



认知科学的 哲学应用

Philosophical Applications of Cognitive Science

[美] 阿尔文·戈德曼 (Alvin I. Goldman) 著
方环非 译



ZHEJIANG UNIVERSITY PRESS
浙江大学出版社

认知科学的 哲学应用

Philosophical Applications of Cognitive Science

[美] 阿尔文·戈德曼 (Alvin I. Goldman) 著
方环非 译



ZHEJIANG UNIVERSITY PRESS
浙江大学出版社

图书在版编目 (CIP) 数据

认知科学的哲学应用 / [美] 戈德曼 (Goldman, A.) 著; 方环非译. — 杭州: 浙江大学出版社, 2015. 7

书名原文: Philosophical Applications of Cognitive Science

ISBN 978-7-308-14668-5

I. ①认… II. ①戈… ②方… III. ①认知科学—应用—科学哲学—研究 IV. ①N02

中国版本图书馆 CIP 数据核字 (2015) 第 095806 号

浙江省版权局著作权合同登记图字: 11—2015—126 号

认知科学的哲学应用

[美] 阿尔文·戈德曼 (Alvin I. Goldman) 著
方环非 译

责任编辑 胡 畔 (lpp_lp@163.com)

封面设计 春天书装

出版发行 浙江大学出版社

(杭州市天目山路 148 号 邮政编码 310007)

(网址: <http://www.zjupress.com>)

排 版 杭州中大图文设计有限公司

印 刷 杭州日报报业集团盛元印务有限公司

开 本 880mm×1230mm 1/32

印 张 7.5

字 数 160 千

版 次 2015 年 7 月第 1 版 2015 年 7 月第 1 次印刷

书 号 ISBN 978-7-308-14668-5

定 价 38.00 元

版权所有 翻印必究 印装差错 负责调换

浙江大学出版社发行部联系方式: 0571-88925591; <http://zjdxbs.tmall.com>

中文版序言

《认知科学的哲学应用》一书出版于二十多年前的 1993 年。自那个年代以来,无论是哲学还是认知科学领域都有很多新的发展,这些发展无疑拓宽并再次赋予了这些领域之间富有成效的交互所具有的前景。因此,有很多机会可以重新确立并培育哲学与认知科学之间的共同界面。这个序言选取了当代哲学中一些重要的论题,并表明当前心理学与神经科学中的研究如何能够被关联到这些或那些论题,以及这些研究发现如何能够活跃并丰富哲学的思考。

这个序言重新审视了本书所考察的四个哲学分支。对于每个领域,我都会提出并简洁地呈现一个或一个以上新的问题,然后将它们与认知科学中的新发现或视角进行交叉、结合。这四个领域包括认识论、心智哲学、形而上学和伦理学。^①

^① 为何只是选择认识论、心智哲学、形而上学以及伦理学等四个领域,而没有对科学哲学的新发展进行阐释,译者曾就此与戈德曼教授进行沟通,他认为“科学哲学相比较其他四个领域而言没有很多新的进展,况且确实没有必要在中文版前言中就每个领域、每一章的主题进行新的讨论”。——译者注

一、认识论

认识论中的经典问题之一就是人们如何获得其信念的**确证**(justification)。究竟是什么东西使得他们有资格相信某些命题而不是其他命题？对这一问题的最常见解决路径就是**基础主义**，它主张我们大部分确证地相信的东西根本上源自于一系列特别的具有**基础性**(或**直接性**)的已然得到确证的命题。那么对于一个信念而言，它以一种基础的或直接的方式得以确证意味着什么呢？用否定性话语来说，它意味着信念的确证并非衍生于**其他信念**(指的是其自身必须得以确证的那些信念)。相反，用更加肯定性的话语来说，确证源自于某些其他类型的内在因素，它们能够**赋予确证**(conferring justification)。这里最基本的形式之一就是**知觉经验**。当人们有个视觉经验——院子里有只袋鼠时，该视觉经验(其自身并不是信念)使得人们确证地相信院子里有只袋鼠。对认识论家而言，问题就变成：究竟是如何，或者在什么条件下，S之拥有一个视觉经验——有这么一个带有属性F的物体，赋予S之相信在其面前的物体就是F以确证呢？

对这一问题的常见回答是由被称为“犬儒主义”的认识论观念所提供的。(在认识论家使用这一术语时，它并不暗含着什么理智上的固执。)根据以上描述，对犬儒主义可以进行如下阐释。

“犬儒主义”：无论何时当你拥有某物 F 对你来说以某种方式存在的经验时，那么你相信那就是 F 便得到直接的确证。

犬儒主义的原则在很多情形中似乎都有道理。然而它是在所有情形中都有效吗？很明显会有一些以下情形的反例，知觉者的经验受到其心智中反常因素的影响，比如欲念（desire）、情绪或先前信念。比方说，假定洁儿（Jill）相信杰克（Jack）对她发火了，这就使得她感受到杰克的脸上有愤怒的表情。犬儒主义对此就意味着，洁儿是确证地相信杰克对她发火了，因为她的经验就是“as of”他在生气（Siegel, 2012）。不过对洁儿认知状态的这个结论似乎并不正确。假设洁儿先前的怀疑——杰克生气自身没有得以确证，那么她的视觉经验——他在生气就是由误导性的、反常因素所引起。洁儿继之而来的信念或被强化的信念——杰克在生气，它由视觉经验所引起，并不是真的得以确证，这与犬儒主义所声称的恰好相反。

犬儒主义的辩护者或许会对这个反驳意见进行如下回应。上述情形只是表明，杰克的表情（对洁儿而言）似乎他是在发火，这只是洁儿对他的部分视觉经验。我们需要更为仔细、准确地对待经验的内容。哲学家或许会主张，视觉经验并不包含诸如“在生气”或“看起来在生气”这样的属性。视觉经验仅仅包含有关事物的形状、事物的颜色、事物的距离等等的内容。后者属于“低级”属性。高级属性如“在生气”不是视觉经验的一部分。它们至多是后经验（post-experiential）认知的属性。一旦我们将洁儿认知的生气这一方面从其（严格意义上的）视觉经验中去除，犬儒主义就不会

表明,她是确证地相信杰克在生气;因为生气这一要素并不处于其经验本身“之中”。

我们该如何解决视觉经验是否包括如生气这样的高级特征,而不单单是像形状、颜色等低级特征呢?在这一点上,我们诉诸认知科学似乎是合适的(且必要的)。麦克格拉斯(McGrath, 2015)表明有**联想性失认症**(Associative agnosia)这样的现象,它是由认知科学家(比如参见 Palmer, 1999)所发现的某些病患的缺陷。某些病患“早期”视觉区域受到损伤,在特定视觉相关的任务中他们或许表现正常,这样在向他们展示有特定物体如香蕉的图片时,他们能够用手准确画出来。因此,他们并非盲人,他们有视觉经验。不过在要求他们对物体进行分类时,他们就完全困惑了,不能完成这一任务。不是说他们似乎已经失去或遗忘香蕉这个概念(或语词)。如果是通过口头问他们有关香蕉这个物体,他们能够说出什么是香蕉。对于联想性失认症来说,似乎就变成他们的高级经验不包括香蕉这样高级特征。换句话说,他们存在**经验上的**(experiential)缺陷。正常人在其知觉经验中确实有像香蕉这样的种类区分。这似乎与犬儒主义者的提议相冲突,他们认为在洁儿的例子中,生气这样的特征必然要被视为**后经验性因素**而不是**经验性因素**。与之相反,对正常人而言,高级特征是经验的构成部分。

二、心智哲学

心智哲学是第三章的内容,可以将其分为两个方面,即

形而上学与认识论。认识论方面(仅指这里所讨论的内容)着重于人们如何知道其自身心理状态以及他人心理状态。我们拥有什么特别的通道把握自身的状态,至少是我们自己的意识状态吗?它会产生有关我们经验、信念与情绪的不可错知识吗?另外,我们如何“读出”他人的心理状态?总而言之,两类读心(mindreading)(即第一人称与第三人称)包含了“民众心理学”的内容,也即那些甚至未经训练的人理解其自身以及他人心智的方式。

有关第三人称读心的理论——“模拟理论”自1993年以来已经得到可观的发展。其中某些发展受到了意大利帕尔马大学瑞佐拉蒂(Giacomo Rizzolatti)所领导的神经科学家团队发现的镜像神经元与镜像所促发。镜像最初是在猴子的前运动皮层(premotor cortex)中发现,而行为计划正是在这里发生的(比如打算抓握一个物体)。前运动皮层要负责将信息发送至像手这样的感受器(effector)以抓住某个物体。在猴子做出某种手部行为,并且当它注意另一个猴子或人做出相同的手部行为时,那种被称为“镜像”神经元的东西具有奇特的(且未得预见的)属性。因此,在观察另一个个体时,猴子大脑有一部分“模拟”它在类似情境中施行的相同行为。这个发现(再加上认为镜像在人类中也同样会出现的原因)使得加莱西(Vittorio Gallese)与本人提出如下假设,镜像或许是能够被用于多种用途(可能包括读心)的大脑的根本特征。(Gallese and Goldman, 1998)

基于镜像的读心的证据或许最好可以说是在另一个领域中被发现的,即通过人们的面部表情读出他们的情绪。戈德曼与斯里帕达(Goldman and Sripada, 2005)回顾了恐惧、

恶心与愤怒识别方面的神经科学文献,这些文献在以下情形中发现了偶对**缺陷**(deficit)(**损伤**[impairment]):(1)经历着某类情绪;(2)将同样情绪归赋于他人。换句话说,没有以正常方式经历恐惧的病患也趋向于缺乏对他人面孔上恐惧的识别,但在识别他人情绪方面则并不缺乏。这表明,正常个体在知觉他人处于恐惧状态时是利用其自身的恐惧经验器官(尤其是杏仁体)。杏仁体损伤(amygdala impairment)的个体则不能像正常个体所表现出的那样成功进行同样的情绪识别。根据本人在一本大部头著作中对于读心这一问题所做的论述,“低层”读心似乎是通过**模拟**来进行的(“Simulating Minds,” Goldman, 2006, chap. 6)。这一观点就是,在看到一個情绪性表情时,大脑试图“**确定**(enact)”或“**模仿**(imitate)”其自身中的情绪,其目的是为了识别发生于被观察个体身上的情绪。

另一种不同的模拟处理在“高级”认知中同样起着一定作用,这种模拟是由**想象**(imagination)来执行的(Goldman, 2006, chap. 7)。人们在自己心智中建构相同的初始状态,认为目标个体会处于相同的状态。比如,如果你在下象棋,想要预测你的对手下一步是什么,你会从对手的视角来想象棋盘,并从这个视角来计算好招数。认知系统就是根据这个假想状态(比如走黑棋而不是白棋)来运作,并从这个座位起点的假想状态来形成计划,然后你就预测这正是你的对手要走的那一步。

针对读心的一个趋同的进路是由两位神经科学家(Bunker and Carroll, 2007)分别提出来的。他们探究了被他们称之为“自我投射”的神经基础,大体来说就是,想象自

己处于一个(不同于实际情景的)其他可能情景之中。人们可以将其自己投射到过去(当他们试图从记忆中找回一些东西时);人们可以将自己投射到未来;人们可以将自己投射到一个不同的地点;人们也可以将自己投射到另一个人的心智之中。邦克与卡罗尔(Bunker and Carroll, 2007)辨别了大脑的大片区域,总体上将其称为“默认网络(default network)”。它包括内侧前额叶皮层(medial prefrontal cortex)、楔前叶(precuneus)、后扣带回皮层(posterior cingulate cortex)以及侧顶叶皮层(lateral parietal cortex)。这个网络被连接到某些任务,在这些任务中人们回忆来自过去的经验,想象经历一些假想的事件,并考虑未来要经历的具体事件。这样,默认网络对于高级模拟读心的催生(substrate)而言是个非常好的备选。

作为论述我们如何读心(或阅读他人心智),模拟进路与所谓的“理论论”进路彼此相互对立。假定这一现象对人类生活与社会交互有着重要意义,那么哲学家就有特别的理由要去理解其运作的模式。有充分依据的理解在部分意义上依赖于认知科学知识。

三、形而上学

我们的第三个领域是形而上学,这是第四章所讨论的内容。传统的面对形而上学的进路是在形而上学与实证性科学之间画一条清晰的界线。形而上学的真理——尤其是本体论的真理在传统意义上被视为必然真理,通过先天的途径

可以获知。科学的发现与科学的方法论在此似乎没有关联。需要承认的是,当代形而上学家在刻画基础本体论时越来越多地转而向物理学寻求帮助。这里并没有什么趋势要将认知科学带到这个领域之中。那么如果对认知科学来说存在这么个切入点的话,会是什么呢?

本文作者(Goldman, 2015)最近论述了这样一种切入点,其他哲学家也在探究有些相似的思路(如 Paul, forthcoming)。下文是我对认知科学如何且为何能够被关联至形而上学的发展作出思考的一部分内容。

形而上学中最根本的争论在于各种常识性物体、属性或关系是否真的存在并拥有我们生来便认为他们就拥有的那些特征。比方说,实践存在吗? 它会“流动”吗? 对这类问题的可能回答就是,某些具体物体并不真的存在(尽管我们对其命名,并“相信”它们);或者它们缺乏通常被归赋于它们的那些属性。例如,时间存在但并不“流动”。这一揭示真相(debunking)的立场能够以多种形式出现:力主完全“消除”所讨论的物体,或者将其降到(demote)形而上学的层面。这样的降级处理亦可能呈现在多重标签之下:虚无主义、响应依赖(response-dependence)、建构论、投射论、表达主义(expressivism)、虚构主义(fictionalism)、误差理论等。这里将所有这些立场均称为“反实在论”,在这一范围中,形而上学的中心任务便是在(某个选定领域的)实在论立场与(那个领域的)某个反实在论立场之间进行选择。

形而上学家应该如何决定赋予某个领域内的实在论立场与反实在论立场多高的置信度(credence)呢? 在努力确定一个合适的置信度时,形而上学家应该依据什么证据呢? 应

该寻找什么样的新证据呢？尤其是，或许是不是会有相关于形而上学家的认知任务的那种来自认知科学的证据呢？

关于实在和/或目标形而上学对象的本体论地位，形而上学家的初始证据通常是一套日常经验、直觉，以及/或者有关推定性对象的信念。比如，说到时间的流动，一项证据可能就是一次经验或者一种时间似乎流动的感觉(feeling of time as flowing)。不过这一点点证据应该如何进行评价呢？运用贝叶斯推理来处理的话，形而上学家会考察拥有这个经验的可能性：(1)如果实在论者的假设为真的话，且(2)如果反实在论者的假设为真。(这些就是正在出现的证据事件的“可能性”或条件概率)。假定时间确实流动，那么拥有一次时间似乎流动的经验的可能性应该会非常高。假定时间并不流动的情况下拥有同样经验的可能性如何呢？形而上学家最初或许对这样的可能性指配极低的置信度。不过心理学理论或许隐含地主张，单单心理因素可能就会生成这样的经验，甚至即使时间自身并不流动。如果形而上学家学会相关心理学知识或理论，他或她就应该提高对这种情形的可能性评估。这正是贝叶斯推理所要实现的，即增加反实在论者的假设的概率。

这样一系列的反思解释了来自认知科学的证据是如何能够做到与形而上学理论有所不同。它们影响了与形而上学核心问题的假设——实在论者与反实在论者的假设相关的随后(posterior)概率。戈德曼(Goldman, 2015)提出了四个形而上学领域的例证，并标明这些反思能够加以有效地、严肃地探究。贝叶斯模型(template)被用于阐释几位不同学者的论证，他们近来均(至少在部分意义上)基于认知科

学提出了“揭示真相”的论证。这些学者与论题包括保罗(Paurie Paul, 2010)讨论时间流动, 斯特里特(Sharon Street, 2006)探究价值实在论, 莱斯利(Sara-Jane Leslie, 2013)探讨本质与自然种类, 以及那仑扎彦等(A. Narenzayan et al., 2012)研究上帝的存在。

四、伦理学

对于伦理学问题, 我用两个例子来表明认知科学的最近研究是如何影响道德理论的。心理学中极富影响的框架就是双加工理论(dual-process theory)。根据这一框架, 人类心智是用两个根本上完全不同的加工/过程或系统类型来运作的(Kahneman, 2011; Evans, 2008)。“系统 1”的过程自动地且快速运作, 很少或根本没有自主控制(voluntary control)地尝试努力与感觉。当某个反应成为系统 1 的产物时, 主体常常没有意识到诱发这一反应的刺激, 甚至没有意识到反应本身。与之相比, “系统 2”过程则是慎重的, 充满着努力的(effortful), 也是有组织的(orderly)。系统 1 与系统 2 过程的二分被认知科学应用至很多不同的论题。然而, 它们还被应用到道德哲学的难题, 即在什么条件下某个行动者会因为其施行的某行为或产生的某结果而应该负有责任(blameworthy)? 道德哲学家努力给出一个令人满意的一般性回答。

史密斯(Holly Smith, forthcoming)诉诸双重系统理论做出了重要的区分, 这个区分能够帮助对道德上的负有责任

与没有责任(甚至即使结果“很糟糕”)做出一般性阐述。大体上说,系统2似乎具有监管甚至有时推翻与系统1过程相关联的行为趋向的功能。然而,系统2有时又无法做到足够迅速地从中干预(intervene),这样的话系统1的过程在一种迅速的、草率的状态下得以进行下去,同时产生与主体(agent)所拥有的其他欲求(desire)以及反感(aversion)“不相一致的”结果。准确地说,在这样的情形中,史密斯认为很自然会将这个行为视作没有责任的(blameless),尽管其仍具有令人反感的(objectionable)特征。当然,在负有责任的行为与没有责任的行为之间进行如此区分并不只是直接源自于认知科学的“出现”。不过科学的心理学在勾画那些我们发现具有本质与直觉特征的道德地位之间的原则性差异上起到一定作用,只是如果没有来自科学的帮助这一点难以清晰、准确地进行区分。

第二个例证涉及哲学家试图运用神经科学来阐明,同时也是批判我们的某些道德直觉。这个例证源自于道德哲学中对道义论与后果论的传统区分。基于fMRI研究,格林(Joshua Greene)与他的同事(Greene et al. 2001)认为,在我们的大脑中有两个对立的系统,它们支撑着不同的道德直觉。第一个系统是利用情绪的神经过程,所形成的判断通常相关于伦理学的道义论进路。第二个系统则利用更具认知性的神经过程,所导致的判断与伦理学中的效用论或后果论相关。这些在运用“电车(trolley)”之例的实验研究中已经得到呈现(Thomson, 1976)。在“电车驾驶员困境”中,驾驶员开着一辆失去控制的电车,它正要撞向五个工人。只要驾驶员拉动拉杆让电车改变行驶路径,这五个工人就会得救,但

那样的话就会撞死在另一条道上的一个工人。在所谓的“天桥困境”中，你正站在那辆失控的、飞速行驶的电车所沿轨道上方的天桥之上。在你边上有个陌生人正好是个大胖子，如果你把他推下天桥掉到轨道上，他的硕大身躯就会阻止电车行进。尽管这个胖子会死掉，但那五个在电车轨道上的工人会得救。上面两种情形中到底应该如何做呢？

格林等人预言称，在周密思考这个选择时，通过扫描大脑表明天桥困境将会使人们引发更多情绪活动，因为他将会对“个人的(personal)”道德困境做出反应。这个脑激活的预测实际上已经得到确认。同时，因为情绪是值得怀疑的(格林等人给出了原因)，他们认为这就倾斜了道义论与后果论之间的天平，意味着构成对道义论的反对而支持后果论。然而，后一个推论引起很多道德理论家的怀疑(如 Berker, 2009)。不过它确实表明了有意思的可能性——将神经科学引入与道德哲学进行的对话。

我很感谢浙江大学出版社以及编辑胡畔小姐、宋旭华先生等在出版本书中文版中所做的工作。特别要感谢本书译者方环非教授对本书所做的细致、认真的翻译。能将本书介绍给中国读者我感到非常荣幸。

阿尔文·戈德曼

Alvin I. Goldman

Bibliography

Berker, S. (2009). "The moral insignificance of neuroscience." *Philosophy and Public Affairs* 37: 293-329.

Buckner, R. L. and Carroll, D. C. (2007). "Self-projection and the brain." *Trends in Cognitive Sciences* 11: 49-57.

Evans, J. S. B. T. (2008). "Dual-Processing Accounts of Reasoning, Judgment, and Social Cognition," *Annual Review of Psychology* 59: 255-278.

Gallese, V. and Goldman, A. (1998). "Mirror neurons and the simulation theory of mindreading." *Trends in Cognitive Sciences* 2: 493-501.

Goldman, A. and Sripada, C. S. (2005). "Simulationist models of face-based emotion recognition." *Cognition* 94: 193-213.

Goldman, A. (2006). *Simulating Minds: The Philosophy, Psychology, and Neuroscience of Mindreading*. New York: Oxford University Press.

Goldman, A. (2015). "Naturalizing metaphysics with the help of cognitive science." In K. Bennett and D. Zimmerman (eds.) *Oxford Studies in Metaphysics*, vol. 8, pp. 171-214. Oxford: Oxford University Press.

Greene, J. D. , Sommerville, R. B. , Nystrom, L. E. , Darley, J. M. , and Cohen, J. D. (2001). "An fMRI investigation of emotional engagement in moral judgment." *Science* 293: 2105-8.

Kahneman, D. (2011). *Thinking, Fast and Slow*. New York: Farrar, Straus and Giroux.

Leslie, S. J. (2013). "Essence and natural kinds." In T. Gendler and J. Hawthorne (eds.), *Oxford Studies in Epistemology*, vol. 4. Oxford: Oxford University Press.

McGrath, M. (2015). "Perceptual justification," in A. Goldman and M. McGrath, *Epistemology: A Contemporary Introduction*. New York: Oxford University Press.

Narenzayan, A. , Gervais, W. M. and Trzesniewski, K. H. (2012). "Mentalizing deficits constrain belief in a personal god," *Plos One* 7(5): 1-8.

Palmer, S. (1999). *Vision Science: From Photons to Phenomenology*. Cambridge, MA: MIT Press.

Paul, L. A. (2010). "Temporal experience," *Journal of Philosophy* 107(7): 333-359.