

经典决策论的规范性问题及其回应

宋子明

摘 要: 决策论哲学中的一个重要问题是如何在规范层面上理解经典决策论。即使作为形式工具, 自创始之初它就被认为符合人类理性选择的某些基本原则, 被用来说明和预测选择行动。这一决策论的心智主义解释预设理论元素具有心理实存对应物。当代对心智主义预设的质疑认为, 它误解了决策论的形式理论。同时, 决策论的规范维度也受到行为经济学等实证研究的挑战。然而, 通过进一步分析决策论的规范性刻画问题, 本文辩护一种从规范层面刻画经典决策论的立场, 旨在表明规范性决策论是评估选择合意性的一种“逻辑”。

关键词: 合理性原则; 经典决策理论; 心智主义; 规范性

中图分类号: B81

文献标识码: A

1 引言

哲学的发展经常从科学理论中汲取养分。决策论 (decision theory) 作为当代社会科学中的重要理论之一, 对我们理解人类行动既提供了素材, 也颇具洞见。经典的决策论又称“期望效用理论” (expected utility theory, [22, 23]), 形成之初便具有实证和规范的双重色彩, 既是经验描述性质的科学理论, 同时, 其中的关键概念 (如效用、偏好、选择) 和其对合理性原则的刻画, 使得该理论可以被视为具有规范内容。哲学界对经典决策论规范性问题的讨论, 立足实证科学而探讨规范问题, 有助于理解合意的选择遵循怎样的规范原则。

决策论的规范性刻画主要面临两大难题。其一, 是来自工具主义解释的挑战。([8, 15]) 规范性刻画依赖心智主义 (mentalism), 认为理论中的概念对应心理状态 (信念、愿望和动机), 因此理论可以表征意图行动者的选择行动。但工具主义认为这是对理论的误解, 否认早期的决策论使用者普遍采纳的期望效用最大化主张, 即理论为行动的合理性所提出的规范主张是“应该最大化期望效用”。可是如果决策论不反映心理状态, 也就难以帮助理解选择的合意性规范。其二, 决策论也面临来自心理学和行为经济学上的实证反驳。证据显示, 人们实际所做的

收稿日期: 2023-03-23; 修订日期: 2023-07-23

作者信息: 宋子明 四川大学哲学系
thatzmsong@126.com

选择与理论预测存在系统性偏差。([2, 6, 11, 19])然而对于规范地解释决策论而言,实证意义上的失败并不自动排除规范的解释。第二节将提出一种以心智主义为基础的规范性决策论。通过回应两个难题(第三、四节),本文进而辩护并说明这种规范的解释怎样帮助我们理解一种评估行动合意性的普遍规范(第五节)。

2 经典决策论及其规范之源

戴维森(D. Davidson)认为,作为理性人的我们通过常识心理学(folk psychology)看待彼此和自己的行动,并以此实现理解。([3])这是一种基于工具-目标合理性并诉诸愿望和信念的理解:如果我有和你一样的愿望和信念,那么我会做出和你相同的决定;当我具有某愿望和信念,那么在某种必然的意义上,我会——并且应该——那样行动。简言之,常识心理学是合理性原则规范理论的起点。

20世纪中叶,社会科学家将决策论发展为一种刻画选择行动的公理化形式理论。作为刻画行动的理论,决策论不是常识心理学的替代品,相反,它是一种以之为起点并精确化、规整化的理论。刘易斯(D. Lewis)很好地阐述了决策论的起源和它所扮演的规范角色:

决策理论(至少如果我们忽略那些修饰)不是艰深的科学,无论在局外人看来是多么陌生。相反,它系统地阐述了关于信念、愿望、偏好和选择的某些精挑细选的陈词滥调。它是我们关于人的常识理论的核心,被剖析并优雅地系统化。([12], 第114页)

常识心理学可以描述(包括说明、预测)一个理性人为什么会做某个选择,也可以规定(包括建议、评估)行动者应当做何选择。而因对常识心理学的精确化,决策论也同时具备了描述和规范双重维度。这一传统理解为早期的决策论使用者所采纳,且理论的初创者们也默认了决策论的双重属性。

本节阐述一种对期望效用理论的规范解释¹,突显常识心理学为理论提供起点,并以此引出传统解释中的一些问题。阐述分为两个阶段,首先考虑确定条件下的简单决策,然后处理不确定的情况。第二阶段强调偏好的更新,彰显从对确定结果的偏好到对不确定选项的偏好的更新过程。

2.1 决策论与愿望/信念

首先,拥有愿望的理性主体关心结果(outcomes)。用命题表征结果,例如“我吃这个苹果”,“我在海南岛度假”,“我的朋友被拘留”。结果可以是好的或坏的,它们是理性主体可以对之进行价值评估的对象。理性主体对其想要结果的评

¹理论的形式模型参考[10, 14, 16, 18]。

估, 通过其偏好序列 (preference ranking) 来体现。该序列反映主体在综合考虑所有因素的情况下, 会选择不同结果的先后顺序。形式上, 设 Z 为结果集, 偏好序列是定义在 Z 上的二元关系, 用 R^* 表示。 R^* 要满足弱序的一致性要求, 即满足以下两个条件。

完备性: 对于每个 $x, y \in Z$, xR^*y 或 yR^*x 。

传递性: 对于每个 $x, y, z \in Z$, 如果 xR^*y 和 yR^*z 那么 xR^*z 。

随后, 决策论中通常再定义一个效用函数 (utility function, U), 从数字上表征偏好序列。 U 分配更大的实数给更好的结果。亦即, $U: Z \rightarrow \mathbb{R}$, 使得 xR^*y 当且仅当 $U(x) \geq U(y)$ 。这是决策论如何开始将常识心理学的愿望部分形式化。信念部分的建模则基于主体认为可行的一集行动。信念函数为每个行动赋值对应结果。形式上, 设 A 是主体认为可从中选择的一集行动, $B^*: A \rightarrow Z$ 表征主体关于行动如何与结果关联的那些信念。

然后, 关于主体如何选择, 决策论的基本思想仍然是常识心理学的: 在主观认为的会导致不同结果的可选行动范围内, 理性主体应该 (而且经常会) 选择她认为最能满足她的偏好的行动。形式上, 主体应该选择某行动 $a \in A$, 使得对所有 $b \in A$, $B^*(a)R^*B^*(b)$ 。或者用效用函数表达为: 主体应该选择某行动 $a \in A$, 使得对所有 $b \in A$, $U(B^*(a)) \geq U(B^*(b))$ 。

以上是确定情况下的决策, 尚未引入概率处理不确定情况, 故理论中的偏好和信念关系都暂时带 “*”。而我们已经看到, 理论中的 R^* 和 B^* 分别对应常识心理学的愿望和信念, 由两者共同导向行动。在此意义上说, 决策模型是对常识理论的精确化。描述维度上可以说, 某人选择了——会选择——行动 a , 是因为, 根据她当下的信念, a (将会) 导致她的价值评估中的最好结果; 规范维度来说, 她应该选 a , 依据 “ a 使得对所有 $b \in A$, $B^*(a)R^*B^*(b)$ ”。

面向决策论使用者的教科书往往从这样的对应关系出发, 开始介绍理论元素的建构。直观上也因贴近常识而更易理解。然而, 这预设了在当代被称为 “心智主义” (mentalism) 的解释, 受到了以哲学家奥卡沙 (S. Okasha) 为代表的批评。([8, 15]) 事实上, 期望效用模型源自应用数学领域中的测量理论 (measurement theory), 并由经济学家们发展应用到社会科学。在冯诺依曼和摩根斯坦 (J. von Neumann, O. Morgenstern)、萨维奇 (L. Savage) 等初创者看来, 理论出发点并非对其中参数的某种解释; 理论是由模型参数和在模型下证明的定理所构成的形式工具。但社会科学之所以能够使用这一数学工具对人的选择偏好进行说明和预测, 仍然需要依赖某种解释桥介。

2.2 表征定理

上述第一阶段关于决策论如何刻画愿望/信念理论对于信念部分的处理不够灵活。即使从主观上，一个理性主体也很少能确定地知道她的某个行动会导致什么结果。一般来说，一个给定的行动可能会导致许多可能的结果。为了处理这种风险情况，决策论诉诸于概率。理性主体有一个概率函数，来表征她对选项行动产生各种结果的可能性的信念。形式上，设 L 是一组前景 (prospects)，有时称“乐透”或“赌博”，它将结果集 Z 的元素作为具有特定概率的“奖品”。 L 包含了 Z 上的所有概率分布。本文用小写的 p, q, r 等来表示 L 的元素。每个概率分布将结果集 Z 映射到区间 $[0, 1]$ ，且所有分布的值加起来为 1。换句话说， L 中的概率函数 p 定义为 $p: Z \rightarrow [0, 1]$ 使得 $\sum_{z \in Z} p(z) = 1$ 。选择主体所面临的每一个风险情况，都可以理解为这样一个前景，由 p 表征。

第一阶段有对确定结果的偏好排序 R^* 。当选择行动与确定结果紧密联系时， R^* 会延续。在第二阶段，行动与前景相联系，因此，我们需要对这些前景形成偏好排序，用 R 表示。根据决策论中的著名结果“表征定理” (representation theorem)，存在这样的偏好序列 R 。具体地，围绕对结果的偏好序列 (关系 R^*) 和由概率构建的前景，存在一集公理，允许我们形成对前景的偏好序列 (关系 R)。这些公理以及从它们导出的引理可以以不同方式呈现，不同的公理系统有技术细节上的差别，例如 von Neumann-Morgenstern 理论的表征定理和 Anscombe-Aumann 理论的表征定理在形式上略有不同。本文忽略不同表征定理形式上的不同，仅以冯氏理论作为正统理论的一个代表。对与本文有关决策论规范内容的论证来说，重要的是对表征定理的解释。

表征定理说：由一个效用函数 u 出发，表征对结果的偏好序列 R^* ，结合一个定义良好的概率函数 p ，当它们满足一集公理，则能够证明，存在一集效用函数，并且它们经线性变换后是唯一的，该函数期望地表征对前景的偏好序列 R 。定义主体的信念函数 $B: A \rightarrow L$ ，在这一阶段， B 将每个可选行动赋值一个前景选项。有了定理对 R 的保证，决策论对理性选择的刻画也就顺理成章：理性主体会选择行动 $a \in A$ ，使得对所有 $b \in A$ ， $B(a)RB(b)$ 。更精确的量化表征为：一个选择行动的效用，是由该行动可能获得的每个前景的效用经概率加权后的总和，数学上称“期望效用”。例如，某行动会以概率 r 导致结果 x ，否则导致结果 y ，那么该行动的期望效用可计算为 $xr + y(1 - r)$ 。一般来说，行动 p 的期望效用是 $u: L \rightarrow \mathbb{R}$ ，使得：

- $u(p) = \sum_{z \in Z} u(z)Pr(z|p)$ ，且
- pRq 当且仅当 $\sum_{z \in Z} u(z)Pr(z|p) \geq \sum_{z \in Z} u(z)Pr(z|q)$ 。

这里，函数 $u(p)$ 期望地表征了主体对前景选项 p 的偏好 R 。在 p 的期望效用大于

q 的情况下, p 是一个比 q 更好的选择。由此说, 表征定理对选择偏好的表征是期望的 (expectational)。虽然有表征定理参与的第二阶段在形式上更复杂, 但其基本想法仍然是常识心理学的。

3 来自工具主义解释的批评及回应

经济学家对期望效用模型的解释, 通常认为是工具主义 (instrumentalism) 的。模型本身来自测量论, 定义一些偏好相关参数并推导出决策的期望效用。经济选择上的应用则建立在对行动主体真实偏好的观测上: 通过向主体分别展示多组前景, 邀请主体比较对它们的偏好; 所表达的偏好被用来构建“标尺”, 进而得以测算主体的效用函数 u 和概率函数 p 。而理论中的关于偏好的公理保证了这样的测量方式对于偏好的期望表征, 是充分且必要的。经济学家瓦克尔 (P. Wakker) 在其决策论教材中, 称此为决策论的“行为基础” (behavioral foundations, [20])。而奥卡沙也使用“行为主义” (behaviorism) 指这里的工具主义。但正如奥卡沙也注意到, “行为主义”一词容易和萨米尔森 (P. Samuelson) 的行为主义经济学混淆。([7, 17]) 因此本文称“工具主义”。 U 和 p 来自对偏好态度的实证测量, 由此看出, 理论参数只不过是一种方便的数学工具, 用来描述一个“好像” (as if) 的故事。主体依其偏好做决策就好像是她在计算着一个期望效用函数的值。理论的描述并不预设效用、信念 (度) 等概念具有心理的实存内容。行动 a 的期望效用值大于 b , 并不从心理实存意义上因果地说明为什么某人更偏好 a 。期望效用的计算与偏好态度并不构成因果关系, 只是表征关系。

本体意义上, 上述工具主义的解释承诺一种反实在论, 不同于心智主义所认为的效用函数对应着某些实在, 例如愿望强度和满足程度。在偏好和效用函数的关系问题上, 两种解释都属于“建构主义” (constructivism), 认为 u 只是理论参数, 由主体的偏好态度所决定。相对地, 强实在论则认为效用是独立于偏好的心理状态。在建构主义的两种解释下, 我们都可以既谈理论通过 u , 怎样描述偏好, 也谈理论怎样规范偏好。甚至在强实在论下, 我们可以谈偏好怎样通过效用而形成。既谈效用怎样说明和预测主体具有的偏好, 也谈效用怎样要求主体应当形成怎样的偏好。

近年来, 随着讨论决策论的哲学家和将之作为工具应用在其他哲学问题的学者逐渐增多, 对理论的解释也不尽相同。例如政治哲学家高契尔认为, 后果主义的幸福最大化原则与期望效用最大化原则是一致的。([5]) 哲学家杰克逊认为, 期望效用最大化的主张是一种后果主义的道德选择理论。([9]) 又如奥卡沙列举的诸多例子表明, 当代哲学界习惯采纳心智主义的解释来赋予决策论以规范内容。德莱尔 (J. Dreier) 也指出了类似问题。([4]) 当代哲学界似乎普遍认为, 决策论通

过表征定理主张“理性主体应该最大化行动的期望效用”：

表征定理经常被用来为决策理论奠基。首先，该定理表征信念度和效用。其次，该定理辩护两个理性的基本原则：我们应该有符合概率的信念度，并且我们应该按照期望效用最大化的方式行事。（[13]）

我们先看对期望效用最大化主张的批评。在经典决策模型中，表征定理的主张是：

某主体的偏好（与由之定义的 u 和 p ）满足了一集公理（弱序、独立性、连续性），当且仅当，该主体的偏好可由一集效用函数期望地（亦即，期望效用最大化地）表征。（1）

命题 (1) 在逻辑上独立于、且严格地弱于下述命题 (2)

主体的偏好应该最大化期望效用。（2）

根据奥卡沙的论证，即使命题 (2) 为假，命题 (1) 也可以为真。即使主体的行动实际上违反期望效用最大化，她的偏好关系依然可以满足理论公理。而一旦满足了公理，那么她的效用函数就是期望地可表征的。因此，完全有可能的是，要么该主体的效用函数与其真实心理意义上的效用函数不同，要么她的概率函数与其心理的信度函数不同，或两者皆是，而两种情况下，她的偏好关系都完全可能符合公理。信念方面，当主体的概率函数和偏好关系符合决策论公理时，其真实信度函数可异于该概率函数。（[21]）换言之，表征定理中出现的 u 和 p ，只是便利的数学工具建构。在主体的偏好满足公理的条件下，仅能证明存在某个 u 作为偏好的期望表征，但是理论并不保证那个 u 与主体真实特定的心理状态有关，因此更不保证理论中构建出来的某些效用和概率函数与主体那些真实的愿望和信念一一对应。

奥卡沙论证的部分结论是合理的，即“偏好应最大化期望效用”这一主张是对决策论规范内容的误解。但他接着关于理论所支持规范内容的推论却不够严谨。奥卡沙认为，决策论的规范主张是“偏好应当满足决策论的公理”，且该规范内容排斥对理论的心智主义解释。需要注意的是，从其文显示，奥卡沙的“心智主义”承诺了强实在论，认为效用是独立于偏好且不可还原的事实。若将对行动的描述和规范，归因于独立实存的愿望等心理状态，这样的常识心理学理解也承诺强实在论。但按照本文已经界定的划分，心智主义完全可以只承诺建构主义的弱实在论。因此，奥卡沙对“心智主义”的排斥，实际上只是排斥了强实在论解释。

而且，以奥卡沙为代表，将“决策论并不承诺某些先于理论而独立存在的效用和信度”这一工具主义解释直接与“偏好应当满足公理”这一规范主张挂钩，有

捆绑工具主义之嫌。奥卡沙并未直接辩护工具主义，只是一笔带过其强实在论版本所面临的困难。决策论经典文献中（[10, 14, 16, 18]），工具主义也通常视为默认选项。这样的做法给人误导性的印象，似乎决策论的规范内容来自于对理论的某种解读。但这颠倒了逻辑顺序，理论的内容属于理论本身，应先于对其的解读。相反，如果我们没有从认定工具主义出发，那么奥卡沙的论证只说明，具有心理实存意义的效用、概率函数的存在，对满足理论公理不必要。这为心智主义留下了空间。因为即使心智主义下的效用、概率函数不必要，也不意味着它们必须不存在。一种混合式解释是开放的：将工具主义和心智主义的部分真相结合起来，刻画决策论的规范内容。

4 来自描述与规范维度不相容论的批评及回应

虽然哲学界对决策论规范内容存在一定误解，但该现象有其充分的原因。心智主义解释的倾向，源自常识心理学合理性原则。虽然社会科学家考虑决策论是为服务于描述性地说明和预测行动，哲学家对决策论的兴趣则主要来自它的规范性。正如历史上亚里士多德、休谟及当代学者对实践理性的关注，哲学家们更关心理性人应遵循怎样的选择规范，而非对选择过程呈现机制的描述。

决策的规范维度，指的是对某行动是否合理的评估判断，包括“厚的”和“薄的”理性评估。外部评估通常是厚的，从评估者视角考虑如，决策问题的定义（选项和结果的认定、信息的收集）、所选行动在道德价值意义上的辩护等。而薄的内部评估是一致性评估，从行动者视角，只考虑所选行动与所想目标是否内在地自洽。因此，规范维度既包含外部评价，也包含内部一致性的评估，都是对选择是否合理的评估判断。

贝穆德兹（J. Bermúdez）讨论理性在不同学科领域的不同维度，及决策论能否为这些不同诉求提供一个统一的理论基础。（[1]）关于描述与规范两个维度能否统一的问题，他分析了相容论与不相容论两种立场。不相容论者认为，规范维度必须建立在决策论公理被满足的基础上，而描述维度则蕴含对公理的违背，因此两者不相容。近数十年来，行为经济学和心理学对人的非理性现象的研究显示，人们实际的行动在很多方面广泛系统地偏离了经典决策论的描述，真实的选择和动机几乎是而非理性的。（[2, 6, 11, 19]）如要决策论发挥其规范作用，则意味着公理必须被满足，但也因此不具备描述准确性。而相容论者持反对意见，认为对决策论描述准确性的否定，独立于其规范内容。

在区分了描述与规范维度的基础上，对于决策论是关于过程（process）或结果（outcome）的理论，我们进一步做出解释上的区分。首先，作为关于过程的描述，决策论描述人们真实的选择依据怎样的机制。例如心理学上描述部分学生写

论文拖延的心理机制。而根据决策论,想要获得好成绩又相信论文成绩很重要的学生,不会延迟论文写作。把决策论解释为关于过程的描述是不合理的,人们真实的心理过程与决策论的描述不符。第二种区分是把决策论解释为关于结果的描述。例如经济学上关于企业总是以最大化利润的方式选择,但有时企业也为了追求社会公平等价值而放弃利润最大化。决策论对于结果的描述性预测有时也与事实不符。因此,合理地解释经典决策论需要离开描述维度。

在规范维度讨论决策论,同样可以区分关于过程或结果的两种解释。作为关于过程的规范理论,决策论扮演规定(prescriptive)选择的作用,主张行动者应当怎样实现效用最大化。而作为关于结果的规范理论,决策论只扮演“答案”的角色,主张理性主体应当以怎样的理想结果为标准。²规定过程的决策论需要预设其与真实选择至少大致相符,毕竟“应当蕴含可能”。但作为理想标准的决策论则不要求现实的准确性。因此实证证据对这种解释并不构成挑战。

事实上,对于决策论的两种规范解释,在哲学文献中常常是模糊的。贝穆德兹虽然看似区分了两者——将规定过程的解释定义为行动者第一人称的内部评估,但却将理想标准的解释局限为厚的外部评估。第三人称描述视角须以满足公理为前提,否则理论无法适用。而对于第一人称的内部评估规范,如果将这种规范理解为对行动的规定,那么同样受限于理论的现实准确性和适用条件。但它们都不是对规范性决策论的唯一解释。

规范性决策论是过程规定,还是理想标准这一区分,可通过表征定理的两种陈述方式显示。回顾第二节的命题(1)发现,该条件句的前件(主体的偏好是否满足决策论公理)是或然的。而根据表征定理,一旦偏好满足了公理,那么偏好和效用函数可期望地表征,这是逻辑必然地。换言之,通常认为,表征定理主张的是下述“狭域”命题,即使我们不预设工具主义解释:

狭域命题 如果某主体的偏好(与由之定义的 u 和 p) 满足一集公理,那么,必然地,该主体的偏好可由一集效用函数(好像,或真实地)期望地表征。

可以看到,表征定理是否适用,需要建立在前件公理被满足的基础上,不论是描述维度或过程规定的解释。按照“应当蕴含可能”原则,如果证据显示人们几乎不可能遵循表征定理的前件,那么该定理也几乎无法用来规定行动。但是,我们可以把决策论的规范内容解释为一种理想标准,也就是说,当我们赋予表征定理以广域的理解,此时不仅规避了实证证据对决策论描述不准确的批评,获得描述与规范维度相容论立场,而且提供了一种理性选择的普遍准则:

广域命题 必然地,如果某主体的偏好(与由之定义的 u 和 p) 满足一集

²当然这里的理性仅指薄的意义上一种工具-目标合理性。

公理，那么该主体的偏好可由一集效用函数（好像，或真实地）期望地表征。

5 一种“评估的逻辑”：决策论规范性刻画

第二节阐述了一种对经典决策论的规范解释，凸显怎样从公理系统内部将愿望、信念和行动联系起来。通过第三节分析看到，决策论对心智主义解释是友好的：愿望和信念分别由效用函数和信念函数表征，行动的选择价值则以行动的期望效用为指标。这使得“决策论提供理性选择的规范”这一解释成为可能。通过第四节对实证挑战的回应，我们区分出一种关于结果而非过程的规范解释。出于这种关于结果的规范解释，亦即对表征定理的广域理解，本节将进一步阐明决策论的规范性刻画：与演绎逻辑对真值的更新、归纳逻辑对信度或证据支持度的更新相类似，这种规范性决策论是关于理性主体更新偏好的普遍规范。

表征定理表征主体怎样由“输入”其对结果的偏好序列 R^* ，到“输出”对前景选项的偏好序列 R 的过程。它提供了一个从 R^* 到 R 的偏好更新机制。通俗而言，如果这些是某人对事态的基本偏好和信念，那么那个行动就是她应该想要选择的。因此称其为一种“评估的逻辑”（logic of evaluation）。这种“逻辑”与传统演绎逻辑和归纳逻辑存在有趣的共通特征。首先，三种理论的规范性都可独立于其描述部分。即使很多人实际上不按照逻辑规范进行推理，也不妨碍逻辑规范成立。例如，当人们的推理实际上符合 *modus tollens* 时，理论家会认为自己既阐明了应该如何推理的规范，又描述了人们实际推理的一些方式。决策论在多个学科领域中的不同研究视角恰恰体现了这点。在经验科学中，它可以被视为对实际决策方式的（有限的）描述，但这并不应该自动定义和穷尽其所能扮演的角色。相比描述事实，经典决策论更成功地扮演了阐明关于评价和选择的“观念关系”的角色。这里借用休谟对“观念关系”和“事实问题”的区分，即使描述决策论事实为假，规范决策论所蕴含的观念关系依然成立。

其次，演绎逻辑表征逻辑有效性，归纳逻辑表征证据支持的强度。演绎系统输入一组真句子，经过系统的公理和推演规则，输出另一组真句子，具有“保真性”（truth-preserving）。或者说，演绎系统是一种更新真值的机制，对两组句子的真值进行“再分配”。归纳逻辑的情况类似，从输入一组假设的证据支持度开始，经贝叶斯更新规则，输出另一组信度或支持度。这种重新分配支持度的过程受到贝叶斯归纳逻辑的约束。具有较强证据支持的假设的后验概率将增加，而具有较弱支持的假设的后验概率将减少。因此，归纳逻辑具有“真值指示性”（truth-indicating）。

相应地，决策论提供了有关合意性（desirability）评估的“再分配”机制，更新主体对于合意性的偏好评估。像演绎逻辑和归纳逻辑一样，决策论形式系统的

内在自洽，可以说具有“有效指示性”（effectiveness-indicating）的优点。我们可以清楚看到演绎逻辑、归纳逻辑与评估的逻辑之间的这种比较，如下表所示：

	输入/输出	过程	优点
演绎逻辑	命题集的真值指派	通过演绎系统重新分配真值	逻辑有效性自洽； 保存真值性
归纳逻辑	信念度集（支持度）	通过贝叶斯更新重新分配信念度（支持度）	证据归纳支持强度自洽； 真值指示性
评估的逻辑	合意性指数	通过表征定理重新分配合意性判断	合意性判断自洽； 有效指示性

6 结语

本文分析并回应了规范地解释经典决策论所面临的两个难题。第一，工具主义解释并不自动排除心智主义解释，相反，规范性决策论依赖一种混合式心智主义。通过分析传统解释对工具主义偏颇的默认，本文开发了心智主义的理论空间。决策论并不只是为了工具式地预测而讲述“好像”的故事。第二个难题来自经验证据：理论的描述预测与现实的选择行动相去甚远。但这并不妨碍决策论作为一种理想规范对理性选择发挥评估作用。进而，本文提出一种刻画这种理想规范的方案，认为规范性决策论是更新偏好的机制。由于规范决策论是对常识心理学的精确化，因此也继承了其中的规范性。如果常识心理学是理性选择的规范尺度，那么规范决策论为这把标尺划上了清晰的刻度。本文并未处理描述层面怎样选择的问题（[24]），只在抽象地理解理性选择问题上具有启发意义。

参考文献

[1] J. L. Bermúdez, 2009, *Decision Theory and Rationality*, Oxford: Oxford University Press.

[2] C. Cherniak, 1986, *Minimal Rationality*, Cambridge, MA: The MIT Press.

[3] D. Davidson, 1985, “Incoherence and irrationality”, *Dialectica*, **39(4)**: 345–453.

[4] J. Dreier, 1996, “Rational preference: Decision theory as a theory of practical rationality”, *Theory and Decision*, **40(3)**: 249–276.

[5] D. Gauthier, 1986, *Morals by Agreement*, Oxford: Oxford University Press.

[6] G. Gigerenzer and D. G. Goldstein, 1996, “Reasoning the fast and frugal way: Models of bounded rationality”, *Psychological Review*, **104(4)**: 650–669.

[7] F. Guala, 2019, “Preferences: Neither behavioural nor mental”, *Economics and Philosophy*, **35(3)**: 383–401.

- [8] F. Gul and W. Pesendorfer, 2008, “The case for mindless economics”, *The Foundations of Positive and Normative Economics*, pp. 3–42, New York: Oxford University Press.
- [9] F. Jackson, 1991, “Decision-theoretic consequentialism and the nearest and dearest objection”, *Ethics*, **101(3)**: 461–82.
- [10] R. C. Jeffrey, 1965, *The Logic of Decision Chicago: The University of Chicago Press*, Chicago: The University of Chicago Press.
- [11] D. Kahneman, P. Slovic and A. Tversky, 1982, *Judgment under Uncertainty: Heuristics and Biases*, Cambridge: Cambridge University Press.
- [12] D. Lewis, 1983, *Philosophical Papers Volume I*, New York: Oxford University Press.
- [13] C. J. G. Meacham and J. Weisberg, 2011, “Representation theorems and the foundations of decision theory”, *Australasian Journal of Philosophy*, **89(4)**: 641–663.
- [14] J. von Neumann and O. Morgenstern, 1947, *Theory of Games and Economic Behavior*, Princeton, NJ: Princeton University Press.
- [15] S. Okasha, 2016, “On the interpretation of decision theory”, *Economics and Philosophy*, **32(3)**: 409–433.
- [16] M. D. Resnik, 1987, *Choices: An Introduction to Decision Theory*, Minneapolis, MN: University of Minnesota Press.
- [17] P. A. Samuelson, 1938, “A note on the pure theory of consumer’s behaviour”, *Economica*, **5(17)**: 61–71.
- [18] L. J. Savage, 1954, *The Foundations of Statistics*, New York: John Wiley & Sons.
- [19] H. A. Simon, 1976, “From substantive to procedural rationality”, in T. J. Kastelein *et al.* (eds.), *25 Years of Economic Theory: Retrospect and Prospect*, pp. 65–86, New York: Springer.
- [20] P. P. Wakker, 2010, *Prospect Theory: For Risk and Ambiguity*, Cambridge: Cambridge University Press.
- [21] L. Zynda, 2000, “Representation theorems and realism about degrees of belief”, *Philosophy of Science*, **67(1)**: 45–69.
- [22] 任晓明, 李章吕, “贝叶斯决策理论的发展概括和研究动态”, 科学技术哲学研究, 2013 年第 2 期, 第 1–7 页.
- [23] 熊卫, “期望效用理论的两个悖论及其消解——兼谈决策论的发展”, 现代哲学, 2013 年第 5 期, 第 82–86 页.
- [24] 张巍, “日常行动的合理决策”, 自然辩证法研究, 第 2013 卷第 6 期, 第 18–22 页.

(责任编辑: 袁之)

The Problem of A Normative Interpretation of Classical Decision Theory and Responses

Ziming Song

Abstract

One important problem in the philosophy of decision theory is how to understand classical decision theory at the normative level. Even as a formal tool, it has been regarded as conforming to certain basic principles of human rational choice since its inception, and has been used to explain and predict choice behavior. The mentalist interpretation of decision theory presupposes that the theoretical elements have corresponding psychological counterparts. Contemporary criticisms of the mentalist presupposition argue that it misconstrues the formal model of decision theory. Meanwhile, the normative dimension of decision theory is also challenged by empirical research such as behavioral economics. However, with further analysis of the problem of normatively characterizing decision theory, I defend a position that characterizes classical decision theory from a normative perspective, aiming to demonstrate that normative decision theory is a “logic” for evaluating the desirability of choices.