佳规则是对功利主义的最好证明。例如，在对待惩罚无辜者问题上，行为功利主义认为，一个行为只要能带来好的结果就是可证的；那么，因为惩罚无辜者能够威慑犯罪（一般威慑），所以惩罚无辜者是可证的。但规则功利主义者坚持认为，一般情况下，因为惩罚无辜者违反了最佳规则，那么惩罚无辜者就是错误的。

行为功利主义者也承认，由于多种多样的原因，行为人并不能总是计算出每一个行为的结果，这样人的行为就不能不依赖于一般规则。经验告诉我们，在大多数情况下，遵循规则如诚实守信，可以通过行为功利主义来证明，因为遵循规则会带来好的结果。久而久之，如果遵循规则总是带来好的结果，行为功利主义者就会遵循规则行事。但这仍然是行为功利主义，不是规则功利主义。行为功利主义的一个缺陷在于，在日常生活中，我们没有能力或者没有足够的时间来计算某个行为的后果，或者我们在计算时发生错误，那么，按照行为功利主义行事并不能总是带来好的结果。为了克服这个缺陷，行为功利主义就需要修正，求助于规则。

于是，规则功利主义对传统功利主义惩罚理论做出修正，其动机是避免行为功利主义理论的困难，在道德思考中考虑道德规则的作用。显然，行为功利主义惩罚理论会导致各种各样的不公正惩罚，甚至惩罚无辜者。规则功利主义惩罚理论试图通过对报应主义和功利主义做出综合，从而超越两种传统理论。这种理论虽然对正义和应得也给予了考虑，但从根本上讲是功利主义的。如前文所述，罗尔斯通过在惩罚制度与惩罚的具体个案之间做出区分，在立法者和法官之间做出区分，使规则功利主义成为一条可能的思考进路。

那么，这种规则功利主义是否成功的调和了报应主义与功利主义呢？要回答这个问题，就看这个理论能否成功回答功利主义惩罚无辜者的问题。在惩罚无辜者问题上，如果即使惩罚无辜者能实现法律所坚持的功利主义目的和预防犯罪效果，而规则功利主义又坚持只有罪犯才应被惩罚的规则，那么，规则功利主义是否解决了行为功利主义惩罚无辜者这个问题呢？一个规则功利主义者

可能提出如下解答：

（一）从惩罚无辜者的长远效果的角度看，规则功利主义会认为它已经排除了惩罚无辜者的可能。尽管惩罚一个无辜者会实现行为功利主义的预防犯罪目的，如一般预防。但是，如果一旦惩罚无辜者的真相最终大白于天下，这种不公正的惩罚势必会损公众的法律信仰，会极大的损害整个刑罚制度。虽然对无辜者的惩罚在当时会起到威慑犯罪、稳定社会秩序的目的，但从长远看，这种惩罚的危害大于其所取得的好处，弊大于利。因此，从长远看，对无辜者的惩罚不会实现社会效益最大化。因此，规则功利主义反对惩罚无辜者，排除了惩罚无辜者的可能。

但是，顺着上述思路，如果假定对无辜者的惩罚这个事实永远不会被公众所知晓，那么，法律的权威就不会受到破坏，对无辜者的惩罚就会实现行为功利主义的威慑犯罪目的而不会产生负面效果。这就意味着规则功利主义仍然容许惩罚无辜者这种可能性的存在。

（二）从法官严格遵循规则这点看，规则功利主义认为不会出现惩罚无辜者的可能。行为功利主义认为，法官在审理具体案件的时候，有权根据功利主义的目的决定案件的裁判结果，那就可能导致对无辜者的惩罚；也就是说，如果法官有这种依据功利主义原则进行自由裁量的权力，那么在逻辑上就会与有罪必罚这个规则发生冲突。规则功利主义者可能说，犯罪与惩罚之间的有着内在的逻辑联系，只有犯罪，才应受到惩罚，这是一个制度性规则，它不允许法官在定罪量刑时根据功利主义的惩罚目的进行自由裁量，不允许在规则外存在例外情形。规则功利主义要求法官在所有案件审理中，都要严格遵循规则，因此，一个人如果没有犯罪，他就不应受到惩罚。

这种解释缺少说服力。大家都知道，只有犯罪人应受到惩罚的规则是一个制度性规则，它是普遍有效的，法官没有权力破坏这个规则。但是，当法官面对一个案件，如果惩罚无辜者会带来最好的社会效果的时候，法官就没有理由不放弃这个规则。也就是说，法官要考虑只有犯罪人应受到惩罚的规则，如果有更好的功利主义的理由，他就可以放弃这个规则。当然，在特殊情形下，法官就可以惩罚无辜者。

罗尔斯也认识到，一个允许惩罚无辜者的刑罚制度，尽管惩罚无辜者会实现社会的最大利益，但功利主义不可能为此提供证明。为什么呢？罗尔斯的理由是，“这个制度下的官员滥用权力的危险太大”。^[xvi]显然，这个理由并不能令人信服，因为任何制度都存在被滥用的可能，根本不存在不被滥用的制度，我们只能假定，滥用被减少到可被公众接受的最低限度。

（三）功利主义的立法虽然证明某些规则是正当的，但这些规则仍然容易导致产生惩罚无辜者的情形。假定一条法律规则明确规定，对于一个具体犯罪行为应该惩罚犯罪人的孩子，显然这种规定会取得很好的威慑效果。绝大多数父母如果知道威慑的结果是惩罚自己的孩子，那他们会被威慑而不进行犯罪行为。在功利主义看来，把邪恶的父亲的罪恶转移到无辜的儿子身上，似乎没有什么不当。那么，规则功利主义的惩罚理论就允许以威慑的名义惩罚无辜者。

可见，规则功利主义并不能彻底解决传统功利主义惩罚理论所面临的难题。

[i] See John Rawls, "Two concepts of Rules," in Hugh LaFollette, Editor, *Ethics in Practice: An Anthology*, Blackwell Publishers Ltd. 1997.

[ii] Plato, Protagoras, Clarendon Press: Oxford, 1982, p16.

[iii] E. Westermarck, The Origin and Development of the Moral Ideas, London, Second edition, 1912, Vol.1, pp.81-82.

[iv] See John Rawls, “ Two concepts of Rules, ” in Hugh LaFollette, Editor, Ethics in Practice: An Anthology, Blackwell Publishers Ltd. 1997.

[v] 同上。

[vi] 同上。

[vii] 同上。

[viii] 同上。

[ix] 同上。

[x] 同上。

[xi] 同上。

[xii] 同上。

[xiii] See Punishing Dangerousness: Cloaking Preventive Detention as Criminal Justice, Harvard Law Review, Vol.114, No.5(2001).

[xiv] 本在一篇文章中采取了类似的区分，见I. Benn, “ An Approach to the Problems of Punishment, ” Philosophy, 33(1958), pp.325-6. 哈特也区分了惩罚的一般证明目的和分配问题，他似乎也赞同罗尔斯的这种区分，根据在于他认为：“如果主张惩罚实践的一般证明目的在于它的良好社会结果，同时对这种一般目的的追求应该以分配原则加以限制，即只有违法者因其违法行为才能受到惩罚，而且承认一般证明目的与分配原则是完全一致的，那么就会避免功利主义与其反对者之间的很多令人迷惑的争论。”见Hart, H.L.A. punishment and responsibility (Oxford: Oxford University Press, 1968), p9.

[xv] 约翰·罗尔斯著，何怀宏等译：《正义论》，中国社会科学出版社，1988年版，第1页。

[xvi] See John Rawls, “ Two concepts of Rules, ” in Hugh LaFollette, Editor, Ethics in Practice: An Anthology, Blackwell Publishers Ltd. 1997.

