

山西大学
认知哲学丛书
魏屹东/主编

心智的 非机械论研究

Φιλοσοφία



常照强/著



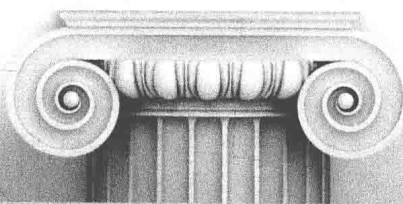
科学出版社

教育部人文社会科学重点研究基地
山西大学“科学技术哲学研究中心”基金
山西省优势重点学科基金
资 助

山西大学
认知哲学丛书
魏屹东 主编

心智的 非机械论研究

常照强/著



科学出版社
北 京

图书在版编目(CIP)数据

心智的非机械论研究 / 常照强著. —北京: 科学出版社, 2016.3

(认知哲学丛书 / 魏屹东主编)

ISBN 978-7-03-047598-5

I. ①心… II. ①常… III. ①认识论—研究 IV. ①B017

中国版本图书馆 CIP 数据核字 (2016) 第 047062 号

丛书策划: 侯俊琳 牛 玲

责任编辑: 朱萍萍 刘巧巧 / 责任校对: 彭 涛

责任印制: 张 伟 / 封面设计: 无极书装

编辑部电话: 010-64035853

E-mail: houjunlin@mail.sciencep.com

科学出版社出版

北京东黄城根北街 16 号

邮政编码: 100717

<http://www.sciencep.com>

北京京华虎彩印刷有限公司 印刷

科学出版社发行 各地新华书店经销

*

2016 年 3 月第 一 版 开本: 720 × 1000 B5

2016 年 3 月第一次印刷 印张: 11 1/4

字数: 224 000

定价: 62.00 元

(如有印装质量问题, 我社负责调换)

从 书 序

21 世纪以来，在世界范围内兴起了一个新的哲学研究领域——认知哲学（philosophy of cognition）。认知哲学立足于哲学反思认知现象，既不是认知科学，也不是认知科学哲学、心理学哲学、心灵哲学、语言哲学和人工智能哲学的简单加合，而是在梳理、分析和整合各种以认知为研究对象的学科的基础上，立足于哲学（如语境实在论）反思、审视和探究认知的各种哲学问题的研究领域。认知哲学不是直接与认知现象发生联系，而是通过以认知现象为研究对象的各个学科与之发生联系。也就是说，它以认知概念为研究对象，如同科学哲学是以科学为对象而不是以自然为对象，因此它是一种“元研究”。

在这种意义上，认知哲学既要吸收各个相关学科的理论成果，又要有自己独特的研究域；既要分析与整合，又要解构与建构。它是一门旨在对认知这种极其复杂的心理与智能现象进行多学科、多视角、多维度整合研究的新兴研究领域。认知哲学的审视范围包括认知科学（认知心理学、计算机科学、脑科学）、人工智能、心灵哲学、认知逻辑、认知语言学、认知现象学、认知神经心理学、进化心理学、认知动力学、认知生态学等涉及认知现象的各个学科中的哲学问题，它涵盖和融合了自然科学和人文科学的不同分支学科。

认知哲学之所以是一个整合性的元哲学研究领域，主要基于以下理由：

第一，认知现象的复杂性，决定了认知哲学研究的整合性。认知现象既是复杂的心理与精神现象，同时也是复杂的社会与文化现象。这种复杂性特点必然要求认知科学是一门交叉性和综合性的学科。认知科学一般由三个核心分支学科（认知心理学、计算机科学、脑科学）和三个外围学科（哲学、人类学、语言学）构成。这些学科不仅构成了认知科学的内容，也形成了研究认知现象的不同进路。系统科学和动力学介入对认知现象的研究，如认知的动力论、感知的控制论和认知的复杂性研究，极大地推动了认知科学的发展。同时，不同

学科之间也相互交融，形成新的探索认知现象的学科，如心理学与进化生物学交叉产生的进化心理学，认知科学与生态学结合形成的认知生态学，神经科学与认知心理学结合产生的认知神经心理学，认知科学与语言学交叉形成的认知语义学、认知语用学和认知词典学。这些新学科的产生增加了探讨认知现象的新进路，也说明对认知现象本质的揭示需要多学科的整合。

第二，认知现象的根源性，决定了认知哲学研究的历史性。认知哲学之所以能够产生，是因为认知现象不仅是心理学和脑科学研究的领域，也历来是哲学家们关注的焦点。这里我粗略地勾勒出一些哲学家的认知思想——奥卡姆（Ockham）的心理语言、莱布尼茨（G.W.Leibniz）的心理共鸣、笛卡儿（R.Descartes）的心智表征、休谟（D.Hume）的联想原则（相似、接近和因果关系）、康德（I.Kant）的概念发展、弗雷格（F.Frege）的思想与语言同构假定、塞尔（J.R.Searle）的中文屋假设、普特南（Hilary W. Putnam）的缸中之脑假设等。这些认知思想涉及信念形成、概念获得、心理表征、意向性、感受性、心身问题，这些问题与认知科学的基本问题（如智能的本质、计算表征的实质、智能机的意识化、常识知识问题等）密切相关，为认知科学基本问题的解决奠定了深厚的思想基础。可以肯定，这些认知思想是我们探讨认知现象的本质时不可或缺的思想宝库。

第三，认知科学的科学性和人文性，决定了认知哲学研究的融合性。认知科学本身很像哲学，事实上，认知科学的交叉性与综合性已经引发了科学哲学的“认知转向”，这在一定程度上从认知层次促进了自然科学与人文科学、科学主义与人文主义的融合。我认为，在认知层面，科学和人文是统一的，因为科学知识和人文知识都是人类认知的结果，认知就像树的躯干，科学和人文就像树的分枝。例如，对认知的运作机制及规律、表征方式、认知连贯性和推理模型的研究，势必涉及逻辑分析、语境分析、语言分析、认知历史分析、文化分析、心理分析、行为分析，这些方法的运用对于我们研究心灵与世界的关系将大有益处。

第四，认知现象研究的多学科交叉，决定了认知哲学研究的综合性。虽然认知过程的研究主要是认知心理学的认知发展研究、脑科学的认知生理机制研究、人工智能的计算机模拟，但是科学哲学的科学表征研究、科学知识社会学的“在线”式认知研究、心灵哲学的意识本质、意向性和心脑同一性的研究，也同样值得关注。因为认知心理学侧重心理过程，脑科学侧重生理过程，人工智能侧重机器模拟，而科学哲学侧重理性分析，科学知识社会学侧重社会建构，

心灵哲学侧重形而上学思辨。这些不同学科的交叉将有助于认知现象的整体本质的揭示。

第五, 认知现象形成的语境基底性, 决定了认知哲学研究的元特性以及采取语境实在论立场的必然性。拉考夫 (G.Lakoff) 和约翰逊 (M.Johnson) 认为, 心灵本质上是具身的, 思维大多是无意识的, 抽象概念大多是隐喻的。我认为, 心理表征大多是非语言的 (图像), 认知前提大多是假设的, 认知操作大多是建模的, 认知推理大多是基于模型的, 认知理解大多是语境化的。在人的世界中, 一切都是语境化的。因此, 立足语境实在论研究认知本身的意义、分类、预设、结构、隐喻、假设、模型及其内在关系等问题, 就是一种必然选择, 事实上, 语境实在论在心理学、语言学和生态学中的广泛运用业已形成一种趋势。

需要指出的是, 与“认知哲学”极其相似也极易混淆的是“认知的哲学” (cognitive philosophy)。在我看来, “认知的哲学”是关于认知科学领域所有论题的哲学探究, 包括意识、行动者和伦理, 最近关于思想记忆的论题开始出现, 旨在帮助人们通过认知科学之透镜去思考他们的心理状态和他们的存在。在这个意义上, “认知的哲学”其实就是“认知科学哲学”, 与“认知哲学”相似但还不相同。我们可以将“cognitive philosophy”译为“认知的哲学”, 将“philosophy of cognition”译为“认知哲学”, 以便将二者区别开来, 就如同“scientific philosophy” (科学的哲学) 和“philosophy of science” (科学哲学) 有区别一样。“认知的哲学”是以认知 (科学) 的立场研究哲学, “认知哲学”是以哲学的立场研究认知, 二者立场不同, 对象不同, 但不排除存在交叉和重叠。

如果说认知是人们如何思维, 那么认知哲学就是研究人们思维过程中产生的各种哲学问题, 具体包括以下十个基本问题。

(1) 什么是认知, 其预设是什么? 认知的本原是什么? 认知的分类有哪些? 认知的认识论和方法论是什么? 认知的统一基底是什么? 有无无生命的认知?

(2) 认知科学产生之前, 哲学家是如何看待认知现象和思维的? 他们的看法是合理的吗? 认知科学的基本理论与当代心灵哲学范式是冲突的还是融合的? 能否建立一个囊括不同学科的、统一的认知理论?

(3) 认知是纯粹心理表征还是心智与外部世界相互作用的结果? 无身的认知能否实现? 或者说, 离身的认知是否可能?

(4) 认知表征是如何形成的? 其本质是什么? 有没有无表征的认知?

(5) 意识是如何产生的? 其本质和形成机制是什么? 它是实在的还是非实

在的？有没有无意识的表征？

（6）人工智能机器是否能够像人一样思维？判断的标准是什么？如何在计算理论层次、脑的知识表征层次和计算机层次上联合实现？

（7）认知概念（如思维、注意、记忆、意象）的形成的机制和本质是什么？其哲学预设是什么？它们之间是否存在相互作用？心－身之间、心－脑之间、心－物之间、心－语之间、心－世之间是否存在相互作用？它们相互作用的机制是什么？

（8）语言的形成与认知能力的发展是什么关系？有没有无语言的认知？

（9）知识获得与智能发展是什么关系？知识是否能够促进智能的发展？

（10）人机交互的界面是什么？人机交互实现的机制是什么？仿生脑能否实现？

当然，在认知发展中无疑会有新的问题出现，因此认知哲学的研究域是开放的。

在认知哲学的框架下，本丛书将以上问题具体化为以下论题。

（1）人工智能的语境范式。在语境论视野下，将表征和计算作为人工智能研究的共同基础，用概念分析方法将表征和计算在人工智能中的含义与其在心灵哲学、认知心理学中的含义相区别，并在人工智能的符号主义、联结主义及行为主义这三个范式的具体语境中厘清这两个核心概念的具体含义及特征，从而使人工智能哲学与心灵哲学区别开来，并基于此建立人工智能的语境范式来说明智能的认知机制。

（2）后期维特根斯坦（L. Wittgenstein）的认知语境论。维特根斯坦作为20世纪的大哲学家，其认知思想非常丰富，且前后期有所不同。对前期维特根斯坦的研究大多侧重于其逻辑原子论，而对其后期的研究则侧重于语言哲学、现象学、美学的分析。从语言哲学、认知科学和科学知识社会学三方面来探讨后期维特根斯坦的认知语境思想，无疑是认知哲学研究的一个重要内容。

（3）智能机的自语境化认知。用语境论研究认知是回答以什么样的形式、基点或核心去重构认知哲学未来走向的一个重大问题。通过构建一个智能机自语境化模型，对心智、思维、行为等认知现象进行说明，表明将智能机自语境化认知作为出发点与落脚点，就是以人的自语境化认知过程为模板，用智能机来验证这种演化过程的一种研究策略。这种行为对行为的验证弥补了以往“操作模拟心灵”的缺陷，为解决物理属性与意识概念的不搭界问题提供了新思路。

（4）意识问题的哲学分析。意识是当今认知科学中的热点问题，也是心灵哲

学中的难点问题。以当前意识研究的科学成果为基础,从意识的本质、意识的认知理论及意识研究的方法论三个方面出发,以语境分析方法为核心探讨意识认知现象中的哲学问题,提出了意识认知构架的语境模型,从而说明意识发生的语境发生根源。

(5) 思想实验的认知机制。思想实验是科学创新的一个重要方法。什么是思想实验?它们怎样运作?在认知中起什么作用?这些问题需要从哲学上辨明。从理论上理清思想实验在哲学史、科学史与认知科学中的发展,有利于辨明什么是思想实验,什么不是思想实验,以及它们所蕴含的哲学意义和认知机制,从而凸显思想实验在不同领域中的作用。同时,借助思想实验的典型案例分析,构建基于思想实验的认知推理模型,这有利于在跨学科的层面上探讨认知语言学、脑科学、认知心理学、人工智能、心灵哲学中思想实验的认知机制。

(6) 心智的非机械论。作为认知哲学研究的显学,计算表征主义的确将人类心智的探索带入一个新的境界。然而在机械论观念的束缚下,其“去语境化”和“还原主义”倾向无法得到遏制,因而屡遭质疑。因此,人们自然要追问:什么是更为恰当的心智研究方式?面对如此棘手的问题,从世界观、方法论和核心观念的维度,从“心智、语言和世界”整体认知层面,凸显新旧两种研究进路的分歧和对立,并在非机械论框架中寻求一个整合心智和意义的突破点,无疑具有重大意义。

(7) 丹尼特(D.Dennett)的认知自然主义。作为著名的认知哲学家,丹尼特基于自然主义立场对心智和认知问题进行的研究,在认知乃至整个哲学领域都具有重大意义。从心智现象自然化的角度对丹尼特的认知哲学思想进行剖析,弄清丹尼特对意向现象进行自然主义阐释的方法和过程,说明自由意志的自然化是意识自然化和认知能力自然化的关键环节。

(8) 意识的现象性质。意识在当代物理世界中的地位是当代认知哲学和心灵哲学中的核心问题。而意识的现象性质又是这一问题的核心,成为当代心灵哲学中物理主义与反物理主义争论的焦点。在这场争论中,物理主义很难坚持纯粹的物理主义一元论,因为物理学只谈论结构关系而不问内在本质。当这两个方面都和现象性质联系在一起时,物理主义和二元论都看到了希望,但作为微观经验的本质如何能构成宏观经验,这又成了双方共同面临的难题。因此,考察现象性质如何导致了这样一系列问题的产生,并分析了意识问题可能的解决方案与出路,就具有重要意义了。

(9) 认知动力主义的哲学问题。认知动力主义被认为是认知科学中区别于认知主义和联结主义的、有前途的一个研究范式。追踪认知动力主义的发展动向,通过比较,探讨它对于认知主义和联结主义的批判和超越,进而对表征与非表征问题、认知动力主义的环境与认知边界问题、认知动力主义与心灵因果性问题进行探讨,凸显了动力主义所涉及的复杂性哲学问题,这对于进一步弄清认知的动力机制是一种启示。

本丛书后续的论题还将对思维、记忆、表象、认知范畴、认知表征、认知情感、认知情景等开展研究。相信本丛书能够对认知哲学的发展做出应有的贡献。

魏屹东

2015年10月13日

前言

20 世纪 50 年代以来，心智研究逐渐突破了行为主义的藩篱，催生了认知主义这一主流范式。作为对行为主义的一种修正，认知主义放弃了一切否认内部心理状态的狭隘观点，转而以内在的符号表征作为其本质特征。但是，近年来，一股“后认知主义”潮流愈演愈烈，其矛头直指认知主义的核心概念——表征。在这一潮流的裹挟之下，人们自然要对心智研究进行一系列的批判性反思，不免要追问：后认知主义潮流的本质何在？什么是一种更为恰当的心智研究方式？面对如此棘手的问题，本书以“心智机械论进路”和“心智非机械论进路”这两个观念为楔子，探究一二。

心智机械论进路脱胎于机械论世界图景。在这一世界图景中，人们一般认为事物的运动遵循着一定的自然规律，逐渐形成了“自然即机器”的隐喻，并且这一隐喻进一步演化为“心智即机器”的隐喻，心智则被认为亦属于自然机制的范畴。伴随着计算机科学、心理学等学科的突破性进展，“心智即机器”的隐喻获得了强有力的科学支撑，逐渐从边缘走向主流，最终为心智机械论进路设定了一个历史背景。在此背景下，一种心智的计算表征主义观念逐渐形成，并最终成为心智机械论进路的理论基础。在实践中，以各种说明隐喻为蓝本，人们构建了多层次的心智计算表征模型，并使之体现这些说明隐喻的本质特征，借此将隐喻机械化地予以实现。

此外，我们还可从以下两个方面来审视心智机械论进路。一方面，它与图灵机模型有着莫大的渊源，因而包含了一种传统的计算观念，即形式化的计算观念。另一方面，它与表征观念有着密切的关联，因而包含了一种基于表征的信息观念，即德雷斯基（Fred Dretske）信息观念。

毋庸讳言，心智机械论进路在实践中获得了巨大的成功，有助于我们将心智纳入科学研究的视野，并将其归于因果说明的范畴，然而在肯定其历史地位

的同时，我们也不能忽视其所面临的主要困境——心智意义的缺失。这一困境的主要根源即是计算表征主义的“标准解释”。在“标准解释”的引导之下，我们误以为心智计算表征理论能够起到“将计算理论与民俗心理学相融合”的作用，换言之，我们一相情愿地将计算系统中的符号表征视为信念、欲望、想法等常识观念的“科学对等物”。

尽管心智机械论进路为摆脱其困境而采纳了一些新的理念，但其终究无法回答意义难题，本质上依然是一种关于心智的碎片化理解方式。一方面，在心智机械论进路中，我们所机械化的无非是人类理性的一个或几个方面而已，因此我们无法回应笛卡儿的责难。另一方面，在心智机械论进路中，我们无法做到将智能行为与有意义的信息绑定起来，所以我们是在以一种碎片化的方式来审视信息、心智和现实之间的本质而已。

为了重获心智缺失的意义，我们一方面要走出中文屋困境，引入非机械论物理学和现象学的有关概念，另一方面要回归皮尔斯的整体性命题，即将“心智、语言和世界”三者视为一个整体，从而在一个整体框架下进行心智研究进路的重构。为此，我们通过两方面的重构而为之，即“激进的具身认知”和“第二次认知革命”。借此，我们找到了一个有别于机械论内涵的心智研究进路，即心智非机械论进路。

事实上，心智研究进路重构的过程即是一个转向的过程。从心智机械论进路到心智非机械论进路，我们不仅实现了世界观和方法论的转向，而且实现了计算和信息观念的转向。最终，我们完成了心智研究的后现代转向。

在心智机械论进路中，人们囿于一种机械论的世界观，固守着一种科学理论的经典结构观念，由于这一结构观念蕴涵着还原主义的思想，因而继承了还原主义所带来的缺陷。与之相对照，在心智非机械论进路中，我们则接受一种非机械论的世界观。基于这一新的世界观，我们放弃了经典结构观念，转而推崇一种整体的动力学方法。

为了破解意义难题，在心智非机械论进路中我们引入了一种自然计算观念，从而将心灵关联于计算观念的源头。与传统的计算观念不同，这种自然计算本质上是一种互动计算。另一方面，借鉴德雷斯基信息观念的优势，同时又弥补了其不足，心智非机械论进路还引入了一种整体信息观念。鉴于心智牵涉着一个动力学的、复杂的关系网络，这种整体信息观念特别强调意义的整体性，反对以任何碎片化的方式来审视信息、心智和现实的本质。换言之，它认为不应将心智从信息的疆域中剔除出来，反而应将其关联于信息过程的源头。

事实上，通过心智非机械论进路的探索，我们最终实现了心智研究的后现代转向，其包含着三重意蕴：一是提出了心智研究的新范式——互助论；二是由心智的机械模型转向心智的自组织模型；三是鉴于身体的脆弱性构成了生物体意义的基础，提出一种有机体意义的观念。

值得注意的是，类似于应对现代主义和后现代主义之间的关系，在认识和评价这两种心智研究进路之时，我们应该保持一种持平的态度。作为一个具有整体性特质的心智研究进路，心智非机械论进路无疑是处理复杂认知系统的最佳方式。但是，同样是这种整体性特质使然，我们在研究人类智能和行为时，在可资利用的一系列复杂的工具之中，机械模型策略仍然是其中不可忽略的一部分，应该审慎地对待。

放眼当今的心智哲学研究，主流学者既不愿重蹈还原论物理主义的覆辙，又想极力避免落入笛卡儿二元论的陷阱，因而迫切地寻求二者之外的第三条道路。心智非机械论进路的提出大体也可以看作这一方向的尝试和努力。无可否认，这一进路才刚刚起步且前途未卜，但是我们希冀它至少能有些许的价值，即示人以启发：若没有形而上学预设、方法论和核心观念等方面的剧烈蜕变，第三条道路的追求恐怕会失之根基而无法成功。

最后，书中难免有谬误之处，敬请读者的指正，是所至盼。

常照强

2016年1月

目 录

丛书序	i
前言	vii
绪论	1
第一节 研究的旨趣	1
第二节 国内外研究现状述评	6
第三节 核心概念的界定	12
一、心智机械论进路	12
二、心智非机械论进路	16
第四节 思路和内容	18
一、本书的思路	18
二、本书的内容	20
第五节 创新与不足	22
第一章 心智机械论进路	23
第一节 机械论的世界图景	23

一、机械论观念的由来	24
二、机械论的隐喻	25
第二节 从心智的表征到心智的计算	27
一、心智的表征	27
二、心智的计算	27
三、心智的计算表征理论	29
第三节 从隐喻到机械模型	30
一、心智研究的隐喻观念	30
二、逻辑隐喻的机械化实现	31
三、比例模型隐喻的机械化实现	33
第四节 从图灵机到形式化的计算	34
一、图灵机模型	34
二、经典的计算观念	36
第五节 基于表征的信息观念	37
一、心智信息观念的介入	37
二、德雷斯基的信息观念	38
第二章 心智机械论进路的评估	41
第一节 历史的定位	41
一、心智的因果描述	42
二、身心问题的回应	44
三、心智的科学研究	47
第二节 陷入“标准解释”误区	48
一、何谓“标准解释”	49
二、“标准解释”的误导作用	51
第三节 意义缺失的批判	53
一、塞尔的质疑	53
二、德雷福斯的责难	55

三、皮奇尼尼的疑问	56
第三章 心智研究进路的重构	58
第一节 心智研究新观念	59
一、下向因果观念的运用	59
二、生物学启发思想的介入	62
三、进化观念的渗透	63
第二节 元规则：一种伪具身进路	64
一、元规则策略	65
二、笛卡儿的启示	68
第三节 重构的必要性	70
一、回归整体性命题	70
二、走出中文屋困境	73
第四节 重构的两个向度	76
一、向度一：激进具身层面	77
二、向度二：社会文化层面	86
第五节 心智研究的中间地带	92
第四章 心智非机械论进路	95
第一节 非机械论世界图景	96
一、非机械论物理学	97
二、整体的世界观	98
三、非机械论的心物观	101
第二节 方法论的嬗变：从“结构”到“动力学”	110
一、经典的结构观念	111
二、结构的消解	116
三、动力学关系的确立	119
四、一个案例：从结构表征到动力学非表征	121

第三节 计算观念的新范式	128
一、何谓“自然计算”	129
二、范式转向：互动计算	130
第四节 走向有意义的信息	131
一、整体信息观念	131
二、隐缠序和信息观念	135
第五节 心智研究的后现代转向	138
一、心智研究的新范式：互助论	138
二、从机械模型到自组织模型	141
三、走向“有机体意义”	143
 结束语	 148
 参考文献	 151
 后记	 161

绪 论

第一节 研究的旨趣

谈论本书的研究旨趣，就不能不提到所谓的“中文屋论证”（the Chinese room argument, CRA），或者更为准确地说“升级版的中文屋论证”。任何接触过心智哲学的人，对于中文屋论证可以说都并不陌生，这一思想实验令它的提出者约翰·塞尔（John R. Searle）享誉哲学界。但是，对于升级版的中文屋论证，恐怕许多人知之甚少。我们之所以称其为“升级版的中文屋论证”，原因在于这一所谓的“新论证”（the new argument, NA）从另一个角度批判了强人工智能，而且在塞尔看来，“中文屋论证”与“新论证”在本质上有着巨大的差异。个中原因在于，前者意图表明：语义并非内在于句法；而后者意图表明：句法并非内在于物理。所以，塞尔认为，这就意味着“新论证”较之于“中文屋论证”更为深刻^①。

在“新论证”中，塞尔有一个重要的预设，即计算是与观察者相关的。塞尔之所以提出这一预设，主要基于两个方面的理由：一是句法是与观察者相关的；二是使得计算具有计算特征的一切事物都是句法的。由于句法是与观察者相关的，所以句法符号的物理属性不会使其独立于任何人的意向性^②。

有了这样的一个预设，塞尔即可顺理成章地进行以下步骤的论证。首先，根据传统定义，凭借符号的操作，计算通过句法而得以定义。其次，句法和符号并不能凭借物理来定义，尽管字符（symbol tokens）总是一些物理的符号，但是字符并不能通过物理特征来定义。简而言之，句法并不内在于物理。再次，计算不能在物理中被发现，其不过是被分派的，这样使得某些物理现象能够通

^① John R. Searle. The Mystery of Consciousness[M]. New York: The New York Review of Books, 1997: 17.

^② John R. Searle. The Mystery of Consciousness[M]. New York: The New York Review of Books, 1997: 239.