

电车难题之思——浅谈后果主义解决伦理困境之优越性

浙江大学

土水交

徐子安

3200100963

摘要

本文最开始由电车难题出发，简要地阐述了以道德义务论为代表的传统道德理论在一些极端伦理困境下的“进退两难”情形，并引出自己构建的思想范式——“后果主义”。通过解释后果主义基本观点、后果主义的本体论认识、后果主义的决策方式、后果主义对道德等“非后果性”因素的包容机制等方面，描述这一套思想范式的主要原则和特点。作者在最后对后果主义做出反思，并针对其缺陷提出完善和规划。

关键词： 后果主义, 伦理冲突, 利己主义

1 引言

这里讨论的“后果主义”，并非学术上著名的功利主义的演化版本，而只是作者对“结果定义价值”这一原则及其演化思想思想的思考。文章的第三章《何为后果主义》将会详细阐释这种思想的具体含义。

相对于称其为一种“哲学”，我更愿意称其为一种思考的范式。它并不全面、并不能完善地用于指导社会治理的各个方面，而只是我个人尝试设计并践行的一种思想模式。

我在设计这种思想的时候，要求它必须能在任何时候为我提供坚定的选择。我尝试用单一的标准来衡量世界上的价值——以便在遇到伦理困境和矛盾冲突的时候，能够坚定且唯一地帮助我做出选择；我要求它能够融入世俗社会，让我的行为贴近绝大多数人的惯常行为，甚至希望通过它来解释大多数人的行为。经过筛选和选择，不断更改和实践，我最终选择了自己称为“后果主义”的这一套思想范式作为研究目标。

在践行和思考的过程中，我也越来越多地发现这种思想方式的局限性。我也在试图去完善这样一套思想理论。对我来说，这思想的

发展完善过程其实也是我自己认识世界、形成自己价值观的过程。这一点将在论文中（如果它能算作论文的话）有所展现。我先阐述我的理论，然后跳出理论本身来审视自己这些想法的可行性，并思考自己产生这些想法的深层原因。不光呈现在文章的章节安排上（最后一章就是对自己生发“后果主义”思想的反思），在每一节、每一段中，我都将尽力去做这件事情。

我为何会如此“设计”自己的世界观？我相信，如此庞大的思想体系其中必然蕴含了深深的个人烙印和个人风格，在审视其形成原因的过程中，我们能透过很多平时我们对自己主观的误解，看到更深刻的、沉睡在无意识之下的那个自己。

全文手打，无任何复制粘贴。若有不甚严谨之处，恳请读者指出。

2 电车难题：“电车”为何成为“难题”

这是一个大家耳熟能详的思想游戏。

在讨论伦理学问题的时候，“电车难题”往往是用来启发学生思考和讨论的良好范例。作者在此便以“电车难题”为例，阐释我所思考之“后果主义”，与康德所秉持的道德，与大家所秉持的道德，究竟有何差异和优劣。

电车难题的传统情景如下：

一辆失控的列车在铁轨上行驶。在列车正行进的轨道上，有五个人被绑起来，无法动弹。列车将要碾压过他们。您正站在车站内，离改变列车轨道的操纵杆很近。如果您拉动此杆，则列车将切换到备用轨道上。但是，您会注意到备用轨道上只有一个人被绑着。您有两种选择：

- 什么也不做，让列车按照正常路线碾压过这五个人。

- 拉下操纵杆，改变为备用轨道，使列车压过备用轨道上的另一个人。

我们先来解读这个场景的含义。在此之前，我们先做一个约定：在理解像“电车难题”这样的思想游戏的时候，我们不应选择无意义地避开这个思想游戏的矛盾核心。例如，在电车难题的情境下，他需要你在不同的生命之间做出选择，你就不应当给出“通过不牺牲生命的方式”解决问题的策略。这不是一个现实问题，不需要机智的答案，它只需要你对这个矛盾核心本身的思考。

在这个问题中，它预设了一个十分极端的生命选择场景，你需要在“五条生命”和“一条生命”之间做出选择。其他的限制条件诸如“列车”“轨道”之类的要素可以忽略，你也不需要知道为什么六个活人会被奇奇怪怪地绑在铁轨上。你只需要知道，现在裁决者是你，你需要做出你的选择。

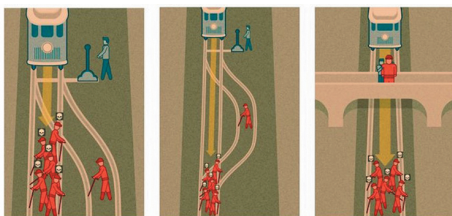


图 1: 电车难题

2.1 经典道德（康德义务论）对电车难题的回答

在课堂讨论中，坚定地抱持义务论观点的同学往往难以做出选择——在这样的理论中，生命应当是绝对的、至高的。任何情况下，直接或者间接导致生命消亡的行为，都是应当受到谴责的。同样你也不能将五条生命和一条生命进行对比——无穷与无穷，没有哪一个更大。

这样的思考矛盾可以这样表述：

- （基本预设）生命是无价的。任何直接或间接导致（无辜的）生命消亡的行为，都是错误的、值得谴责的。

- （情景难题）我们的行为无论如何都会导致（无辜）生命的死亡。
- （结论）我们的行为无论如何都是错误的。

2.2 我们为何不能解决电车难题

为什么我们无法解决电车难题？

电车难题的核心是在拷问我们对生命的看法。你在你的道德体系中，将生命置于一个怎样的地位。生命是否能用数量来简单衡量？当生命遭到破坏的时候，我们如何行动？

电车难题反映出的其实是经典道德义务论在一些极端场景下，自身原本具有的矛盾性。

这套道德伦理本身并不保证自治。“生命无价”的观点，“生命不可被侵犯”的戒律，原本是基于人们的审美、基于人们既有的直觉的道德观念，从诞生之初便不以“解决问题”为根本使命，不保证在任何情况下互不冲突，更不保证在任何情况下都能给出人们“给如何做”的指导。

我觉得“电车难题”是传统道德伦理不完备性的一个小小体现。在更广的维度上，在更多的哲学思想游戏中，甚至是在一些社会问题上，我们可以看到更多这样困境的发生。那些受到热议的、无法被轻易决定对错的实例，往往是因为类似的原因：当双方都有对有错，双方都不应当被牺牲，可是矛盾发生了，我们如何处置这种情形？

我无法忍受这种矛盾和冲突。出于个人的角度来说，我希望自己奉行的思考方式坚定而稳固，时时刻刻为我提供有力的决策参考。这就是为什么我更偏爱以后果来决定价值的这样一套行为方法论——我的“后果主义”。

3 何为后果主义：以后果定义价值

在解释我所说的后果主义的基本原则之前，先来谈谈我对它的一些理解。

“电车难题”本身是道德伦理学中，用来展现功利主义对行为义务论的优势的一个情景。可在这样的情境下，我强调的是后果主义的优越性。我来解释一下其中的原因。

我认为是后果主义，只规定“对‘后果’的无条件追求”这一条戒律。至于什么样的后果是好的，什么样的后果会被定义为是不尽人意的，不是后果主义的规定范围。我更愿意将它看作是更多伦理学规范的一个关节——当你将“实现最大多数人的利益”作为最优的后果的时候，那它就是功利主义；当你将这个主体限制在自己、限制在极少的个体的时候，那它就是利己主义。

我觉得利己主义和功利主义，在这样的情形下都能给出坚定而明确的回答。为什么？因为对后果的明确指向和坚定追求。因此，我将这一部分单独提出来，作为一个单独的关节用来讨论。

3.1 后果主义的基本观点

后果主义者抱持以下观点：

- 一件事情的正确性应当由它的后果来决定，而不是其他任何与后果无关的元素。
- 后果主义者在任何时候都应当坚定地追求，在现有的条件下所能获得的最好的后果。

这两条规定十分明确易懂但其实并不简单。

首先，是对一系列价值的颠覆。第一条规则其实否定了我们习以为常的价值定义。我们在做出个人决策的时候，常常不仅将后果作为唯一的参考因素——诸如道德观点，诸如喜恶，等等等等。

但后果主义的这一条规定消解了所有这样的思考元素——除了“后果”占据最高的地位，除了后果追求的对象（也许是个人或者他人的利益）具有先验的价值，其他任何元素应当平权、平行地纳入考量。

就像是在电车难题的情境下，也许我们最终会做出某样选择，但往往在考虑的过程中，将“生命”本身也作为十分重要的因素纳入考量，带有十分沉重的负罪感或者内疚感进行决策。可是对于后果主义者而言，一样行为，不会因为“符合传统道德”，或者是“因为拯救了生命”而成为正确的行为，而只能因为“对我达成后果有利”而间接地具有正确性。或者说，后果主义者不承认“传统道德”的突出地位，尤其是在“后果”面前

的突出地位。

因此，后果主义者在对电车难题做出回答的时候，应当是毫不犹豫、毫无负担地做出决策——如果我是功利主义者，我也必须选择大多数人的利益，无论五个人在哪条轨道上，我必须尽一切可能去拯救他们，只因为人数上的区别；如果我是利己主义者，我也许会选择袖手旁观，因为这样会造成对我最少的舆论谴责。但无论如何，行为都是坚定的、自洽的，在他们自己的道德逻辑中，自己已经完满地解决了这个问题，并不应当抱负任何负罪感。

接下来，我们将通过一些例子，来深入地理解后果主义的以上观点和一些衍生观点。

3.2 本体论回答：后果主义不关心“同一”

后果主义强调对第一条原则的坚定贯彻。不仅仅在行为的维度上，甚至在认知的维度上。我们对第一条基本观点作如下引理：

- 当两件事情造成的影响相同，我们就认为它们一样。
- 如果事情不产生任何后果，我们就认为它们不存在。无论这件事情是否真的存在。

暂借另一个十分有趣的思想游戏——忒修斯之船悖论，来略作说明。忒修斯之船悖论的背景如下：

这是一则非常古老的思想实验，也称为认知“悖论”。忒修斯是传说中的雅典国王。忒修斯号是以他的名字命名的一艘船，它标志着战争胜利的荣耀。在克里特岛归还时，国王与雅典年轻人搭乘了这艘有 30 个桨的船，后被雅典人作为纪念。

随着时间的流逝，忒修斯号船的木板逐渐腐朽，而只要哪块木板老旧腐烂，雅典人就会立刻换上新的木板来替代。终于有一天，该船的每根木板都被换了一遍。

古希腊哲学家对此发问：“这艘船还是原来的那艘忒修斯号船吗？如果是，但它已经没有最初的任何一根木头了；如果不是，那它是从什么时候开始就不是了？”

哲学家托马斯·霍布斯将这个问题弄得更加复杂：如果有人将换下来的老旧配件收集起来，再做成一艘船，与那艘由新配件组

成的“式修斯号”相比，哪一艘才是真正的式修斯号呢？

在课堂对这个问题进行讨论的时候，大家对这个问题的理解仅仅局限在“是否同一”，以及“为何同一”。在大家没有抱持相同基础观点的情况下，其实讨论者之间并没有办法彻底理解和接受对方的观点。后果主义便提供了一个这样的基础，并且简单而明确地解决了这个问题。

后果主义不关心同一。因为“同一”这个概念本身不造成任何影响，“同一”从后果的视角来看完全可以由非本质性的“无穷相似”来替代。我们认为无限接近和同一是一样的，或者更明确一点，直接从理论体系中删掉“同一”这个东西。

如果我是乘船者，我只需要关心这两艘船是否具有相同的航行性能，如果在这个层面上一样，那就是一样的；如果我是考古学家，我只需要关心这两艘船是否都有我需要的历史遗留，既然没有，那显然不一样；如果我是想把它做成旅游景点的投资人，我只需要考虑换了木板是否会对景区的火爆程度造成影响。不论在哪种情形下，我认为只要仿照到了极致，那就是一样的。

这不是一种退避，对问题视而不见，并不是忽略了或者避开了问题本身。后果主义已经给出了相当后果主义的回答。

在这个问题中，后果主义不同于其它思考方式的地方在于，它选择舍弃这种理想的“绝对”。这种“舍弃”，对无必要无意义概念的舍弃，对“本质”本身的否定，使得它在解决许多伦理困境下展现出强有力的优越性。我认为这是一种创造性的策略，它是后果主义解决问题优越性的重要来源之一。

“沙子问题”也同样可以得到回答：

一顿沙子堆在一起，是一顿沙子。现在你从里面拿掉一粒，它还是一堆沙子。你不断地拿走一粒一粒，请问在被拿完之前，什么时候它不是一堆沙子了？

对于这样的问题，后果主义同样保持相同的回答：当这个概念对我们理解沙子无意义的时候，就可以将它删去。“一堆沙子”是一个概念，指向某个数量范围内的沙子。它能够存在的前提，就是说话者和倾听者不关心什么时候它不是一堆沙子。

在另外一些场景下，我们可以选择其他恰当的描述方式：无论什么时候，沙子的集

合都必定具有确定的质量、确定的数目，会造成确定的后果。当我需要知道它的重量，我就去称重；当我需要知道它的数目，我就费劲去数。但在任何情况下，它“是不是一堆沙子”都没有任何影响。买沙子的人会关心这里有多少吨，但绝不会关心着是不是一堆沙子。

后果主义者抱持这样的观点：概念只是为了我们更好地描述这个世界。当概念本身成为了解世界和描述世界的阻碍时，我们就应该放弃它们。后果主义否定“概念”本身的价值。反对将价值与“概念”本身直接绑定，而选择在需要的时候，用解构的、用精确的方式来描述事物。沙子什么时候成为一堆不重要，多重、多少粒才重要；就像行为是否符合道德不重要，是否能获取想要的后果才重要。

3.3 后果主义：“有限决策”原则

这样的后果主义显然地会引起很多质疑。

在功利主义的受到的批评之中，较为严重的一个问题是，你如何计算最大的幸福，如何计算最大的利益。在遇到复杂的情境的时候，我们必须要做到最好，意味着必须全知全能。而这是不可能的。相同的批评我也会发生在后果主义身上。

我的后果主义第二条基本观点，对此问题做出了修正的回答——“有限决策”原则：我只做我认知中最好的决策：

混沌便是相等，未知便是平权。我只选择我视角中当下能做到的最好。

计算机解决问题有一个策略，叫近似最优解。当你没有办法在整个过程中得到最优解的时候，没有办法找到代价最小的那条路，那就在走到一个地方的时候，做出当下最好的选择。也许“当下最好”走完的路径代价是“最好路径”的百分之一百二十、百分之一百五十，但你花费的算力可能是它的千分之一甚至是万分之一。

在我受限的视角中，我不在乎如何决策是事实上的最好。在我的视角中，根据我现在掌握的信息判断，这样的后果是最好的，那我就这么办。事实上究竟是不是，我不关心。

例如，我需要为天下大多数人谋福祉，我现在可以选择医生或者老师这两个职业。当然它们都是前途远大的职业，都能给他人带

来光明和希望。我不知道哪一个会带来更多的福祉，我可以做更多的事情，这两个选项的价值对我来说是混沌的，那我就随便选。因为在大多数情况下，在混沌的问题上投入思考是低效的——你永远无法精确预测世界会如何变化，而根据混沌的蝴蝶效应，任何一点小小改变，都会使得你的预测完全无效化。有这个工夫，你不如思考如何把自己选择的职业干好。

我想，“没有达到最优”不应当成为批评理论的原因。如果想达到“最优”，你应该去问科学，而不是哲学。没有哪一种哲学体系可以帮助你任何情形下做出最优的选择——后果主义在努力地这样做，它在试图排除掉一切干扰性的因素，排除掉一切阻挡在科学决策面前的障碍，我相信，相比其他的行为范式，后果主义做得更好。

4 后果主义包容性：套接利己主义为例

后果主义不应当是“反世俗”的或是“激进”的。成功的后果主义者并不总是看上去和他人格格不入。

我相信后果主义是始终存在于人们思想和决策中的思想范式，只不过贯彻的程度有所区别。在这一章节中，我将以利己版本的后果主义为例子，阐述它是如何与我们在生活中惯常遵循的那些准则完美地融合在一起的。

4.1 利己的后果主义：强大包容性

后果主义本身具有十分强大的包容性。的确，它从根本上蔑视除了后果（在这里是自我利益）以外的所有价值，从本质上否定了“生命”“道德”等元素的价值，可这并不代表它倡导人们不遵循传统道德。

事实上，成功的后果主义者应当能够很好地融入社会，在各方面做得突出而出色，而不会遭到他人的非议，即使这个人是后果主义者中的利己主义者也一样。

利己主义者如何理解道德？就和理解杯子理解水一样，它无非是无形无质的存在，但存在便必有因果、必有影响。利己主义者这样认知道德：

道德是重复社会博弈的最优解。道德是通行的社会价值交换准则。道德是人与人之间价值交换的契约。

后果主义者的唯一目的，只有后果本身。他并没有将“传统道德”作为仇视的目标，而只是去除其中规定“不可违反”的那部分，将它当做一个事实存在来看。

它具有方法论意义：人世的本质是混沌的。即使是利己主义者也不可能通过纯粹的计算方式来获取最优解。而往往，对公认传统道德的“出格”会导致不良的后果，绝大多数情况下，“符合传统道德的表现”会让你的结果更出色。

它具有“认证”意义：商品的价值有交换必须借由货币作为尺度，理性的经济人之间的交换必定需要一定程度的诚信作为保障，而大多数人进行情感和物质交换的前提便是“这个人是否在一定程度上具有道德”。无论道德本身是怎么来的，是否正确（是否正确在后果主义看来是不需要讨论的无意义内容），它至少对人们的生活产生了影响，具有必要的利用价值。从这个角度，利己主义者应当选择遵循道德。

你可能会问，那是不是代表，这个人在不与人接触的情形下，就可以不遵循道德呢？答案出乎意料地，是否定的。他人观测你是否遵循道德，并不是观测你是否在某一次、某一地遵循道德，而是“是否有违背道德”的行为。如果你将人前和人后的自己分为两种状态，彻底地将道德作为一种伪装，你的行为将很容易被戳穿。

因此，利己主义者应当倡导对道德的近乎绝对实践。从行为上看，除了在一些极度特殊的情况下（例如电车难题或是面临重大选择的时候），利己主义者应当表现得跟“心奉道德”常人无异。

利己主义并不是仇视道德，相反地，它通过十分具有后果主义色彩的方式，十分逻辑的方式，在试图赋予道德一种“非先验”的地位。而在后果主义者的眼中，只要去坚定地执行这种行为，无所谓其动机是否是“道德的”，他便是实现了道德。

再比如与“逻辑”似乎并不相容的美学价值，也能通过相似的机制融入到后果主义者的生活中。或者说，对于一些在其他学说中被“先验定义有价值”的这些事物，比如说道德和美，后果主义在以一种“逻辑”的

方式解释其意义。譬如你无法用道德义务的理论，去逻辑地回答“为什么我要遵守这些道德”这样的问题，但是这样的问题可以在后果主义的学说中找到答案。

5 后果主义观察小记：自我反思

(如果这不适合加入论文里，我会把它删掉。)
我要开始胡说八道了。

5.1 “追求走向反面”的迷思

我一直在疑惑一个问题。

世界上有没有什么东西，是必须纯粹，不能用后果来度量的？

我所找到的结果，最可能接近的结果，是爱和人性的。也许它们必须要被无条件地相信，才能展现出自己的价值。当“爱”不再被无条件地接受，而必须与“满足孤寂的需要”和“向往其中的美好”这样的动机挂钩，它可能就成为了永远追求不到的东西。

后果主义的解构太过于强大，我越来越无法接受相反的逻辑：我愿意和你待在一起，不是因为需要什么，不是因为要你给我什么，它毫无缘由、毫不在乎影响，仅仅是存在，所以有理由，如此而已。

当你坚定地秉持后果主义，将一个人的福祉作为最高的要素，通过适当的调控和配置，或许你能拥有最好的人缘，拥有最好的财富和能力，但你一定摆脱不了寂寞。后果主义者会需要一种热烈的激情，需要那个目标一直引诱你向前冲刺，当他停下来看这个世界的时候，只能看见无底的空虚和寂寞。

我一开始给自己制定的后果，究竟是否正确？我选择解构一切，不相信一切，追问一切，究竟是否正确？在某一个瞬间，我想换一种思考方式，我不需要最大化的结果，甚至不需要计算到的最大的幸福和快乐，我只想单纯地相信些什么，也许这样就已经足够幸福了。

5.2 后果主义的出身背景

这思想有非常浓重的个人色彩。

我被训练成为一个优秀的掠夺机器。出身于平凡得不能再平凡的家庭，父母终日为生活琐事所忙碌。我从来没有被教育过“寄予”，我无法理解“寄予”和纯粹的利他，无法理解纯粹的美好，我理解世界的运行本质为“交换”。

我习惯性地“获得”作为最大的目的。后果主义天生默认“获得”为最大的义务，没有给出具体而平等的论证。后果主义天生是一种工具，我反复强调它是一种思维范式，而不是一个具有指导性的世界观。从电车难题那里开始，我选择祛魅生命的价值，选择祛魅道德，最终究竟丢弃了什么？我不知道，我也很惶恐。

在我看来，后果主义的思维方式更像是一种狭隘的世界观的产物——被生活本身的烟火气所限制，被欲望和贪婪限制，尤其是它们被美化为“积极和进取”之后——像是阶级跃迁或者是生活不易的牺牲品。

我从小被认为具有成熟的思想。我不喜欢玩，不喜欢很多别人喜欢的东西，而会希望自己比别人做得好一点、看得远一点。毫无理由地。

这是一种一无所有的害怕。对将来的不确定的害怕，对自己生活不稳定的害怕，对人生本质是否虚无的害怕。（所以我一直在强调这种思想范式必须帮助我做出决策）这是一种全阶级甚至全社会都共有的情绪，驱使着我们不断去掠夺，去获得更多的东西。我只是这股洪流中的小小一者，后果主义不过是诸多精致利己者的一个托辞而已。

我在这里遇到了一些我曾经无法想象的思想。也许并不是所有人都害怕失去，并不是所有人都在如此强烈地寻求安全感。总有人有权利或者是有想法，去诗意地栖居。说实话这使我十分震惊。

我觉得，后果主义不应该是我的希望。至少在被生活的重压击倒之前，在社会污浊侵扰之前，我想明白什么是真正的幸福和快乐。

5.3 后果主义完善规划：“阶级化”的后果主义

后果主义最大的问题在哪？它为了逻辑的连贯性和坚固性，选择无视个人自己的情感需求，在本质上选择否认“美”和“爱”的地位。

我想，是不是有可能，我们在划定一个固定的区域，让它不受这种范式支配，或是更少地被这种范式支配。我觉得这样的情况可能更适合具有正常情感的人类，其实大多数人往往在某个固定的圈子里（例如面对自己的挚友，自己的家人）和圈子外，的确对“结果”的追求程度不太一样。

我将尝试如下更改这套理论：以某种方式划定不同的阶级区域，在其中践行不同程度的后果主义；在选择至高价值的对象的时候，并不局限于选择自己或者选择天下所有人，而是去选择一个与自己情感密切相关的对象，也许是家庭，也许是爱人。

致 谢

这后果主义也许有凡此种种的问题，亟待我去完善和解决。但我相信其中的一条是应当被践行和保留的思想原则：抛开他人给予的观点、自己的成俗陋见，我们应当无歧视地观照每一个思考本身。

陈炜老师的讲课过程给了我很大的启发和鼓励，见到了哲学的自由与深邃，让我爱上了浙江大学的通识课。在我的眼里，她始终秉持一种开放包容的态度，耐心细致地对同学不甚完善的想法作出指点。

“哲学思考必须要有现实意义吗？”“功利主义甚至是利己主义应当被怎样看待？”对这类问题，她从来不会明确给出一个“正”的结论，而是鼓励我们从零、从无偏见开始思考这些问题，这使我大受鼓励。往往很多教授都不会这么干，他们负责的是“将自己的思想印入学生的脑海”，而非更让同学们“自由地开出自己的智慧之花”。

论文仅为个人的小小浅见。您见笑了。（这句话在正式论文中将会删去）。