

订阅DeepL Pro以编辑此演示文稿。  
访问[www.DeepL.com/pro](https://www.deepl.com/pro?cta=edit-document)，了解更多信息。

规则后果主义

*2003 年 12 月 31 日（周三）首次发布；2023 年 1 月 15 日（周日）进行了实质性修订*

我们可以称之为完全的规则后果论的道德理论，完全根据规则后果的善恶来选择规则，然后声称这些规则决定了哪些行为在道德上是错误的。乔治-伯克利可以说是第一个规则后果论者。他写道："在制定一般自然法则时，我们当然必须完全以人类的公共利益为指导，但在我们生活中的普通道德行为中却不是这样。......规则的制定是为了人类的利益；但我们的实践必须始终紧随规则"（伯克利 1712 年：第 31 节）。

* [1.功利主义](https://plato.stanford.edu/entries/consequentialism-rule/" \l "Uti)
* [2.福利](https://plato.stanford.edu/entries/consequentialism-rule/" \l "Wel)
* [3.其他待促销商品](https://plato.stanford.edu/entries/consequentialism-rule/" \l "OthGooPro)
* [4.完全规则后果主义](https://plato.stanford.edu/entries/consequentialism-rule/" \l "FulVerParRulCon)
* [5.全球后果论](https://plato.stanford.edu/entries/consequentialism-rule/" \l "GloCon)
* [6.提出完全规则后果主义](https://plato.stanford.edu/entries/consequentialism-rule/" \l "ForFulRulCon)
  + [6.1 实际货物与预期货物](https://plato.stanford.edu/entries/consequentialism-rule/" \l "ActVerExpGoo)
  + [6.2 合规和验收](https://plato.stanford.edu/entries/consequentialism-rule/" \l "ComAcc)
  + [6.3 完全验收与不完全验收](https://plato.stanford.edu/entries/consequentialism-rule/" \l "ComAccVerIncAcc)
* [7.规则后果论的三种论证方法](https://plato.stanford.edu/entries/consequentialism-rule/" \l "TwoWayArgForRulCon)
* [8.规则后果论是否必须为崩溃、不连贯或规则崇拜负责？](https://plato.stanford.edu/entries/consequentialism-rule/" \l "OldObjRulCon)
* [9.对规则后果论的其他反对意见](https://plato.stanford.edu/entries/consequentialism-rule/" \l "NewObjRulCon)
* [参考书目](https://plato.stanford.edu/entries/consequentialism-rule/" \l "Bib)
* [学术工具](https://plato.stanford.edu/entries/consequentialism-rule/" \l "Aca)
* [其他互联网资源](https://plato.stanford.edu/entries/consequentialism-rule/" \l "Oth)
* [相关条目](https://plato.stanford.edu/entries/consequentialism-rule/" \l "Rel)

1.功利主义

当且仅当一种道德理论仅从后果的善恶角度来评估行为和/或性格特征、做法和制度时，它就是后果论的一种形式。在历史上，功利主义是最著名的结果论形式。功利主义只从整体净利益的角度来评估行为和/或性格特征、做法和制度。总体净收益通常被称为总体福祉或福利。总体福利的计算方法是，将任何一个人获得的利益或受到的伤害与任何其他人获得的相同大小的利益或受到的相同伤害相等，然后将所有利益和伤害相加，得出一个总和。后果论者对福利的最佳解释存在很大争议。

2.福利

古典实用主义者（即杰里米-边沁、J.S. 密尔和亨利-西奇威克）认为，利益和损害纯粹是快乐和痛苦的问题。这种认为福利是快乐减去痛苦的观点一般被称为享乐主义。这种观点已日趋成熟（帕菲特，1984 年：附录 I；Sumner 1996；Crisp 2006；de Lazari-Radek 和 Singer 2014：第 9 章），但仍然坚持这样一个论点，即一个人生活的好坏*完全*取决于他或她的快乐减去痛苦，尽管快乐和痛苦的解释非常宽泛。

即使对快乐和痛苦的理解非常宽泛，享乐主义也会遇到困难。主要的困难在于，许多人（如果不是所有人的话）都非常关心自己的快乐和痛苦之外的其他事情。当然，这些其他事物作为获得快乐和避免痛苦的手段可能是重要的。但是，许多人对享乐主义工具价值之外的东西非常关心。例如，许多人想知道各种事情的真相，即使这不会增加他们（或其他人）的快乐。另一个例子是，许多人对成就的关注超出了这些成就可能带来的快乐。同样，许多人以非工具性的方式关心家人和朋友的福祉。关于这些观点，尤其是最后一点，有一种对立的说法是，人们关心的不仅仅是自己的福利，还有许多其他的东西。

根据任何可信的福利观，当人们的欲望得到满足时，他们所能感受到的满足感会增加他们的福利。同样，根据任何一种合理的观点，由于欲望得不到满足而产生的挫败感也会造成福利的减少。有争议的是，除了满足某人的欲望对其满足感或挫败感的影响之外，满足该人的欲望是否会给该人带来好处。享乐主义的回答是 "否"，声称只有对满足感或挫败感的影响才是重要的。

另一种福利理论的回答是 "是"。该理论认为，行为人任何欲望的实现都会给行为人带来好处，即使行为人永远不知道该欲望已经实现，即使行为人没有从欲望的实现中获得任何快乐。这种人类福利理论通常被称为*福利的欲望实现理论*。

显然，福利的欲望满足理论比享乐主义更为宽泛，因为欲望满足理论承认，能够构成福利的东西比仅仅是快乐更为宽泛。但是，我们也有理由认为这种更宽泛的理论过于宽泛。首先，人们可能会有与自身生活脱节的合理欲望，而这些欲望与他们自身的福利无关（Williams 1973a: 262; Overvold 1980, 1982; Parfit 1984: 494）。我希望远在他乡的饥民能够得到食物。但实现我的这一愿望并不能使*我*受益。

另外，人们可能会对自己的荒唐事产生欲望。假设我想数一数这条路上草坪上的所有草叶。如果我在做这件事时得到了满足感，那么这种满足感对我来说就是一种利益。但仅仅满足我数清这条路上草坪上所有草叶的愿望并不构成利益（Rawls 1971: 432; Parfit 1984: 500; Crisp 1997: 56）。

经过仔细思考，我们可能会认为，如果且*只有*当某人的愿望具有某组内容之一时，满足该 人的愿望才会增加其福利。例如，我们可能会认为，满足某人对快乐、友谊、知识、成就或自主的渴望*确实*会增加她的福利，而满足她对不属于这些类别的事物的任何渴望并不会直接使她受益（尽管她从满足这些渴望中获得的快乐会使她受益）。如果我们这样想，似乎我们就认为有一系列事物构成了任何人的福利（帕菲特，1984 年：附录 I；Brink 1989：221-36；Griffin 1996：第 2 章；Crisp 1997：第 3 章；Gert 1998：92-4；Arneson 1999a）。

只要要推广的商品是福利的一部分，该理论就仍然是功利主义的。功利主义有很多可取之处。显然，生命的去向是重要的。道德从根本上说是不偏不倚的，也就是说，在道德的最基本层面上，每个人都是同等重要的--女人和男人、强者和弱者、富人和穷人、黑人、白人、西班牙裔、亚洲人等等。功利主义貌似将这种同等重要性解释为，在计算整体福利时，任何一个人所获得的利益或受到的伤害，与任何其他人所获得的同样大小的利益或受到的同样大小的伤害相比，都不算多，也不算少。

3.其他待促销商品

结果论家族中的非功利主义成员是指那些仅从所产生的 "善 "的角度来评估行为和/或性格特征、做法和制度的理论，*这里的 "善 "并不局限于福利*。这里的 "非功利主义 "指的是 "非纯粹功利主义"，而不是 "完全非功利主义"。当作者把自己描述为结果主义者而非功利主义者时，他们通常是在表明，他们的基本评价不仅要考虑福利，还要考虑其他一些物品。

这些其他商品是什么？最常见的答案是正义、公平和平等。

柏拉图认为，正义就是 "各得其所"（《*共和国*》上卷）。我们可以认为，"各得其所 "是指人们被亏欠的东西，或者是因为他们应得的，或者是因为他们在道义上有权利得到这些东西。假设我们将这些观点纳入结果论。那么我们就会得到这样一种理论，即对事物的评估不仅要看产生了多少福利，还要看人们在多大程度上得到了他们应得的东西，以及在多大程度上道德权利得到了尊重。

然而，结果论采取这种思路，就等于限制了它的解释野心。一种理论仅仅预设了什么，就无法解释什么。一个结果论理论既预设了正义是由这样那样的东西构成的，又预设了正义是要促进的事情之一，但它并没有解释为什么正义的组成部分是重要的。它没有解释什么是沙漠。它没有解释道德权利的重要性，更没有试图确定这些道德权利的内容是什么。这些问题都太重要了，也太有争议了，因此后果主义理论不能不加以解释，也不能不加以讨论。如果结果论要提及正义、荒漠和道德权利，它就需要分析这些概念，并证明它赋予这些概念的作用。

关于公平，也可以有类似的说法。如果一个结果论理论预设了对公平的解释，而只是规定要促进公平，那么这个结果论理论就不是在解释公平。但是，公平（就像正义、沙漠和道德权利一样）是一个非常重要的概念，因此结果论不能不加以解释。

结果论者处理正义与公平问题的一种方法是认为，正义与公平是由符合某种合理的社会实践构成的，而这些实践的合理性在于它们普遍促进了整体福利与平等。事实上，我们可以认为，人们应得的、人们在道义上有权享有的、正义和公平所要求的，就是遵守任何促进整体福利和平等的做法。

然而，是否需要在公式中加入平等的内容却存在很大争议。许多人认为，纯粹的功利主义公式具有充分的平等主义含义。他们认为，即使目标是促进福利，而不是促进福利加平等，人类的一些偶然但普遍存在的事实也会推动物质资源的平等分配。根据 "物质资源边际效用递减规律"，一个人从某一单位的物质资源中得到的好处越少，他已经拥有的该单位物质财富就越多。假设我从只能步行到拥有一辆自行车，或者，虽然我住在一个会很冷的地方，但我从没有保暖大衣到拥有一件大衣。与我从拥有九辆自行车或十件大衣到拥有十辆自行车或十件大衣相比，我从获得第一辆自行车或第一件大衣中受益更多。

边际效用递减规律也有例外。在大多数例外情况中，一个额外的物质资源单位会使某人超过某个重要的临界点。例如，一顿饭、一粒药片或一口空气挽救了某人的生命，或者一辆汽车的购置使竞争激烈的收藏家跃居首位。在这种情况下，使人跨过门槛的单位很可能与之前的任何单位一样对人有益。不过，一般来说，物质资源的边际效用确实是递减的。

除了物质资源的边际效用递减这一假设之外，我们还要加上一个假设，即不同的人从相同的物质资源中获得的收益一般大致相同。同样，也有例外。如果你生活在寒冷的气候中，而我生活在炎热的气候中，那么你从一件温暖的大衣中得到的好处要比我多得多。

但假设我们生活在同一个地方，那里的冬天很冷，有适合骑自行车的道路，却没有公共交通。假设你有 10 辆自行车和 10 件大衣（尽管你不是在争夺什么自行车或大衣收集奖）。与此同时，我却穷得一无所有。那么，把你的一辆自行车和一件大衣重新分配给我，对你的伤害可能小于对我的好处。在资源分配不均的社会中，这种现象比比皆是。在出现这种现象的地方，从根本上说不偏不倚的道德就会受到压力，要把资源从富人那里重新分配给穷人。

然而，人类也有一些偶然但普遍存在的事实，这些事实有利于那些可预见会造成物质不平等后果的做法。首先，更高水平的整体福利需要更高水平的生产率（想想农业生产率的提高所带来的福利收益）。在许多经济领域，为提高生产率提供物质奖励似乎是最有效、最可接受的提高生产率的方法。一些个人和群体的生产率会高于其他个人和群体（特别是在有激励计划的情况下）。因此，为提高生产率提供物质奖励的做法会造成物质上的不平等。

因此，一方面，物质资源的边际效用递减对资源分配更加平等产生了压力。另一方面，提高生产率的需要又迫使人们支持那些会造成物质不平等后果的激励方案。功利主义者和大多数其他后果论者发现自己正在平衡这些相反的压力。

请注意，这些压力涉及资源的分配。还有一个问题是，福利本身应如何平等分配。近来许多作家都认为功利主义对福利分配漠不关心。试想一下，在一个总体福利很大但分配不均的结果和一个总体福利较小但分配均等的结果之间做出选择。功利主义者被认为倾向于总体福利更大的结果，即使它的平均分配也较少。

为了说明这一点，让我们人为地将人口简单地分为两组。

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | 福利单位 | | | | 两组的福利总额 |
| **备选方案 1** | 每人 | | 每组 | |  |
| A 组 10 000 人 | 1 | | 10,000 | |  |
| B 组 100 000 人 | 10 | | 1,000,000 | |  |
|  |  | |  | | *精确*计算1,010,000 |
|  | | 福利单位 | | | 两组的福利总额 |
| **备选方案 2** | | 每人 | | 每组 |  |
| A 组 10 000 人 | | 8 | | 80,000 |  |
| B 组 100 000 人 | | 9 | | 900,000 |  |
|  | |  | |  | *精确*计算980,000 |

许多人会认为备选方案 2 优于备选方案 1，并可能认为这些备选方案之间的比较表明，总是存在着支持更大福利平等的压力。

然而，正如德里克-帕菲特（Derek Parfit，1997 年）特别指出的那样，我们不能操之过急。请看下面的选择：

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | 福利单位 | | | | 两组的福利总额 |
| **备选方案 1** | 每人 | | 每组 | |  |
| A 组 10 000 人 | 1 | | 10,000 | |  |
| B 组 100 000 人 | 10 | | 1,000,000 | |  |
|  |  | |  | | *精确*计算1,010,000 |
|  | | 福利单位 | | | 两组的福利总额 |
| **备选方案 3** | | 每人 | | 每组 |  |
| A 组 10 000 人 | | 1 | | 10,000 |  |
| B 组 100 000 人 | | 1 | | 100,000 |  |
|  | |  | |  | *精确*计算110,000 |

福利平等是否如此重要，以至于备选方案 3 优于备选方案 1？举一个帕菲特的例子，假设让每个人在视力方面平等的唯一办法是让每个人都完全失明。这种 "削平 "是道德所要求的吗？事实上，这在道德上有任何可取之处吗？

如果我们认为答案是否定的，那么我们可能会认为福利平等本身并不是一个真正的理想（参见 Temkin 1993）。只有在有利于穷人的情况下，富人的损失才是合理的。我们所认为的有利于福利平等的压力，其实是有利于福利水平提高的压力。我们可以说，福利越差的人，其福利受到的影响越大。这种观点被称为*优先主义*（Parfit 1997；Arneson 1999b）。它具有巨大的直观吸引力。

举一个简单的例子来说明优先主义是如何发挥作用的：假设最贫困者的福利是较富裕者福利的五倍。那么上表中的备选方案 1 的结果是

(1×5×10,000)+(10×100,000)(1×5×10,000)+(10×100,000),

总计 1 050 000 个福利单位。同样，由于最贫困者的福利是普通人的五倍，备选方案 2 的结果是：

(8×5×10,000)+(9×100,000)(8×5×10,000)+(9×100,000),

总福利单位为 1 300 000。这符合人们的普遍反应，即备选方案 2 在道德上优于备选方案 1。

当然，在实际例子中，社会从来都不是只有一种分化。相反，社会中存在着一个等级，从最贫困的人到不太贫困的人，一直到最富裕的人。优先主义致力于对处于这一尺度上不同位置的人的利益给予不同程度的重视：一个人的境况越差，对其收益的重视程度就越高。

这引起了人们对优先主义的两个严重担忧。第一个问题是，优先主义难以*非武断地*决定对处境较差者的福利给予多大的重视。例如，最贫困者的一个福利单位是否应比最富裕者的同等福利单位多 10 倍，比一般富裕者的同等福利单位多 5 倍？或者，乘数应该是 20 倍和 10 倍，或者 4 倍和 2 倍，或者其他数额？对优先主义的第二个担忧是，更重视某些人的福利增长，而不重视其他人的同等规模的福利增长，是否与基本的公正性相矛盾（Hooker 2000: 60-2）。

这里不是进一步讨论先验主义与其批评者之间争论的地方。因此，本文的其余部分将搁置这些争论。

4.完全规则后果主义

结果论者将其理论分为三个部分：(1) 他们关于什么行为在道德上是错误的论断，(2) 他们关于行为人在做出道德决定时应使用的程序的论断，以及 (3) 他们关于在什么条件下道德制裁（如责备、内疚和赞美）是适当的论断。

我们可以称之为*完全的规则后果论，它*包含了规则后果论的所有三个标准。因此，完全的规则后果论认为，当且仅当一种行为被其后果所证明的规则所禁止时，这种行为在道德上就是错误的。它还声称，行为人应根据后果证明合理的规则做出道德决策。它还声称，实施道德制裁的条件是由后果证明合理的规则决定的。

完全规则后果论者可能会认为，关于这三个不同的主题事项，实际上只有一套规则。或者，他们可能认为有不同的规则，这些规则在某种意义上相互对应或互为补充。

比区分不同种类的完全规则后果主义更重要的是区分完全规则后果主义和*部分*规则后果主义。部分规则后果主义可能有多种形式。让我们重点讨论最常见的形式。最常见形式的部分规则后果论主张，行为人应该参照后果证明合理的规则来做出关于做什么的道德决定，但并不主张道德上的错误是由后果证明合理的规则决定的。部分规则后果论者通常赞同这样一种理论，即道德上的不法性是直接根据行为后果与其他可能行为后果的比较来确定的。这种不法性理论被称为行为后果论。

区分完全规则后果主义和部分规则后果主义可以澄清行为后果主义和规则后果主义之间的对比。行为后果主义最好被理解为仅仅坚持以下几点：

行为后果主义的不法性标准：当且仅当一个行为所带来的好处少于某个可供选择的行为所带来的好处时，该行为才是错误的。

面对道德错误的标准，许多人自然而然地认为，决定做什么的方法就是应用该标准，即：

行为后果主义道德决策程序：在每种情况下，代理人应通过计算哪种行为会产生最大的善来决定做什么。

然而，结果论者几乎从未为这种行为后果论决策程序辩护，将其视为做出道德决策的一般和典型方式（密尔 1861：第 2 章；西奇威克 1907：405-6、413、489-90；摩尔 1903：405-6、413、489-90）：405-6, 413, 489-90; Moore 1903：162-4；Smart 1956：346; 1973: 43, 71; Bales 1971: 257-65; Hare 1981; Parfit 1984: 24-9, 31-43; Railton 1984: 140-6, 152-3; Brink 1989: 216-7, 256-62, 274-6; Pettit and Brennan 1986; Pettit 1991, 1994, 1997: 156-61, 2017; de Lazari-Radek and Singer 2014: ch. 10）。有许多令人信服的结果论理由可以解释为什么行为后果论决策程序会适得其反。

首先，代理人往往不知道各种行为会产生什么后果。

其次，获取此类信息所涉及的成本往往高于所做决定的利害关系。

第三，即使代理人掌握了计算所需的信息，他也可能在计算中出错（当代理人的自然偏差出现时，或当计算很复杂时，或当计算必须匆忙进行时，这种情况尤其可能发生）。(当代理人的自然偏见侵入，或当计算复杂，或当计算必须匆忙进行时，这种情况尤其可能发生）。

第四，我们可以称之为期望效应。想象一下，在这样一个社会里，人们知道别人天生偏袒自己和自己所爱的人，但却试图通过计算整体利益来做出每一个道德决定。在这样的社会里，每个人都可能会担心，只要别人相信自己的行为会产生最大的整体利益，他们就会到处违背诺言、偷窃、撒谎，甚至殴打他人（Woodard 2019: 149）。在这样的社会中，人们会觉得无法相互信任。

与前三个因素相比，第四个因素更具争议性。例如，Hodgson 1967 年、Hospers 1972 年和 Harsanyi 1982 年认为信任会瓦解。辛格 1972 年和刘易斯 1972 年则认为不会。

然而，大多数哲学家都认为，由于上述四个原因，使用行为后果论的决策程序不会使善最大化。因此，即使是支持行为后果论道德错误标准的哲学家，也拒绝接受行为后果论道德决策程序。取而代之的是他们通常主张的以下方法：

规则后果主义决策程序：至少在通常情况下，行为主体在决定做什么时，应采用那些接受后会产生最佳后果的规则，如 "不要伤害无辜的他人"、"不要偷窃或破坏他人财产"、"不要违背诺言"、"不要撒谎"、"特别关注家人和朋友的需要"、"普遍为他人做好事 "等。

由于关于不法性标准的行为后果论者通常接受这一决策程序，因此行为后果论者实际上是部分规则后果论者。通常，作者所说的间接后果论就是这种关于不法性的行为后果论与关于适当决策程序的规则后果论的结合。

标准而言，完全的规则后果主义所认可的决策程序是*社会*最好接受的决策程序。之所以需要 "标准地 "这一限定词，是因为规则后果论的一些版本允许规则相对于小团体甚至个人（D. Miller 2010; Kahn 2012）。而行为后果论则坚持*个人*最好接受的决策程序。因此，根据行为后果论，由于杰克和吉尔的能力和情况可能截然不同，杰克接受的最佳决策程序可能不同于吉尔接受的最佳决策程序。然而，在实践中，行为后果论者通常会在大多数情况下忽略这种差异，并认可上述规则后果论的决策程序（Hare，1981 年，第 2、3、8、9、11 章；Levy，2000 年）。

当行为后果论者认可上述规则后果论的决策程序时，他们承认遵循这一决策程序并不能保证我们会做出具有最佳后果的行为。例如，有时我们遵循的决策程序不允许伤害无辜的人，这就会阻止我们做出会产生最佳后果的行为。同样，在某些情况下，偷窃、违背诺言等行为也会产生最佳后果。尽管如此，从长远来看，我们遵循一种普遍排除此类行为的决策程序所产生的后果，总体上可能要比我们试图逐个行为地进行后果论计算所产生的后果好得多。

由于行为后果论者通常同意规则后果论者的决策程序，因此将某些哲学家归类为行为后果论者还是规则后果论者可能会有问题。例如，G.E. 摩尔（1903, 1912 年）有时被归类为行为后果论者，有时被归类为规则后果论者。与其他许多人一样，包括他的老师亨利-西奇威克（Henry Sidgwick），摩尔将行为后果论的道德错误标准与规则后果论的程序相结合，以决定做什么。摩尔只是比大多数人更进一步地强调了偏离规则后果主义决策程序的危险性（见 Shaw 2000）。Hare（1981）和 Pettit（1991、1994、1997）也是特别有影响力的行为后果论者，他们认为日常决策应该根据熟悉的规则来进行，而不是只关注后果。

5.全球后果主义

一些作家提出，后果主义最纯粹、最一致的形式是这样一种观点，即绝对任何事情都应根据其后果来评估，不仅包括行为，还包括规则、动机、实施制裁等。让我们效仿佩蒂特和史密斯（2000 年），将这种观点称为*全球后果主义*。卡根（Kagan，2000 年）将其称为多维直接后果主义，即直接根据每件事本身的后果是否与其他选择的后果一样好来评估。

这种全面后果主义与我们所说的部分规则后果主义有何不同？我们所说的部分规则后果主义不过是行为后果主义的道德错误标准与规则后果主义的决策程序的结合。按照这样的定义，部分规则后果主义对道德制裁何时合适的问题没有定论。

一些部分规则后果论者认为，只要行为主体没有选择会导致最佳后果的行为，他们就应该受到指责并感到内疚。部分规则后果论者更合理的立场是，只要行为主体选择了规则后果论决策程序所禁止的行为，就应该受到指责并感到内疚，无论该行为是否未能产生最佳后果。最后，我们所定义的部分规则后果论与以下主张是一致的，即行为主体是否应该受到责备或感到内疚，并不取决于他们的行为是否错误，也不取决于所推荐的道德决策程序是否会导致他们选择他们所选择的行为，而是完全取决于这种责备或内疚是否会带来任何好处。这正是全球结果主义对制裁的看法。

对全面后果主义的一种反对意见是，同时对行为、决策程序和实施制裁适用后果主义标准会导致明显的悖论（Crisp，1992 年；Streumer，2003 年；Lang，2004 年）。

假设从整体和长远来看，你所能接受的最佳决策程序是导致你现在采取 *x* 行为的程序。因此，总体后果论告诉你：（a）使用最佳决策程序，但（b）不要做这个决策程序选出的行为。这似乎是自相矛盾的。

如果我们考虑到自责和内疚，情况就会变得更糟。假设你遵循了可能的最佳决策程序，但却没有做出具有最佳后果的行为。你应该受到责备吗？你应该感到内疚吗？总体结果论认为，只有当指责你会产生最佳后果时，你才应该受到指责；只有当这样做会产生最佳后果时，你才应该感到内疚。假设由于某种原因，责备你遵循规定的决策程序（因此做了 *x*）会产生最好的后果。但是，道德理论要求责备你，尽管你遵循了该理论规定的道德决策程序，这肯定是自相矛盾的。或者，假设由于某种原因，责备你故意选择具有最佳后果的行为（*y*）会导致最佳后果。同样，道德理论要求指责你，尽管你有意选择了理论所要求的行为，这肯定是自相矛盾的。

因此，全球结果论的一个问题是，它在它所宣称的必要行为与它告诉行为人使用的决策程序之间，以及在每种行为与无责之间，都造成了潜在的差距。(对这一攻击思路的明确回应，见 Driver 2014：175 and de Lazari-Radek and Singer 2014：315-16.)

这并不是人们最熟悉的全球结果论的问题。人们最熟悉的问题反而是它的最大化行为后果主义的错误标准。根据这一最大化标准，当且仅当某一行为未能带来最大的善时，它才是错误的。根据这一标准，一些看似错误的行为被判定为没有错误。它也把一些似乎没有错的行为判定为错的。

例如，考虑一个谋杀行为，它所产生的善比任何其他行为所产生的善稍多一些。根据我们最熟悉的行为后果最大化主义的错误标准，这种谋杀行为并不是错误的。其他许多行为，如殴打、偷窃、违背诺言和撒谎，即使做了会比不做产生的善稍多，也可能是错误的。同样，我们所熟悉的行为后果论的最大化形式否认了这一点。

或者考虑一下，有人把自己的一些资源给了孩子，或者自己留着，而不是拿出来帮助一些陌生人，因为他们本可以从这些资源中获得更多。这样的行为似乎并没有错。然而，行为后果最大化标准却判定它是错误的。事实上，想象一下，一个普通的富裕者要做出多大的自我牺牲，她的进一步行动才能满足行为后果最大化的错误标准。她必须付出到这样一个地步，即为了他人的利益而做出进一步的牺牲，对她造成的伤害会大于对他人的利益。因此，最大化行为后果主义的不法性标准常常被指责为不合理的要求。

刚才针对最大化行为结果论的反对意见，可以通过一种不要求最大化善的行为结果论来回避。这种行为后果论现在被称为满足后果论。有关这种理论的更多信息，请参阅[后果](https://plato.stanford.edu/entries/consequentialism/)主义条目。

6.提出完全规则后果主义

规则后果论有多种不同的表述方式。例如，它可以从规则实际产生的善的角度来表述，也可以从规则后果的合理预期善的角度来表述。它可以从遵守规则的后果角度来表述，也可以从接受规则的更广泛后果角度来表述。它可以从绝对人人都接受规则的后果的角度来表述，也可以从并非人人都接受规则的角度来表述。规则后果主义也可以从向下一代的每个人传授准则的角度来表述，但要充分认识到由此产生的接受程度会有所不同（Mulgan 2006, 2017, 2020; D. Miller 2014, 2021; T. Miller 2016, 2021）。规则后果主义在某些方面的表述比在其他方面的表述更可信。以下三个小节将对此进行解释。关于规则后果论的反对意见，也与表述问题有关。

6.1 实际货物与预期货物

如前所述，完全的规则后果论包括对三个问题的规则后果论回答。第一个问题是，是什么使得行为在道德上是错误的？第二个问题是，行为人应该使用什么程序来做出道德决定？第三个问题是，在什么条件下，责备、内疚和赞美等道德制裁是适当的？

正如我们所看到的，完全规则后果论者对决策程序问题的回答与其他类型后果论者对该问题的回答大同小异。因此，让我们聚焦于对比点，即另外两个问题。这两个问题--关于什么行为是错误的和关于什么时候制裁是适当的--之间的联系比我们有时意识到的要紧密得多。

事实上，结果论之父米尔（J.S. Mill）肯定了它们之间的紧密联系：

我们不称任何事情为错事，除非我们的意思是说，一个人如果做了错事，就应该受到某种惩罚；如果不是法律的惩罚，就是同类的舆论的惩罚；如果不是舆论的惩罚，就是自己良心的谴责。(1861：第 5 章第 14 段）

让我们假定，密尔认为 "应该受到惩罚，至少受到自己良心的惩罚，如果不受到别人的惩罚的话 "与 "值得责备 "大致相同。有了这个假设，我们就可以把密尔解释为把 "不法性 "与 "可责性 "紧密联系在一起。稍后，我们可以考虑一下，如果密尔错误地把不法性与可责性紧密联系在一起，会出现什么情况。首先，让我们考虑一下，如果密尔认为不法性与可责性紧密相连是正确的，那么接下来会发生什么。

请看下面的论证，它的第一个前提来自密尔：

如果一种行为是错误的，那么它就应该受到谴责。

当然，代理人不能因为接受和遵循了其无法预知会产生次优后果的规则而受到责备。由此，我们得到了第二个前提：

如果一个行为应受到指责，那么允许该行为的规则所产生的次优后果一定是可以预见的。

从这两个前提中，我们可以得出结论：

因此，如果一个行为是错误的，那么允许该行为的规则所产生的次优后果一定是可以预见的。

当然，接受一套规则的实际后果可能与接受这套规则的可预见后果不同。因此，如果完全的规则后果论认为，当且仅当允许某一行为的规则的*可预见后果*是次优的时候，该行为就是错误的，那么规则后果论就不能认为，当且仅当允许某一行为的规则的*实际*后果是次优的时候，该行为就是错误的。

现在，假设不法性与可责性之间的关系比密尔所说的要松散得多（参见索伦森，1996 年）。也就是说，假设我们的不法性标准可以与我们的可责性标准大相径庭。在这种情况下，我们可以认为

*实际主义*规则后果主义的*不法性*标准：当且仅当一种行为被规则所禁止，而这种规则*实际上*会带来最大的善时，这种行为才是错误的。

和

*可责性的预期*规则后果主义标准：当且仅当一种行为被*可预见会*带来最大利益的规则所禁止时，这种行为才是可责的。

让我们把 "可预见地带来最大的好处 "改为 "带来最大的预期好处"。下面是一套规则的预期利益的计算方法。假设我们可以确定一套规则可能产生的每种结果的价值或不价值。将每种可能结果的价值乘以出现该结果的概率。将所有乘积相加。得出的数字就是这套规则的预期效益。

预期利益不是通过人们对可能结果的疯狂概率估计来计算的。相反，预期利益的计算方法是将可能结果的价值或不价值乘以*合理*或*正当的*概率估计。

不同的行为者有不同的证据，因而有不同的理性和合理的概率估计。这种差异有时恰恰解释了为什么人们对现有道德准则的修改会产生分歧。在某些分歧案例中，造成分歧的原因是至少有一方没有意识到或没有完全吸收现有的证据。期望清单规则后果论者很可能希望将合理或正当的概率估计与当时可用的证据联系起来，即使有些人没有意识到或没有完全理解其含义。

对于如何计算预期效益可能存在相当大的怀疑（Lenman，2000 年）。即使有可能进行这种计算，它们也往往是相当印象化和不精确的。尽管如此，我们还是有理由希望至少能对其他可能规则的*可能*后果做出*一些*明智的判断（伯奇-布朗，2014 年）。我们可以以这些判断为指导，就像立法者经常说的那样。与此相反，哪些规则*实际上*会产生*最好的*后果，通常是无法获得的。因此，预期清单规则后果主义的可责性标准很有吸引力。

现在，我们回到这样一个提议，即虽然可责性的标准是期待主义的规则后果论标准，但道德错误的正确标准却是现实主义的规则后果论标准。这一提议摒弃了密尔将道德不法性与有责性联系在一起的做法。然而，有人对这一提议提出了强烈的反对意见。如果将道德不法性与可责性割裂开来，那么道德不法性的作用和重要性何在？

为了保留道德不法性的明显作用和重要性，那些致力于期望清单规则-后果主义的*有责性*标准的人很可能会赞同：

*简单的预期*规则后果主义道德错误标准：当且仅当一种行为被规则所禁止时，这种行为在道德上是错误的，而这种规则会带来最大的预期利益。

事实上，一旦我们面前有了(a) *实际结果的*价值量和(b) 理性*预期的*善之间的区别，完全的规则后果论者就很可能会去选择道德错误、可责备性和决策程序的预期清单标准。

如果据我们所知，没有一个代码的预期价值高于其竞争对手呢？我们需要修改我们的预期值标准，以适应这种可能性：

*复杂的预期*规则主义道德错误标准：当且仅当一种行为被可带来最大*预期*善的规则所禁止，或者当两种或两种以上可供选择的规则在预期善方面同样最佳时，被其中最接近传统道德的规则所禁止，这种行为在道德上就是错误的。

利用与传统道德的亲近程度来打破原本同样有前途的准则之间的联系的论点，首先是观察到社会变革经常会产生意想不到的后果。而这些意想不到的后果通常似乎都是负面的。此外，新准则与已被普遍接受的准则之间的差异越大，产生意外后果的可能性就越大。因此，在我们认为具有同样高预期价值的两个准则之间，我们应该选择与我们已经知道的准则最接近的一个。(关于两套准则具有同样高的预期价值且似乎同样接近传统道德的讨论，见 Hooker 2008：83-4.)

其含义是，人们应该改变现状，但前提是这些改变比维持现状具有更大的预期价值。规则后果主义显然有能力建议变革。但它并不赞成为了变革而变革。

规则后果论的表述肯定需要处理预期价值的联系问题。不过，在本文的其余部分，我将忽略这一复杂问题。

6.2 合规和验收

规则后果论者还面临其他一些重要的表述问题。其中一个问题是，规则后果论应该从遵守规则的角度来表述，还是从接受规则的角度来表述，或者从教授规则的角度来表述。诚然，教授和接受规则的最重要后果是遵守规则。规则后果论的早期表述确实明确提到了遵守规则。例如，他们说，当且仅当一种行为被规则所禁止时，这种行为在道德上是错误的，而*遵守规则*将使善（或预期善）最大化。(见 Austin 1832；Brandt 1959；M. Singer 1955，1961）。

然而，接受一项规则可能会产生遵守规则之外的其他后果。正如卡根（2000: 139）所写，"规则一旦嵌入，就会对结果产生影响，而这种影响与规则对行为的影响无关：比如说，可能仅仅是对一套规则的思考就会让人们感到安心，从而促进幸福。(关于我们可以称之为规则的'超越遵守的后果'的更多信息，请参见 Sidgwick 1907：405-6, 413; Lyons 1965：140; Williams 1973: 119-20, 122, 129-30; Adams 1976: 470; Scanlon 1998: 203-4; Kagan 1998: 227-34）。

接受规则的这些后果无疑应成为规则成本效益分析的一部分。用接受规则的后果来表述规则后果主义，可以使它们成为这种分析的一部分。共同接受一套规则的重要后果包括保证效应、激励效应和威慑效应。对保证和激励效应的考虑在规则后果论的发展中发挥了重要作用（Harsanyi 1977, 1982: 56-61; 1993: 116-18; Brandt 1979: 271-77; 1988: 346ff [1992: 142ff.]; 1996: 126, 144; Johnson 1991, 特别是第 3, 4, 9 章）。因此，我们不应对规则后果论从遵守规则到接受规则的转变感到惊讶。

然而，正如我们需要从思考遵守规则的后果转向思考接受规则的更广泛后果一样，我们还需要更进一步。单纯关注接受规则的后果，会忽视教师教授这些规则的 "过渡 "成本，以及教学对象内化这些规则的成本。然而，这些代价肯定是巨大的（Brandt 1963: section 4; 1967 [1992: 126]; 1983: 98; 1988: 346-47, 349-50 [1992: 140-143, 144-47]; 1996: 126-28, 145, 148, 152, 223）。

举例来说，假设一个相当简单、要求相对较低的规则代码 *A* 被接受后，代码 *A 的*期望值为 *n*；假设一个更复杂、要求更高的替代代码 *B* 被接受后，代码 *B 的*期望值为 n+5。因此，如果我们只考虑两个备选代码被接受的预期值，代码 *B* 胜出。

但是，现在让我们把教授这两套准则和让新一代人内化这两套准则的相对成本考虑进我们对立准则的成本/效益分析中。由于守则 *A* 相当简单，要求也相对较低，因此使其被内化的成本为-1。由于代码 *B* 更复杂，要求更高，使其被内化的成本为-7。因此，如果我们在比较这两套代码时考虑到让它们被内化的各自成本，代码 *A* 的预期值为 n-1-1，代码 *B 的*预期值为 n+5-7。一旦我们把代码内化的各自成本计算在内，代码 *A 就*赢了。

如前所述，成功传授一种准则，使其被广泛内化的成本是 "过渡成本"。当然，这种过渡总是从一种安排过渡到另一种安排。我们设想的过渡安排是接受某项拟议的守则。我们想象中的过渡安排*是从......*嗯，什么？

一种答案是，过渡的起点应该是社会已经接受的道德准则。这似乎是一个自然的答案。然而，这个答案遭到了强烈的反对，即规则后果主义不应该让拟议守则的成本/收益分析受到让人们放弃他们可能已经内化的任何规则的成本的影响。这有两个原因。

最重要的是，规则后果主义对法典的评估需要避免直接或间接地重视那些源于其他道德理论而非规则后果主义本身的道德观念。假设某个社会的人们从小就认为女性应该服从男性。规则后果论者在评价拟议的无性别歧视准则时，是否应该考虑让人们放弃他们已经内化的性别歧视规则，从而接受新的无性别歧视规则的成本？既然性别歧视规则是不合理的，那么就不应该让它们被接受的事实影响到规则后果主义的评估。

拒绝我们正在考虑的答案的另一个原因是，它有可能支持一种缺乏吸引力的相对主义。不同的社会在其现有的道德信仰方面可能存在很大差异。因此，如果评估所提出的准则的方法考虑的是让人们已经承诺遵守其他准则的成本，那么最终就不得不考虑为遵守同一准则而付出的不同过渡成本。例如，从已被接受的种族主义准则过渡到非种族主义准则的成本要比从已被接受的非种族主义准则过渡到非种族主义准则的成本高得多。制定规则后果主义，使其为 20 世纪 60 年代的密歇根州和 20 世纪 60 年代的密西西比州认可相同的准则，是可取的。(关于应制定规则后果主义以维护社会相对主义的反对论点，见 Kahn 2012；D. Miller 2014, 2021；T. Miller 2021）。

为了避免出现刚才提到的社会相对主义，我们可以从*新一代人的*接受程度的角度来阐述规则后果论。我们可以提出这样的建议：我们可以比较其他准则各自的 "教学成本"，假设这些准则将被教授给新一代的儿童，即尚未接受过接受道德准则的教育的儿童。我们可以想象，孩子们一开始就会自然而然地（非道德地）倾向于对自己和其他一些人非常偏袒。我们还应该假设，学习每一条规则都需要付出认知成本。

这些都是现实的假设，具有重大意义。其一，对其他规则守则进行成本/效益分析时，有理由倾向于采用更简单的守则，而不是更复杂的守则。当然，制定更多或更复杂的规则也会带来好处。然而，如果把教学成本计算在内，规则的复杂程度可能是有限度的，但其预期价值仍然大于简单的规则。

另一个影响涉及到为帮助他人而做出牺牲的预期规则。由于孩子们一开始就只关注自己的满足，让他们内化一种要求他们为了他人的利益而一再做出巨大牺牲的大公无私精神，会付出极高的代价。当然，这种规则的内化也会带来巨大的利益--主要是对他人的利益。利益会大于成本吗？至少自西奇威克（1907：434）以来，许多功利主义者理所当然地认为，人类的本性是这样的，即真正的可能性是：（1）人类热衷于关心某些人，而较少关心人类的其他每个人；或者（2）人类对每个人的关心微弱但不偏不倚。换句话说，根据这种人性观，不现实的可能性是人类强烈而公正地关心世界上的每一个人。如果这种人性观点是正确的，那么成功地让人们完全公正的一个巨大代价就是，这样做会让他们只剩下微弱的关心。

即使这种对人性的描绘不正确，也就是说，即使可以在不耗尽人们的热情和激情的情况下使人们完全公正，但要成功地使人们像关心自己一样关心其他每个人，其代价也是高昂的。在从完全偏袒到完全公正的光谱上的某一点上，推动和诱导每个人进一步沿着光谱前进的成本超过了收益。

值得一提的一个复杂因素是，道德教育显然是发展性的。教给儿童多少规则，这些规则有多复杂，规则的要求有多高，以及如何解决规则之间的冲突，都取决于儿童所处的发展阶段。因此，教给年幼儿童的较简单、要求较低的规则，与在较高发展阶段教给他们的较复杂、细致入微、要求较高的规则之间，可能存在冲突。当然，规则后果论的最佳表述是发展末期的规则，而不是早期阶段的规则。

6.3 完全验收与不完全验收

对准则进行规则后果主义的成本/收益分析时，应计算使这些准则被新一代内化的成本，但这种分析也应承认，内化不会在每一个人身上实现（D. Miller 2021: 130）。无论教得多么好，结果都是不完美的。有些人学到的规则在某种程度上会与教给他们的规则不同，其中有些人最终会对什么是道德要求、什么是道德选择或什么是道德错误产生非常错误的看法。还有一些人最终会缺乏足够的道德动机。还有一些人最终根本没有道德动机（精神病患者）。规则后果论需要有一些规则来应对道德教育不可避免的不完美结果。

这些规则主要包括有关惩罚的规则。从规则后果论的角度来看，惩罚的一个目的是阻止某些行为。惩罚的另一个目的是让没有受到威慑的危险分子离开街道。也许规则后果论可以承认，惩罚的另一个目的是安抚此类行为受害者及其亲友的原始复仇欲望。最后，惩罚规则还具有表达和强化的力量。

规则后果论的某些表述方式使得关于惩罚的规则难以解释。其中一种规则后果论的表述方式是：

一种行为在道德上是错误的，当且仅当这种行为被绝对每个人都完全接受并能产生最大预期善果的规则所禁止。

想象一下，在这个世界上，绝对每一个成年人都完全接受禁止（例如）对无辜者进行人身攻击、偷窃、违背诺言和撒谎的规则。假设这些规则已经被这个世界上的每个人深深地内化，以至于在这个世界上，人们完全遵守这些规则。还假设，如果这个世界上的每个人都始终遵守这些规则，那么这种完美的遵守就会成为常识。在这个世界上，很少或根本不需要关于惩罚的规则，因此制定这些规则的好处也很少或根本没有。但是，每一条规则的传授和内化都需要付出传授和内化成本。因此，纳入任何有关惩罚的规则都会产生教学和内化成本。有成本而无收益的组合是令人厌恶的。因此，对于一个完全遵守规则的世界来说，上文提到的规则后果论版本不会赞同关于惩罚的规则。

我们需要一种规则后果主义，其中包括处理不遵守正确规则的人的规则，甚至包括处理不可救药的人的规则。换句话说，规则后果论的表述需要把社会概念化为包含一些对正确规则不够执着的人，甚至包含一些不执着于任何道德规则的人。这里有一种方法：

当且仅当一种行为被一种规则所禁止时，这种行为才是错误的，而每一代人中的绝大多数人接受这种规则将产生最大的预期价值。

请注意，规则后果主义既不赞同也不宽恕绝大多数人之外的人不接受准则。相反，规则后果主义认为这些人在道德上是错误的。事实上，如此表述规则后果主义的全部意义就在于为如何消极回应这些人的规则留出空间。

当然，"绝大多数 "这个词非常不精确。假设我们选择一个精确的社会百分比，比如 90%，来消除这种不精确性。选择任何精确的百分比都有明显的随意性。例如，如果我们选择 90%，为什么不选择 89%或 91%呢？

也许，我们可以主张将 90%这个数字作为两种压力之间的合理折衷。一方面，我们所选的百分比应足够接近 100%，以保留理想情况下道德规则会被所有人接受的想法。另一方面，这个百分比又必须远远低于 100%，以便为有关惩罚的规则留出相当大的余地。考虑到平衡这些因素的需要，90% 似乎是一个可以接受的范围。(反对意见见 Ridge 2006；对 Ridge 的答复见 Hooker and Fletcher 2008。H. Smith 2010；Tobia 2013, 2018；T. Miller 2014, 2021；Toppinen 2016；Portmore 2017；Yeo 2017；Podgorski 2018；Perl 2021）。

霍利-史密斯（2010 年）指出，对低于 100%的正百分比人口接受任何特定守则的成本/效益分析，取决于其他人口的接受程度和意愿。请看下面的对比。一种假想情况是，90% 的人接受一种准则，另外 10%的人接受一种非常相似的准则，因此这两种准则在实践中很少出现分歧。第二种假想情况是，90% 的人接受一种准则，另外 10%的人接受各种准则，这些准则在实践中经常与 90%的人接受的准则发生严重冲突。与第二种情况相比，第一种情况下冲突的解决和执行并不那么重要。因此，如果规则后果主义是以少于 100%的人接受一种准则的方式来制定的，那么对不接受这种准则的人口比例做出什么样的假设就很重要了。

一些理论家提出，以教学具有最大预期价值的准则来表述规则后果论，要优于以固定或可变的准则接受率来表述该理论（Mulgan，2006，141，147；2017，291；2020，12-21；T. Miller，2016；2021；D. Miller，2021）。规则后果论的 "教学表述 "认为：

一个行为在道德上是被禁止的、必须的或可选择的，当且仅当且因为它是被规则准则所禁止、必须或允许的，而规则准则对每个人的教导至少与任何其他准则的教导具有同样的预期价值。

需要立即澄清两点。首先，这里的 "每个人 "需要加以限定，以便不包括 "有重大认知或内涵缺陷 "的人（D. Miller 2021: 10）。其次，向每个人传授准则并不意味着每个人都会内化这一准则。因此，即使每个人都被灌输了某种准则，但可以预见的是，有些人内化的规则或多或少会与灌输的规则有所不同。也会有一些人的道德动机不可靠，甚至无法辨别。也会有一些人对道德规则，甚至对道德的其他方面持怀疑态度。

规则后果论的教学有一些明确的优势。可以肯定的是，我们在进行成本/收益分析时，必须考虑到要付出足够的成本，使教学至少在一定程度上取得成功。换句话说，我们必须坚持，教学必须足够成功，才能让足够多的人内化规则，从而实现良好的合作与安全。但是，我们并不需要精确到有多大比例的被教导者仍然是无道德主义者，有多大比例的被教导者最终内化了与被教导者有些不同的准则，或者有多大比例的被教导者最终没有足够的动力去遵守他们内化了的道德准则，除非有效的执行到位（D. Miller 2021; T. Miller 2021）。

教导表述的另一个好处是，向每个人传授准则的想法与这一准则应成为公共知识的想法紧密相连。道德规范必须适合成为公共知识的想法非常吸引人（拜尔 1958: 195-196；罗尔斯 1971: 133；格尔特 1998: 10-13、225、239-240；希尔 2005；库雷顿 2015；佩蒂特 2017：39, 65, 102).规则后果论者主张道德规则受制于这种 "公共性条件"（Brandt 1992: 136, 144; Hooker 2010; 2016, forthcoming; Parfit 2011, 2017a; D. Miller 2021: 131-32）。

7.规则后果论的三种论证方法

规则后果主义认可什么规则呢？它赞同禁止攻击无辜者、夺取或损害他人财产、违背诺言和撒谎的规则。它还赞同要求人们特别关注自己的家人和朋友的需要，但更普遍的是愿意帮助他人完成（道德上允许的）项目的规则。一个在公共准则中突出这些规则的社会，可能比一个缺乏这些规则的社会有更多的善。

这些规则得到了规则后果主义的认可，这使得规则后果主义很有吸引力。因为从直觉上讲，这些规则似乎是正确的。然而，其他道德理论也认可这些规则。最明显的是，我们熟悉的一种道德多元论认为，这些直观上具有吸引力的规则构成了道德的最基本层面，也就是说，这些规则背后并不存在更深层次的道德原则。这种观点被称为罗斯多元论（为了纪念其倡导者 W.D. Ross（1930 年，1939 年））。

规则后果主义可能同意罗斯的多元论，赞同禁止攻击无辜者、偷窃、违背诺言的规则，以及要求各种忠诚和更普遍地为他人做好事的规则。但是，规则后果主义超越了罗斯的多元论，它明确了一个基本的统一原则，为这些规则提供了公正的理由。其他道德理论也试图做到这一点。这些理论包括某些形式的康德主义（Audi 2001, 2004）和某些形式的契约主义（Scanlon 1998; Parfit 2011; Levy 2013）。无论如何，论证规则后果论的第一种方法是，它指明了一个基本原则，为直观上看似合理的道德规则提供了公正的理由，而且没有任何其他理论也能做到这一点（厄姆森，1953 年；勃兰特，1967 年；霍斯珀斯，1972 年；胡克，2000 年）。(对规则后果论这一论证思路的攻击包括 Stratton-Lake 1997 年；Thomas 2000 年；D. Miller 2000 年；Montague 2000 年；Arneson 2005 年；Moore 2007 年；Hills 2010 年；Levy 2014 年）。

规则后果论的第一种论证方式可以被视为借鉴了这样一种观点，即一种理论对我们来说更有道理，因为它增加了我们信念的一致性（罗尔斯，1951 年，1971 年：19-21，46-51；德保罗，1987 年；埃伯兹，1993 年；赛尔-麦考德，1986 年，1996 年）。[参见 "[认识论合理性的一致性理论](https://plato.stanford.edu/entries/justep-coherence/) "条目。]但这种方法也可被视为中度基础主义，因为它以一系列信念（各种道德规则）为起点，并赋予这些信念以独立的可信性，尽管这些信念并非无懈可击（奥迪，1996 年，2004 年；克里斯普，2000 年）。[诚然，与我们的道德信念一致并不能使道德理论成为*真理，*因为我们的道德信念可能是错误的。然而，如果道德理论与我们的道德信念严重不一致，就会削弱该理论对我们的合理性。

沃尔夫（2016）和科普（2020）认为，关于道德是一种具有特殊功能的社会实践的元伦理学观点可能会引导我们走向规则后果主义。然而，胡克（即将出版）认为，关于道德是一种具有特殊功能的社会实践的元伦理学观点，需要从它是否与我们深思熟虑的道德原则和更具体的深思熟虑的道德判断相一致的角度加以论证。换句话说，关于道德功能的元伦理学观点需要根据它是否有助于我们实现信念之间的反思性平衡来判断。

规则后果论的第二种论证方式则截然不同。它以对后果主义评估的承诺为开端。以这一承诺为第一前提，然后指出，*间接地*评估行为，如关注共同接受规则的后果，实际上会比直接评估行为本身的后果产生更好的后果（奥斯汀 1832 年；勃兰特 1963 年、1979 年；哈桑尼 1982 年：58-60；1993 年；莱利 2000 年）。毕竟，决定做什么才是对行为进行道德评估的重点。因此，如果一种对行为进行道德评估的方式有可能导致错误的决定，或更普遍地导致不良后果，那么，根据结果论者的观点，这种评估行为的方式就更糟糕了。

前面我们看到，现在所有的结果论者都认为，以预期价值来单独评估每种行为一般来说是一种糟糕的道德决策程序。行为主体在决定如何行动时，应该求助于某些规则，如 "不对他人进行人身攻击"、"不偷窃"、"不违背诺言"、"特别关注家人和朋友的需要 "以及 "普遍乐于助人"。这些都是规则后果论所认可的规则。然而，许多后果论者认为，这并不能说明完全的规则后果主义是后果主义的最佳形式。一旦区分了 "做什么的最佳道德决策程序 "和 "道德对错标准"，所有后果论者都会承认，我们需要规则后果论的规则作为决策程序。但不是规则后果论者的后果论者则认为，这些规则在道德是非标准中不起作用。因此，这些结果论者拒绝本文所称的完全规则后果论。

然而，我们刚才考虑的对规则后果论的第二种论证方式的反对意见是否是一个好的反对意见，取决于区分适合做出道德决定的程序与道德对错标准是否合法。这个问题仍然存在争议（Hooker，2010；de Lazari-Radek和Singer，2014：第10章）。

然而，规则后果论的第二种论证方式遇到了另一种截然不同的反对意见。这种反对意见认为，规则后果主义论证的第一步是对后果主义评估的承诺。这第一步本身就需要理由。为什么假定以后果主义的方式评估事物是唯一合理的呢？

也许有人会说，后果主义评估是合理的，因为促进公正的善具有明显的直观吸引力。但这是行不通的，因为除了结果论评估之外，还有其他方法也具有明显的直观吸引力，例如 "根据没有人会合理拒绝的准则行事"。我们需要的是一种论证道德理论的方法，而不是从乞求哪种理论最合理这个问题开始。

规则后果论的第三种论证方式是契约论（Harsanyi 1953, 1955, 1982, 1993; Brandt 1979, 1988, 1996; Scanlon 1982, 1998; Parfit 2011; Levy 2013）。假设我们可以指定合理的条件，在这些条件下，每个人都会选择，或至少有足够的理由选择相同的规则准则。直观地说，这种理想化的协议将使规则规范合法化。现在，如果这些规则的内部化能使预期利益最大化，那么契约主义就会把我们引向规则后果主义的规则。

关于在各种可能的道德规则中做出选择的合理条件，存在着不同的观点。一种观点认为，每个人的公正性都必须通过强加一层假定的 "无知之纱 "来保证，在这层 "无知之纱 "后面，没有人知道自己的任何具体事实（哈桑尼，1953 年，1955 年）。另一种观点认为，我们应该设想人们在选择道德准则时所依据的是：(a) 关于对每个人的不同影响的全部经验信息；(b) 正常的关注（利己和利他的）；(c) 大致相同的讨价还价能力（勃兰特，1979 年；参见格特，1998 年）。帕菲特（Parfit，2011 年）建议寻求每个人都有（个人或公正）理由选择的规则或每个人都接受的意愿。如果公正的理由总是充分的，即使遭到个人理由的反对，那么每个人都有充分的理由希望每个人都接受其普遍接受将产生公正考虑的最佳后果的规则。同样，Levy（2013）假设没有人会合理地拒绝这样一种规则，即它给自己带来的负担加起来小于其他每种规则给他人带来的负担的总和。这样的论证表明契约主义与规则后果主义是等同的。(关于帕菲特对规则后果主义的契约论证是否成功的评估，见 J. Ross 2009；Nebel 2012；Hooker 2014。关于规则后果主义与斯坎伦的契约论之间的相似之处，以及在这两种理论之间做出决定的困难，见 Suikkanen 2022）。

8.规则后果论是否必须为崩溃、不连贯或规则崇拜负责？

从 20 世纪 60 年代中期到 90 年代中期，对规则后果论的反对意见中经常提到的一点是，这一理论被以下两难困境中的一角或另一角致命地卡住了：要么规则后果论与更简单的行为后果论在实践上等同，要么规则后果论不连贯。

有人认为规则后果主义与行为后果主义在实践上是等同的，原因就在这里。考虑一条规则后果主义声称赞成的规则--例如 "不要偷窃"。现在假设一个代理人处于某种情况下，偷窃比不偷窃会产生更多的好处。如果规则后果主义根据预期的好处来选择规则，那么规则后果主义似乎不得不承认，遵守 "除非在......或......或......的情况下不要偷窃 "的规则比遵守 "不要偷窃 "这一更简单的规则更好。这一点具有普遍性。换句话说，在遵守某些规则不会产生最大预期利益的每一种情况下，规则后果主义似乎都会倾向于遵守某种修正后的规则，而这种规则在当前情况下不会错过产生最大预期利益的机会。但是，如果规则后果主义是这样运作的，那么在实践中，它最终会要求行为后果主义所要求的行为。

如果规则后果论最终要求与行为后果论所要求的行为相同的行为，那么规则后果论就确实陷入了可怕的困境。在这两种理论中，规则后果主义更为复杂。这导致了以下反对意见。如果我们可以用更简单的行为后果论更有效地得到同样的实际结果，那么规则后果论及其无限修正的规则又有什么意义呢？

有人反对规则后果论者的理论与行为后果论在实践上等同。这个回答的依据是，最好的规则后果论*不是根据遵守规则的*预期好处，而是根据*教导*和*接受*规则的预期好处来排列规则体系。举例来说，如果一条禁止偷窃的规则被附加了一个又一个例外条款，那么包含所有这些例外条款的规则就会为诱惑提供太多机会，使行为人相信其中一个例外条款是适用的，而事实上偷窃对行为人是有利的。而关于诱惑的这一点也会削弱其他人对其财产不会被盗的信心。大多数其他道德规则也是如此：纳入过多的例外条款可能会削弱人们对其他人会以某些方式行事（如遵守承诺和避免偷窃）的信心。

此外，在比较其他规则时，我们还必须考虑让新一代人内化这些规则的相对成本。显然，让新一代学习数量庞大的规则或极其复杂的规则的成本是高昂的。因此，规则后果主义倾向于制定一套没有太多规则、规则内部也没有太复杂的规则。

让新一代人接受要求自己为与自己没有特殊关系的其他人做出巨大牺牲的规则，也需要付出代价。当然，遵守这些苛刻的规则会带来许多好处，主要是对他人的好处。但是，在权衡遵守这些规则所带来的好处时，应该权衡内化这些规则的相关成本。在某种苛刻程度上，将这些苛刻的规则内化的成本会超过遵守这些规则所带来的收益。因此，对要求苛刻的规则进行仔细的成本/收益分析，会得出反对规则过于苛刻的结论。

规则后果主义所推崇的规则准则是由数量不多、不太复杂或要求不高的规则组成的准则，有时会导致人们做出预期价值并不最大的行为。For example, following the simpler rule “Don’t steal” will sometimes produce less good consequences than following a more complicated rule “Don’t steal except when … or … or … or … or …”.另一个例子可能是，遵循一条允许人们在一定程度上优先考虑自己的项目的规则，即使他们可以通过牺牲自己的项目来帮助他人，从而产生更多的善果。尽管如此，规则后果论的论点是，使人们广泛接受更简单、要求更低的代码，即使接受这种代码有时会导致人们做出后果次优的行为，但从长远来看，其预期价值要高于使人们广泛接受最复杂、要求最高的代码。由于规则后果论可以告诉人们遵循这种更简单、要求更低的准则，即使遵循这种准则不会使预期利益最大化，因此规则后果论摆脱了与行为后果论的实际等同性。

只要规则后果论规避了崩溃，这一理论就会被指责为不连贯。规则后果论之所以被指责为不连贯，是因为它坚持认为，尽管某一行为未能最大限度地实现预期善，但在道德上是允许的，甚至是必须的。这种指责的背后必然是这样一种假设，即规则后果主义包含着对最大化善的压倒一切的承诺。有了这种压倒一切的承诺，却反对承诺所要求的行为，这是不一致的。(关于这一思路的发展，见 Arneson 2005；Card 2007；Wall 2009）。

为了对反对规则后果主义的不一致性进行评估，我们需要更清楚地了解最大化善的首要承诺的假定位置。这一承诺是规则后果论者道德心理的一部分吗？还是规则后果论理论的一部分？

那么，规则后果论者不需要把善的最大化作为他们最终的、压倒一切的道德目标。相反，他们可以有如下的道德心理：

他们的基本道德动机是做公正合理的事。

他们认为，按照公正合理的规则行事是公正合理的。

他们还认为，总的来说，规则后果论是对公正合理的规则的最佳解释。

具有这种道德心理--即道德动机和信念的结合--的代理人会在道德动机上按照规则后果论的规定行事。这种道德心理当然是可能的。而且，对于拥有这种道德心理的代理人来说，在遵守规则不会使预期利益最大化的情况下，遵守规则并没有什么不连贯的地方。

那么，即使规则后果*论者*不需要对预期利益最大化做出压倒一切的承诺，他们的*理论*是否包含这样的承诺呢？不，规则后果论本质上是两种主张的结合：（1）规则的选择完全取决于其后果；（2）这些规则决定了哪些行为在道德上是错误的。这就是该理论的全部内容--特别是，并不存在由最大化预期利益这一压倒一切的承诺所构成或意味着这一承诺的第三个组成部分。

如果没有最大化预期利益这一压倒一切的承诺，规则后果论禁止某些行为，即使这些行为能够最大化预期利益，也没有什么不连贯的。同样，规则后果论要求其他类型的行为，即使它们与预期利益最大化相冲突，也没有任何不一致之处。一旦我们意识到规则后果论的代理人和理论本身都不包含最大化善的首要承诺，那么对规则后果论最著名的反对意见就不攻自破了。

针对不连贯反对意见为规则后果论辩护的可行性，可能部分取决于规则后果论的论据是什么。如果规则后果论的论证是从对后果主义评估的承诺出发，那么这种辩护似乎就不那么可行。因为以这种承诺为出发点似乎非常接近于以最大化预期利益的首要承诺为出发点。然而，如果规则后果论的论据是这一理论比其他任何道德理论都能更好地为直观上看似合理的道德规则提供公正的理由，那么对不一致反对意见的辩护似乎要稳妥得多。(有关这方面的更多信息，请参见 Hooker 2005, 2007; Rajczi 2016; Wolf 2016; Copp 2020）。

另一个反对规则后果论的老观点是，规则后果论者必须是 "规则崇拜者"--也就是说，即使这样做显然会带来灾难，他们也会坚持规则。

对这一反对意见的回答是，规则后果主义赞同要求人们防止灾难的规则，即使这样做需要破坏其他规则（Brandt 1992: 87-8, 150-1, 156-7）。可以肯定的是，什么是灾难有很多复杂的因素。想一想，当 "防止灾难 "规则与禁止撒谎规则相冲突时，什么才算灾难？再想一想，当 "防止灾难 "规则与禁止偷窃的规则竞争时，什么算作灾难；当 "防止灾难 "规则与禁止伤害无辜者的规则竞争时，什么算作灾难。在这些问题上，规则后果主义可能需要更加明确。但至少它不能被正确地指责为可能导致灾难。

要避免的一个重要混淆是，不要以为规则后果主义包含了 "防止灾难 "规则，就意味着规则后果主义与最大化行为后果主义在实践上是等同的。最大化行为后果主义认为，只要撒谎、偷窃或伤害无辜者会比不这么做产生哪怕多一点点的预期好处，我们就应该这么做。要求人们防止灾难的规则并不具有这种含义。相反，"防止灾难 "规则只有在预期价值的差异非常大时才会发挥作用。

Woodard (2022) 认为，"防止灾难 "规则需要重新构思。许多规则都确定了原因的种类。例如，防止伤害的理由、不偷窃的理由和信守承诺的理由。但是，伍德尔德认为，我们不应该把 "防止灾难 "规则看作是一个确定了某种原因的规则，然后声称它具有压倒一切的力量。我们更应该认为，"防止灾难 "规则区分了两种情况，一种是防止伤害的理由压倒了对立的道德理由，另一种是防止伤害的理由弱于对立的理由。

即使伍德在这一点上是正确的，"防止灾难 "规则也会被指责为模糊和不确定。事实上，在防止伤害的道德理由*压倒*对立道德理由的情况下，与防止伤害的道德理由弱于对立道德理由的情况下，两者之间的界限很可能是不确定的。沃拉德（2022）认为，承认这种模糊性和不确定性并不会削弱规则后果论。她借鉴了自己早先的研究成果（2015），解释了规则后果主义的 "防止灾难 "规则如何不必使该理论陷入要么要求过高、要么对救助他人的要求施加反直觉限制的两难境地。

9.对规则后果论的其他反对意见

从 20 世纪 60 年代中期到 20 世纪 90 年代中期，大多数哲学家都认为规则后果论经不起上一节讨论的反对。因此，在这三十年间，大多数哲学家都没有去理会对该理论的其他反对意见。然而，如果规则后果论对刚才讨论的三种反对意见都有令人信服的回答，那么一个很好的问题就是，该理论是否还有其他致命的反对意见。

还有一些反对意见试图表明，鉴于该理论选择规则的标准，在某些条件下它会选择直觉上不可接受的规则。例如，汤姆-卡森（Tom Carson，1991）认为，规则后果论在现实世界中要求极高。Mulgan（2001，特别是第 3 章）同意 Carson 的观点，并进而指出，即使规则后果论在现实世界中的含义没有问题，但该理论在可能世界中却有反直觉的含义。如果穆尔根的观点是正确的，那么这就会使人怀疑规则后果论是否能解释为什么某些要求在现实世界中是适当的。关于这些问题的争论仍在继续（胡克，2003 年；劳勒，2004 年；帕菲特，2011 年，2017a，2017b；沃拉德，2015 年：181-205; Rajczi 2016; Portmore 2017; Podgorski 2018; D. Miller 2021; Perl 2021, 2022）。而穆尔根已经成为规则后果主义的发展者而非批评者（穆尔根 2006, 2009, 2015, 2017, 2020）。

对规则后果论的一个相关反对意见是，规则后果论使人们熟悉的规则的正当性取决于各种经验事实，如人性是什么样的，有多少人需要帮助或有多少人有能力帮助。对规则后果论的反对意见是，一些人们熟悉的道德规则是必然的，而不仅仅是偶然的（McNaughton 和 Rawling，1998 年；Gaut，1999 年，2002 年；Montague，2000 年；Suikkanen，2008 年）。这一反对意见的一个兄弟姐妹是，规则后果论使规则的正当性取决于错误的事实（Arneson，2005 年；Portmore，2009 年；参见 Woollard，2015 年：特别是第 185-86 页、第 203-205 页；第 2022 页）。

新准则的教学机制给规则后果论带来了严重的问题，因为规则后果论计算的是让新一代人内化规则的成本。如前文所述，将教导对象限定为新一代人，是为了避免计算现有一代人内化规则的成本，因为他们已经内化了其他一些道德规则和观念。但是，我们能对那些应该对这些新一代人进行教育的人做出一致的描述吗？如果设想教师本身已经内化了理想准则，那么这又是如何发生的呢？如果设想这些教师尚未内化理想准则，那么理想准则与他们已经内化的准则之间的冲突就会产生代价。(这一反对意见由约翰-安德鲁斯、罗伯特-埃曼和安德鲁-摩尔提出。参见 Levy 2000; D. Miller, 2021）。与此相关的一个反对意见是，规则后果论的表述方式尚未使其能够合理地处理规则之间的冲突（Eggleston，2007 年）。但见 Copp 2020; D. Miller 2021; Woodard 2022。