

自我欺骗问题研究 ——意向论 vs. 非意向论

On Self-Deception: Intentionalism vs. Non-Intentionalism

马翰林 /MA Hanlin 陈刚 /CHEN Gang

(华中科技大学哲学学院, 湖北武汉, 430074)

(Department of Philosophy, Huazhong University of Science and Technology, Wuhan, Hubei, 430074)

摘要: 关于自我欺骗的道德心理学理论主要分为意向论和非意向论两大阵营。以戴维森为首的意向论阵营强调自欺是一种意向行为, 按照一种人际间欺骗模型来认识自欺。意向论者认为人的心灵可以分为不同的人格系统, 试图用一种非理性的心理因果性来解释自欺。以梅勒为首的非意向论者反对人际间欺骗模型, 从而也反对把自欺看作一种意向行为, 梅勒试图将自欺与认知心理学对接, 认为自欺是由人类的认知特点造成的, 而不是人自主造成的。梅勒提出偏差信念论之后, 英美学界对自欺的讨论进入了一个新的阶段, 但是这些新的观点又各有各的问题。我们将看到, 引入不同的自我知识的理论能够帮助完善以上各个理论, 为解决它们的问题提供帮助。

关键词: 意向 心理分割 偏差信念

Abstract: There are two ways to define self-deception, which are intentionalists' approach and non-intentionalists' approach. As an intentionalist, Davidson emphasizes self-deception is intentional, and describes it as an interpersonal-deception. Intentionalists claim that human mind can be divided into different self-nets, and we should use irrational mental causation to explain self-deception. Non-intentionalists argue against the model of interpersonal deception, and also argue against that self-deception is intentional. Mele treats self-deception as a phenomenon of human cognition, and regards it as a kind of non-autonomy. After Mele, the debates of self-deception have reached a new level. But the new opinions seem to have issues of themselves. And if we introduce some theories of self-knowledge, it may be helpful for both the intentionalists and Non-intentionalists to handle their problem.

Key Words: Intention; Psychological Partitioning; Biased belief

中图分类号: N0 文献标识码: A DOI:10.15994/j.1000-0763.2016.05.007

自欺是一种对我们的生活具有极大影响力的现象。在自省的视角下, 自欺有时被看作一种自戕的毒药, 有时被看作自我抚慰的良药。柏拉图说: “自欺是最糟糕的, 因为你的欺骗者永远不会离开你的心房半步。” ([1], p.144) 而有人则认为自欺是一种帮助人自我控制、获得自我统一性的不自知的认知策略。乔纳·森巴朗 (Jonathan Baron) ([2],

p.71, 238) 认为帕斯卡的赌注 (Pascal's Wager) ([3], p. 153) 就是这样的一个例子: 相信上帝存在对我有利, 不是因为这个策略可以让我获得更多利益, 而是它能帮我获得一种良性的自我控制。

自欺是什么? 它是一种意向性的心理因果作用吗? 还是个体在面对自我、他者或环境时的一个认知策略? 如何回答这些问题, 将决定我们如

基金项目: 国家社科基金项目“当代青年学生马克思主义信仰研究”(项目编号: 11BKS06)。

收稿日期: 2015年7月12日

作者简介: 马翰林 (1985-) 男, 山东青岛人, 华中科技大学哲学系博士生, 研究方向为心灵哲学与科技哲学。Email: fanfan2011cn2000@aliyun.com

陈刚 (1964-) 男, 湖北鄂州人, 华中科技大学哲学系教授, 研究方向为科技哲学、心灵哲学。Email: mgchen@hust.edu.cn

何确定自欺的道德意义。自欺问题已经成为了认知科学以及道德心理学的一个重要的研究领域,但是这个问题还没有在国内学界引起足够的重视。本文将自欺问题的哲学研究为主线来阐释关于自欺本质的两种主流的解释。

一、自欺是否是意向性的?

为了提出一个关于自欺的合理定义,或者说为了论证自欺是如何可能的,英美的哲学家中存在两大类理论倾向,这两类理论的分歧在于是否认为自欺是一种意向行为,其中一类被称为意向论(Intentionalism),持有该观点的有戴维森(Donald Davidson),大卫·皮尔斯(David Pears),贝穆德斯(José Luis. Bermudez)等。另一类被称为非意向论或动机论(Non-Intentionalism、Motivationism),持有该观点的主要是梅勒(Alfred Mele),内尔金(Dana K. Nelkin)等。

认为自欺是意向的(或故意的)行为的哲学家一般都持有这样的一个理解:假如一个人有意识地自欺,那么他必然持有相悖的两个信念, p 和非 p 。意愿地自欺行为是一个故意的欺骗行为,但是在微妙的心理过程中,这个“故意地”的心理内容可能不是太明显。而非意向论的主要工作之一是批判意向论的不足,非意向论者认为意向论存在无法避免的困难。梅勒建立了替代意向论的所谓“偏差信念论”,认为人不可能故意地自我欺骗,但是会在渴望某个信念的时候对证据进行有偏差的认知。这种认知方式不完全是主观造成的,它还受人的客观心理特征的影响。这种认知特征不可避免地会导致一些认知偏差,自欺就是这种偏差。

二、意向论

1. 心理分割

戴维森在《欺骗与区分》一文中认为自欺起码是这样一种情况:当事人自己引发了某种“根据的薄弱”(weakness of the warrant)。所谓根据的薄弱是指:同时存在支持 p 和非 p 的证据,但是当事人的判断却不符合卡尔纳普或亨普尔的归纳推理。即虽然多数证据支持非 p ,但是当事人还是相信 p 是更可信的,有更多的证据支持。([4], p.201)但是,

一个人只是具备根据的薄弱并不一定就是自我欺骗。比如,戴维森举例说,他有一次去国家公园游玩,回来之后在与同伴聊天中认为导游是个男人,但同伴提醒他导游是个女人。戴维森认为这种情况不一定是自欺。因为他可能根本没有意识到要使用归纳推理来证伪“导游是男人”这个信念。或者他确实意识到了这个归纳推理会推出那个人是女人的信念,但是他没有重视这个推理。这些情况可能都是一些失误(mistake),还算不上非理性。

那么自欺的情况是不是当事人故意地要去相信某个信念,而没有顾及相反的证据呢?戴维森认为不完全是这样。这种情况被我们称作一厢情愿(wishful thinking),一厢情愿的人会非常希望某个信念是真的,但是他可能并不是故意地不顾及相反的证据,只是他过于热衷某个信念罢了。如果是这样,一个一厢情愿的人也就算不上是非理性的。此外,自欺并不总是和一厢情愿一样展示了一种有益的愿望。有些自欺者相信的信念是对他来说是痛苦的。比如一个悲观主义者总是认为自己会失败,但实际上他很有可能成功。([4], p.206)

那到底什么是自我欺骗?戴维森认为自我欺骗在某个层面上和人际间的欺骗有点像。他说欺骗并不是(1)说谎者一定要让被欺骗者相信某个信念;也不是(2)说谎者一定要让被欺骗者相信他自己也相信自己说的内容。而是(3)说谎者一定要表明自己相信自己说的那些自己不相信的东西;以及(4)说谎者一定要隐瞒自己的真实意图。如此,自欺就是让自己向自己表明自己相信自己不相信的那些东西,并且向自己隐瞒自己骗自己的意图。但是如果一个人的内心是具有某种内容的一致性的话,自欺就不是自己对自己说谎,因为自己不可能不知道自己的意图,他也不可能相信自己“表面上相信的非 p ”(但实际上不相信)就是正确的信念。([4], p.207)

戴维森认为有一种方法能解决这个问题,那就是把人的心灵分割开来。一部分“欺骗”,一部分“被骗”。两部分之间没有正常的心理因果联系,也就是说没有内容的统一性。这里的欺骗是一种意向性的行为,但是这个意向性不再预设一个同一性的人格。不同的心理内容可以代表不同的人格,如其中一个人格相信自己是秃的,另一个人

格“表面上”相信自己是不秃的。但是戴维森无法回答自我欺骗是如何形成的,或者说一个人如何造成了自己的根据的薄弱?因为根据的薄弱本身并不蕴含当事人要故意地抛弃某些证据或者抬高某些证据的价值。它只是指一种信念尽管在别人看来是错的,但在当事人这里似乎有很客观的证据支持。戴维森认为自欺并不是一个“我自己想要相信某个信念的问题”,而是“我使得我相信某个信念是真的”的问题。但是,这里“我使得我……”中间所带有的意向似乎没有明确的心理内容。

此外,戴维森所说的心灵分割不是一种切蛋糕式的分割,而是一种多元层次式的分割。像是一个人可以在地上被投射出不同的影子——有的时候我们会同时有好几个影子,这些影子有重叠,但又明显不同一。正如他所说:

我们没有必要认为这些边界是永久的,分离的。相对立的关于能否通过驾照考试的信念可能分别隶属于不同的信念网络。它们可能隶属于不同的、但又重叠的领域,但是这些领域毕竟不是相同的领域。如果抹杀掉这些领域之间的区别,那么这些不同信念的对立也会消失掉。我没有明确的理由认为其中的一个领域对意识(认知)是不开放的,但是如果一个人想要看到意识的整体,他就必须去掉这些领域的边界。([4], p.211)

这种边界是无法通过内省发现的,但是它似乎又是真实存在的,戴维森的理论可以很好地解决自欺中的信念冲突,也就是后来梅勒所说的“静态困境”,但是由于自欺的意向是无法获知的,所以戴维森没有办法说明自欺具体是怎么形成的。

2. 主次系统

皮尔斯是另一个重要的意向论者,他将心灵区分为“主系统”与“次系统”两部分,主系统就是正常的信念和欲望系统,而次系统中存在着一种所谓“警戒信念(cautionary belief)”,这种警戒信念正是用来监控主系统的非理性倾向的——因此这个次系统是一个“内在的理性中心”。比如一个人身上没有钱,却想要吃大餐,这时这个次系统就会起作用,把这个一时性起的欲望的不合理性提交给主系统。然后由主系统下判断认定这个想法是错的。而自欺信念的形成是警戒信念失

效的后果,这种失效是因为心灵的分割,即次系统与主系统之间非完全的因果断裂。

尽管两个系统可能分别都是有意识的,但是次系统的意识不是一种自我意识,也就是说我们意识不到我们意识到了这个系统。或者说对次系统的意识是一种无意识,我们没法对它进行因果性的操作。所以,关于自欺,我们就有如下推理:由于(1)警戒系统失灵不被呈现到自我意识中;(2)主系统把那个错误的信念 p 当作是正确的;(3)“纠正 p 的非 p ”这个信念没有办法成为一个意识中的内容。所以,我陷入了自欺状态。

放弃某个警戒信念的意思进一步说就是不采纳某个证据,但是这种消极行为不足以构成一个积极的目的,更不足以直接产生某个信念“非 p ”。正如皮尔斯所说:“我们不需要假设次系统所具有的信息能够表征一个相反的心理行为,那种行为通过慎思展开,并指向某个目的。”([5], p.400)但是他又认为这种次系统已经足以成为一个行为中心(agency center)了,因为这个次系统的确引发了一个特殊的心理动作,这种心理动作只有在自欺的这种非理性的情况下才存在,它是不以自我意识为前提的,但是它是“意向性”的。如若不然,我们所说的就不是自欺了,而可能只是一种一般的一厢情愿而已。也就是说我的内心总是有个我听不到的声音告诉我自己这个信念是错的,而我似乎通过某种隐藏的方式“听到”了这个声音。也就是说,这个次级系统自己“停了下来”,然后还把这个自己停下来的消息以某种方式告诉了我。由于这种意向性无法被表征,所以自欺过程就无法通过自省得到。但我们很难理解这里说的所谓“意向性”是什么意思,或者这个次级系统到底是怎么运作的。皮尔斯本人([6], p.63)、马克·约翰斯顿(Mark Johnston) ([7], p.64)和梅勒都发现了这个问题。尤其是梅勒,他向皮尔斯提出了这样的问题:“到底是什么使次级系统的警戒信念丧失了功能?”;“到底是一种什么样的心理因素周旋于我们的欲望和避免非理性的信念之间?”([8], p.142)他把这种诘难归为心理分割论无法解决他所谓的“策略困境”。

3. 给意向论的建议

意向论的困难在于他们无法清楚地解释“自欺意向”的本质,这是由于(1)意向论者在构建理论时侧重使用与心理动力学协调的本体论概念,

这导致他们使用的概念是模糊的,比如戴维森的“人格”、皮尔斯的“次级系统”,这些概念无法简要清楚地被定义;(2)这里的意向被同“无意识”联系起来,后者也没有得到清楚地定义,戴维森在《非理性的悖论》([4], pp.169-188)一文中试图建立分层的人格同无意识或非理性的关系,这使得问题更加复杂且模糊。笔者的建议是意向论者可以回到知识论,使用心灵哲学的资源来解决自欺问题,暂时抛开非理性和人格问题。

与自欺关联比较大的是“自我知识”问题,该问题牵扯甚广,此处只做粗略引入。根据戴维森的定义,自欺是“让自己向自己表明自己相信自己不相信的那些东西”,其实就是自己向自己表演(通过实现某个心理状态)。经过筛减处理之后,我们把这个条件转化成一个心灵哲学的术语的句子,可以得到“我希望我获得‘我相信p’的心理状态。但同时我不相信p”。自欺意向的目标就是实现“我相信p和我相信非p”的心理状态,如果自欺成功,那么这个状态就应该实现。

处于自我知识讨论中心的外在主义者如普特南和柏奇等人认为“意义不在头脑之内”,这是通过著名的孪生地球思想实验得到的,在我们的语境下可翻译为“意义不等于某个内在(窄的或亲知的)心理状态”(这里省略了一些复杂的问题)。假设“非p”的指称是真实的,那就可以应用外在主义,认为“我相信非p”的指称不是心理状态,而是客观意义(甚至是真值)。而心灵相信p(believing p)的亲知状态在外在主义的意义上是可错的,但是作为窄内容的“相信p”可能遵循另外一套“亲知真”的标准。从而自欺就不再是“主体同时具备两个内在的矛盾心理状态”,而是主体处于“意义”和心理状态的“矛盾”中。这里的“矛盾”不是经典逻辑矛盾,或者说不矛盾。因为“意义真”和“亲知真”之一可能不是经典命题的真(很可能后者不是),或许可以用达科斯塔(N. da Costa)的弗协调逻辑运算。当然还有一种极端的理解,就是无论“相信p”还是“相信非p”都是一种心理状态,但是遵循不同的真理标准“亲知真1”和“亲知真2”,所以不矛盾。

但戴维森不会接受亲知状态是可错的,他的自我知识理论完全不同,他认为亲知反思即是主观的又是客观的,不存在相对的真理标准。自我知识权威性的基础实际上是随附性,([9], p.20)

随附性是不可错的,这与戴维森使用事件本体论,否弃任何观念对象([9], p.36)有关。所以他要创造分层的心灵,每一层有相悖的内容,而不是用“宽”和“窄”来给心灵分类,它们同时是可反思的、权威的、甚至是“正确的”。所以心理分割论可以被整理如下:逻辑的地位低于心灵的本体属性,自欺只是两段(或更多)心灵本体属性(可分割)允许的、逻辑不连贯的心理事件,而且我们完全能从这个事件中看出这个人在故意地自欺。最终解决戴维森的问题的可能是某种脑科学和随附理论。但在此出现之前,他还是要面对自我同一性问题。

三、非意向论

1. 两个困境

梅勒认为如果按照人际间欺骗模型来理解自欺的话,自欺就是不可能的。因为这会导致两个无法解决的困境。第一个困境被称作“静态困境”或“状态困境”(static puzzle),这个困境正是来自“自欺”的词典定义,对此梅勒有如下论述:

对于任何两个人A和B,当A欺骗B使他相信p,而A在使B相信p的时候自己则知道或者确切的相信~p。所以当A欺骗A(他自己)去相信p,而他在欺骗自己的同时又知道并且确切的相信~p。因此,A必须同时相信~p和p。但这怎么可能呢?([8], p.121)

这个困境实际上是在指出自欺的人是有人格同一性的,而人际间欺骗模型中的自欺者似乎没有人格同一性。第二个困境被称作“动态困境”或者“策略困境”(strategy paradox),这个困境的核心在于如果我们按照“自己骗自己”这种定义来理解自欺的话,那么我们是如何把那个知道我们的欺骗的目的和手段的自己欺骗成功的呢?用梅勒的话来说“一般而言,如果B知道A的目的和计划,则A很难成功地实施一次欺骗。”([8], p.138)另一方面,如果说“一个人可以不知道自己的目的而对别人进行欺骗,这也是荒谬的,因为一般而言行为者似乎很难在缺乏一个目标的情况下制定一个计划”。([8], p.138)梅勒认为意向论也无法回答动态困境,因为自欺的人不可能知道自己故意骗自己的那个意图,如果他知道,“我们就很难想象他所知道的这个情报如何在阻止他自欺的时候

失效了?”。([10], p.92) 如果他不知道这个意向,那么他在什么意义上是“故意的”自欺呢? 这个问题确实很难回答。所以他转向了另外一种理论,来回避这两个困境。

2. 偏差信念

梅勒的理论可以被称为有偏差的信念(biased belief)论(简称“偏差信念论”)。他提出,通常我们都会有这样的直觉,即我们的欲望会影响信念的形成,这会导致下面这些现象:1. 消极曲解,不承认不利于自己欲望的事实;2. 积极曲解,在解释问题的时候会支持一些不存在的事实。因为这些事实如果不存在的话,会对动机的实现不利;3. 选择聚焦,忽略不利证据,聚焦有利证据;4. 选择性证据收集,忽视易收集的不利证据,寻找不易收集的有力证据。([11], p.26)

动机是如何导致并维持有偏差的信息证据收集活动呢? 梅勒指出人类的一些认知习惯以及一些认知对象的特征是实现这一点的基础,它们被称作偏差信念的非动机性基础或“冷”基础。它们包括:1. 对信息的鲜明性的积极反应,“对一个个体来说,信息的鲜明性只是他个人兴趣的一个功能,一种信息的具体化,一种‘激起意象’的力量('imagery-provoking' power),或者一种感觉的、暂时的、空间临近关系”([11], p.28)而我们更容易对一些鲜明的信息进行处理;2. 探启的效应性(the availability heuristic),顾名思义,我们在形成信念的时候,往往比较容易收集那些容易接近的信息,它们可能和我们的某些记忆的关系很近,也可能更容易被用来构建想象。例如,比起晦涩的哲学书,我们更愿意去看直白一些的小说;3. 确证的偏见,比起反驳一个假设,人们往往更愿意去证实一些假设,也就是说人们习惯于去验证关于p而不是~p的假设。这些规律大多来自心理学的实验。([12], [13], [2], pp.181-182) 梅勒认为,动机可以通过这些人类的认知习惯来影响信念,比如一个动机可以让一个认知对象变得鲜明起来,从而使得其更容易被识别,而这种更容易被识别的对象就更容易被当作某个信念的证据,然后一个可能是偏差的信念就形成了。偏差信念的冷基础是非意向性的,这些冷基础或者说自然功能可以受到动机的影响。

梅勒提到了一个日常假设检验模型,用来说明动机和假设检验过程中的证据收集的关系。由于这个模型来自弗雷德里希(J. Friedrich)^[14],雅各布·特普(Yaacov Trope)和阿齐瓦·利伯曼(Akiva Liberman)^[15],所以其被称作Friedrich-Trope-Liberman (FTL)模型。弗雷德里希认为“检查和最小化核心差错实际上是假设检验中核心的组织原则。”([11], p.31)而不被期望的差错被称作“昂价差错”。比如,如果我们要证实一个自欺的假设,这个假设的错误很大,但是否认这一点却代价高昂。这种差错被如何评估是受一个人的动机影响的。但是动机只是能够确定哪些差错是被允许的,哪些是不被允许的。被允许的差错就不需要被进行最小化处理,它们会被明显化。这个模型不仅仅被用来解释自欺行为,它可以用来解释一切人类信念的形成方式。它意味着人类的认知受到一种自动的(在弗雷德里希看来是生物性的)功利主义原则影响,这种功利主义的标度在不同的人身上是各异的,如果一个人偏好寻求真理,则那些不利于发现真理的差错就是“昂价差错”,而自欺者则拥有着另外的一些价值标准。梅勒认为动机在这其中并不是一个起独立作用的角色,我们可以这么来理解:不存在这么一个欲望,它会有意识地想要避免某些差错。避免差错本身不是意向性的。它只是用来触发或者维持一个相应的假设,不利于这个假设的一些“昂价差错”就会被我们的认知机制“最小化”。

“特普和利伯曼1996年提供了一个对于假设检验过程中的代价的一个精密的解释”([11], p.34),这个理论和上面弗雷德里希的PEDMIN (Primary Error Detection and MINimization)理论相似。我们可以将他们的理论称为“信任阈值”(confidence)理论或者“阈值”理论。这个理论是说,“阈值越低,对于相应的证据的要求也就越低。”不同的假设有着不同的阈值,而那些拥有着“昂价差错”的假设往往阈值比较低。也就是说一些很低劣的证据也能够越过这个阈值为这个假设提供支持。具体而言阈值又分为“接受阈值”和“拒绝阈值”。“接受阈值”是在接受某个假设前对其真实程度的最低要求,而不只是继续测试它的要求,而‘拒绝阈值’是在拒绝假设并终结测试前对其不真实程

①对弗雷德里希的理论介绍主要转引自Mele ([11], pp.31-46)

度的最低要求(或最高限度)。”([11], p.34) 同一个假设的两个阈值可能不一样。这是由于误拒绝和误接受的成本不一样,比如我担心自己的驾照考试是否可以通过。我可能会倾向于(错误地)接受一些能让我相信我会通过考试的信息。我发现有些水平不如我的人已经通过了考试,但是这个消息很可能是不准确的,这个人可能是通过贿赂考官而通过考试的。但是我更愿意相信这个证据,我的相关的接受阈限就很低。而如果有人想让我拒绝这个证据,可能就很困难,假如有人告诉我那个水平比我低的人是通过贿赂考官通过考试的,我可能会让他给出证据,甚至要他拿出照片或录像作为证据来证明他的说法,只有这样我才能拒绝认为那人凭本事通过了考试。在这里我的拒绝阈限就很高。

最终,梅勒给出了四个自欺的充分非必要条件:(1) S获得的p信念是错的;(2) S在收集信息评价p的真实度时,受动机的影响,产生了偏差;(3) 这种“有偏的”方式是一种“非不正常的”(non-deviant)方式;(4) 这些证据对非p的支持程度高于p。([10], p.95) 可见梅勒可能会认为自我知识是可错的(p可以是自我知识)。他很可能是个外在主义者,但是没有形成自觉,这导致了偏差信念论的危机。

3. 偏差信念论之后的讨论

偏差信念论被认为会遇到一些困难,多数批评者认为梅勒所描述的自欺无法与一厢情愿的情况区分开。贝穆德斯认为可能存在三个自欺的定义,由弱到强排列如下:(A) 自欺者意图获得某个特定信念;(B) 自欺者认识到持有某个信念是错的,但是依然意图获得这个信念;(C) 自欺者意图获得一个错误的信念。([16], p.310) 而偏差信念论所描述的自欺只能是(A)类。这一类自欺可能只是一厢情愿而已,因为偏差信念者并不知道他的信念是错的。而且偏差信念论无法说明为什么一个人会在某个特定时间自欺,或者说梅勒的充分不必要条件的出现并不意味着自欺会出现。只是“想要p”加上“冷基础”并不足以产生一个自欺,要不然我们的生活中就到处都是自欺了。为什么“这个”欲望就偏偏会产生自欺呢?贝穆德斯认为意向论可以很好的解释这一点:因为我们有意地要在此时自欺。([16], pp.316-318) 意向论可以很好地符合(B)和(C)这两种情况,但这个理论对意向论的理解比起戴维森的理论来又过

强了,根据戴维森的观点,欺骗不是要让对方相信一个错的信念,而是向对方表明我相信某个(错误的)信念是对的。所以说,戴维森的自欺顶多可以符合(B),但是不符合(C)。如果认为自欺必须符合(C)的话,那会是一个最强的自欺意向论的观点。除贝穆德斯之外,对梅勒展开批判的还有尼克森(A. Nicholson)^[17]、以及拜纳(Tim Bayne)和费尔南德斯(Jordi Fernández)等。拜纳和费尔南德斯提出,我们很难想象梅勒的模型能解释为什么一个认为自己没有得癌症的癌症患者会逃避去医院向医生咨询病情,并且不按时去医院检查等等([18], p.11)。一个一厢情愿的人可能会积极地根据他想要的信念去实践,但不会去做出逃避能实现相反信念的行为。

内尔金对贝穆德斯的那种批评给出了一个回应。他修改了梅勒的模型,认为导致偏差信念的那个欲望是“希望p为真”而不是“希望p”。“希望p”是一厢情愿的欲望,而自欺的核心应该是“希望p为真”,这是在要求真值条件,而不是简单地欲求心理状态。([19], p.396) 他认为自欺要求一种“相对高阶的知性水平”,([19], p.398) 所以说自欺是有道德责任的。小孩子往往不具备这种能力,所以小孩往往只能产生一厢情愿。成人也会产生一些其他类型的非理性情况,如果那只是涉及一阶的“态度”或欲望,那他就不是在自欺。内尔金的问题在于,我们很难想象“希望p为真”如何同“希望p”在心理上区分开来,除非这个人有一个清醒的意识“p为真”。然而如果是这样,我们就有理由认为他会产生另外一个意识“非p为假”。按照内尔金的说法,似乎“非p为假”是不存在于意识当中的,但是我们很难想象“希望p为真”在意识中而“非p为假”不在意识中的情况,除非这个人的心理有某种分割,正如戴维森所设想的那样。

整个非意向论的问题在内尔金的讨论中集中体现了出来,问题在于他没有自觉地预设一个外在主义视角,把真值条件和自我知识的“窄内容”或亲知的心理内容区分开来,所以他永远也无法区分“希望p”和“希望p为真”。因为除非把“希望p为真”所表示的真值条件理解为由外在对象或宽内容决定,否则他所表示的真永远只能是窄内容自己内部的“真”,这个“真”是和窄内容无法区分的。

四、结 论

我们看到意向论和非意向论各自的问题都可以被归结到一些心灵哲学预设上。意向论是一种在哲学上更完备的理论,它能够更好地解释几乎所有日常的自欺现象,如果我们引入戴维森的一些关于自我知识等问题的理论,或许就能解决它在知识论上的困难;有偏信念论可以很好地避免两种困境,但是却只能给出自欺的充分条件,也就是说这个理论无法给出自欺一个完备地说明。但是由于这个理论与心理科学联系紧密,而心理学的解释往往就是不完备的。所以从科学的角度来看,有偏信念论是可以接受的。此外,一个恰当的外在主义知识论或许会使它在哲学上更加完备。

[参考文献]

- [1] Plato. *Plato Complete Works*[M]. Cooper, John M.(Eds) Hackett Publishing Company Indianapolis: Cambridge Press, 1997.
- [2] Baron, J. *Thinking and Deciding*[M]. Cambridge: Cambridge University Press, 2007.
- [3] Pascal, B. *Pensees and other Writings*[M]. Levi, Honor(Trans.) New York: Oxford university press, 1995.
- [4] Davidson, D. *Problems of Rationality*[M]. New York: Oxford University Press, 2004.
- [5] Pears, D. 'Belief and Rationality'[J]. *Synthese*, 1991, 89(3):393-405.
- [6] Pears, D. *Motivated Irrationality*[M]. New York: Oxford University Press, 1984.
- [7] Johnston, M. 'Self-Deception and the Nature of Mind'[A], McLaughlin, B., Rorty, A.(Eds) *Perspectives on Self-Deception*[C], Berkeley: University of California Press, 1988.
- [8] Mele, A. R. *Irrationality: An Essay on Akrasia, Self-Deception and Self-Control*[M]. Oxford: Oxford University Press, 1987.
- [9] Davidson, D. *Subjective, Intersubjective, Objective*[M]. New York: Oxford University Press, 2001.
- [10] Mele, A. R. 'Real Self-Deception'[J]. *Behavioral and Brain Science*, 1997(2):91-136.
- [11] Mele, A. R. *Self-Deception Unmasked*[M]. Princeton: Princeton University Press, 2001.
- [12] Klayman, J., Y.-W. Ha. 'Confirmation, Disconfirmation, and Information in Hypothesis Testing'[J]. *Psychological Review*, 1987 (94):211-228.
- [13] Nisbett, R., Ross, L. *Human Inference: Strategies and Shortcomings of Social Judgment*[M]. Englewood Cliffs, N.J.: Prentice Hall, 1980.
- [14] Friedrich, J. 'Primary Error Detection and Minimization (PEDMIN) Strategies' in *Social Cognition: A Reinterpretation of Confirmation Bias Phenomena*[J]. *Psychological Review*, 1993(100):298-319.
- [15] Trope, Y., Liberman, A. 'Social Hypothesis Testing: Cognitive and Motivational Mechanisms'[A]. E. Higgins and A. Kruglanski(eds), *Social Psychology: Handbook of Basic Principles*, New York: Guilford Press, 1996: 239-270.
- [16] Bermúdez, J. L., 'Self-Deception, Intentions, and Contradictory Beliefs'[J]. *Analysis*, 2000, 60(4): 309-319.
- [17] Nicholson, A. 'Cognitive Bias, Intentionality and Self-Deception'[J]. *Teorema*, 2007.26(3): 45-58.
- [18] Bayne, T., Fernández, J. *Delusion and Self-Deception, Affective and Motivational Influences on Belief Formation*[M]. Psychology Press, 2008.
- [19] Nelkin, D. 'Self-Deception, Motivation, and the Desire to Believe'[J]. *Pacific Philosophical Quarterly*, 2002, 83(4):384-406.

[责任编辑 王巍 徐竹]