

物理因果闭合性、罗素主义泛心论与心理—神经同构性^{*}

王 帅 翔

摘 要:物理世界的因果闭合性是关于心灵的物理主义受到广泛支持的理由。物理主义者认为物理学和神经生理学的经验研究表明,物理世界中的一切作用力都是基于基本物理作用力的因果作用。但这给罗素主义的泛心论留下了逻辑上的空间。因此物理因果闭合性论证难以为关于心灵的物理主义提供决定性的支撑。神经科学的大量经验证据表明,意识经验的现象结构同构于相应“神经关联物”的物理结构。对于这类神经—心理同构性现象,物理主义提供了比罗素主义泛心论更好的解释,因此仍有理由认为物理主义是心身问题上可行的形而上学主张。

关键词:物理因果闭合性;罗素主义泛心论;神经—心理同构性;最佳解释推理

DOI: 10. 11714/jysu. sse. 202604015

在当代分析哲学领域,物理主义是大部分哲学家在心身关系问题上持有的立场。概括地讲,关于心灵的物理主义认为一切意识体验与意向状态都不过是大脑中的神经结构和动力学活动。关于心智的物理主义之所以能够获得普遍的接受和拥护,无疑在于一种有赖于经验证据的论证,这就是“因果闭合性论证”(the argument from causal closure)。该论证的前提是:现实的物理世界是因果上闭合的;心灵具有因果效力;在心物因果中不存在因果过度决定(overdetermination),由此可以得出心灵就是“物理的”结论^①。这个论证中重要的前提是物理世界的因果闭合性。物理主义者支持物理因果闭合性的理由在于物理学和神经生理学提供的经验证据,因为这些科学的经验研究没有发现任何特有精神作用力的证据。但是,基于物理学基本作用力的因果作用本质上仍有可能是非物理的心理因果作用,这就是某些版本构成性罗素主义泛心论所呈现的图景。因此,“因果闭合性论证”难以为物理主义提供决定性的支撑。

但本文认为物理主义者仍有回应这个挑战的对策,这是因为神经科学的大量研究提供了神经活动结构与意识结构同构性的证据。相对于构成性的罗素主义泛心论,物理主义能为这些现象提供更好的解释。基于最佳解释推理,我们仍有理由认为物理主义是一种关于心智的可行形而上学方案。

一、物理因果闭合性与罗素主义泛心论

物理主义的拥护者认为,来自经验科学的证据为物理因果闭合性提供了强有力的归纳支持^②。物理学将所有表面上的特殊作用力(如浮力、推力和摩擦力等)都还原地解释为少数几类保存能量的基本物

^{*} 收稿日期:2026—03—09

作者简介:王帅翔,北京大学哲学系(北京 100871)。

① Daniel Stoljar, “Physicalism”, *The Stanford Encyclopedia of Philosophy* (Spring 2024 Edition), Edward N. Zalta & Uri Nodelman eds., URL = <<https://plato.stanford.edu/archives/spr2024/entries/physicalism/>>.

② David Papineau, *Thinking about Consciousness*, Oxford: Oxford University Press, 2002, pp. 232–256.

理作用力(即引力、电磁力、强相互作用以及弱相互作用)。生理学发现引起各种身体行为的只是神经电活动(即电磁力),并未发现任何特有的(*sui generis*)“精神力”或“生命活力”。据此,“因果闭合性论证”已经预设了一种对什么是“物理因果作用”的理解。根据这种理解,所谓“物理因果作用”就是基于已知的(保存能量的)基本物理作用力(即引力、电磁力、强相互作用和弱相互作用)且传递守恒量(能量、动量、电荷等)的因果相互作用。但是,这种诉诸基本力的说明难以作为界定“物理因果作用”的充分条件,因为它为构成性的罗素主义泛心论(*constitutive Russellian panpsychism*)^①留下了逻辑上的空间。

构成性的罗素主义泛心论认为,现实世界由物理学告诉我们的理论实体(如基本粒子)构成,现实世界的因果作用也由基本的物理作用力所支配,但在这些基础的物理性质和作用力背后却“潜藏”着原始的心理性质,其凭借与物理作用力的“完美接合”在物理世界中发挥着因果影响^②。也就是说,在构成性罗素主义泛心论者看来,尽管物理世界的一切因果作用都是基于基本作用力的因果作用,但原始的心理性质(特别是现象性质)与基础物理学所推设的结构与倾向性质之间(如动量和能量)存在特定的形而上学关系,因此原始的心理性质同样发挥了因果上的效用。正如高夫(Philip Goff)等论者所说:

如果像二元论者所认为的那样,意识存在于物理世界之外,那么就很难看出它何以影响因果闭合的物理系统。但如果像泛心论者所相信的那样,意识融入了物质世界的内在本质,那么意识及其结果就是因果闭合系统的一部分。^③

因此,根据罗素主义泛心论的基本观点,由于基础物理学告诉我们的结构和倾向性质(如能量、电荷等)凭借特定的形而上学关系(如随附关系)奠基(依赖)于基础现象性质,因此在涉及守恒量传输的基本力因果作用中,基础现象性质仍然是具有因果相关性的。由于基础现象性质的因果相关性,基于物理基本作用力的因果作用也就不再是纯粹的物理因果作用了。

反驳者大概会说,即便接受了罗素主义的泛心论,基础的现象性质仍然是因果上不相关的,因此也就是副现象的。豪威尔(Robert J. Howell)就试图通过“可设想性论证”(the conceivability argument)说明基础现象性质与结构(倾向)性质之间并不存在形而上学的必然关联,因此基础现象性质是因果上不相关的。对此他给出了如下的论证说明:

想象某个世界w1,在其中R,也就是现象红(phenomenal redness),基于w1中统摄R的因果定律奠基了负电荷性质。再来想象另一个世界w2,在其中G,也就是现象绿(phenomenal greenness),由同样的定律所覆盖。因此G奠基了与负电荷相关联的因果效力,而R相反却奠基了与负自旋相关联的因果效力。最后再来想象第三个世界w3,其中的定律是这样的一种情况:要么R要么G能够奠基负电荷的因果效力——R和G都由完全相同的定律以完全相同的方式所统摄。^④

① 罗素主义泛心论一般可区分为构成性(*constitutive*)的版本和非构成性的版本。构成性的版本认为基础物理学层面的现象性质凭借特定组合关系构成生物大脑层面的现象意识,前者与后者之间具有特定非因果的解释性关系。非构成性的版本则认为,基础物理层面的现象性质在动态因果活动中突现出有机体的现象意识。由于意识的突现涉及特有的心理作用力,因此意识的突现意味着物理因果闭合性的破坏。据此,物理因果闭合性论证足以在逻辑上排除非构成性(突现论)版本的罗素主义泛心论。在未经特别说明的情况下,本文所讨论的罗素主义泛心论专指构成性的罗素主义泛心论。

② Torin Alter and Derk Pereboom, “Russellian Monism”, *The Stanford Encyclopedia of Philosophy* (Fall 2019 Edition), Edward N. Zalta ed., URL = <<https://plato.stanford.edu/archives/fall2019/entries/russellian-monism/>>.

③ Philip Goff, William Seager and Sean Allen-Hermanson, “Panpsychism”, *The Stanford Encyclopedia of Philosophy* (Summer 2022 Edition), Edward N. Zalta ed., URL = <<https://plato.stanford.edu/archives/sum2022/entries/panpsychism/>>.

④ Robert Howell, “The Russellian Monist’s Problems with Mental Causation”, *The Philosophical Quarterly*, 2015, 65 (258): 22–39.

豪威尔试图通过这个可设想性论证说明,两个在因果结构(倾向)上完全同构的世界却可能具有不同的奠基性的基础现象性质,因而特定的结构(倾向)性质和特定的基础现象性质之间并不具有必然的形而上学关联,因此现象性质在因果上也就是不相关的了。除此之外,我们似乎还可以设想两个在因果结构上同构的世界,其中一个具有奠基性的基础现象性质,而另一个世界则是没有任何奠基性范畴性质的纯因果结构世界,因此结构(倾向)性质“承担”了所有的“因果工作”,而基础现象性质只是多余的“副现象”。这样一来,由于基础现象性质因果上的不相关性,“物理因果闭合性原则”得到了保全。

那么这个反驳意见真的成立吗?笔者高度怀疑“可设想性”是通达形而上学可能性的路径,鉴于这是一个极为复杂的问题,本文难以就此展开讨论。但即便抛开这一点,豪威尔的论断仍旧面临着其他责难^①。在笔者看来,豪威尔论证的最大问题就在于他忽视了基础现象性质与物理倾向(结构)性质之间确实存在着形而上学必然关系的可能性,而“可设想性论证”所揭示出的或然性的表象不过是出于我们对这种形而上学必然关系的无知。倘若我们理解了这样的形而上学必然关系,豪威尔所说的情形自然是不可设想的。对此,我们需要对绝对(质的)性质与倾向性质的基础本体论问题进行说明。根据罗素主义一元论者的基本立场,他们所能接受的倾向和范畴(质的)属性本体论可归结为如下两大类^②:

(1)倾向和范畴(质的)性质的类型二元论:根据这一观点,存在范畴(质)和倾向两类性质。若某一性质适用“是倾向的”或“是范畴(质)的”这一谓词,则这一谓词适用于该性质的全部,而不是该性质的特定部分。这种对性质的划分既是互斥的,也是穷尽的,某一性质要么是完全的倾向性质,要么是完全的范畴(质的)性质。罗素主义泛心论者如果采纳了这种观点,则会认为某基础物理性质(如负电荷) c_1 与某种现象性质 p_1 之间具有特定的形而上学依赖关系,如随附性等。

(2)倾向—范畴(质)的双面一元论:这类学说认为倾向的与质的描述范畴(谓词)并不是互斥的,同一个性质的全部既适用于倾向的描述也适用于质的描述^③。这种学说也被称为“效力性的质”(powerful qualities)学说。这就好比我们熟知的格式塔转换,对于同一图案,其既可被描述为“兔子头”也可被描述为“鸭子头”。对上述概念、语言界说的形而上学解读就是同一性质兼具质的方面和倾向的方面,任何一种单一的描述范畴都仅仅揭示了该性质部分而不是全部的本质。性质的两个方面并不是等同的,虽然我们能够凭借思维的抽象仅仅描述性质的其中一个方面,但是这种描述并不是形而上学意义上完备的。而且质的方面和倾向的方面虽然是同一个性质的不同方面,但是它们之间却存在着形而上学意义上的必然关联^④。正如三角形的“三角性”和“三边性”尽管是不同的性质,但是它们之间的共存却是必然的。这样一来,由于某一性质的倾向方面与它质的方面存在着形而上学的必然关联,那么性质的方面自然具

① 比如,批评者可以否认特定的基础现象性质与特定的结构(倾向)性质之间存在必然的形而上学关联,也就是说二者之间的关联仅仅是形而上学上或然的(contingent),但是基础现象性质仍然是因果上相关的。这就是说,因果关系并不要求关系者(relata)之间存在形而上学意义上必然的关联。比如,物理事件A引起物理事件B这一点并不蕴含A在形而上学上必然化(metaphysically necessitates)B,因为或许将A与B关联起来的只是某种形而上学或然的因果定律。当然,这种回应方案需要我们接受关于因果性的特定形而上学说明(如休谟主义的因果观)。

② George Molnar, *Power: A study in metaphysics*, Oxford: Oxford University Press, 2003, pp. 148–157.

③ J. Giannotti, “The identity theory of powers revised”, *Erkenntnis*, 2021, 86: 603–621.

④ 这显然不同于马丁(Charles Martin)与海尔(John Heil)倡导的质—倾向等同论方案。等同论的主张是:“若P为某一具体对象的内在性质(intrinsic property),P同时既是倾向的也是质的;P的倾向性和质性(qualitativity)不是P的方面或性质,P的倾向性 P_d 就是P的质性 P_q ,而且其中的任何一者都是 $P:P_d=P_q=P$ 。”(John Heil, *From an Ontological Point of View*, Oxford: Oxford University Press, 2003, p. 111)等同论方案不仅面临概念矛盾的指责,还有滑向纯倾向本体论之嫌(J. Giannotti, “The identity theory of powers revised”, *Erkenntnis*, 2021, 86: 603–621)。

有了因果上的相关性^①。

由此不难看出,诉诸质与倾向二元论方案的罗素主义泛心论或许要遭受形而上学或然性问题以及因果不相关问题的责难,然而罗素主义泛心论者只要接受了质与倾向的双面一元论就可以轻易规避这些指责。这是因为根据双面一元论,同一性质的倾向方面与质的方面存在必然的形而上学关联,因果不相关的问题自然烟消云散。因此,在接受了质与倾向双面一元论的罗素主义泛心论者看来,物理学告诉我们的电荷、能量等守恒量本身就是基础的现象性质(心理性质)。由此可见,根据这种罗素主义泛心论的观点,基于基本作用力传递守恒量的因果作用就是基础现象性质所参与的因果作用。因此,通过“因果闭合性论证”得出的“物理主义”竟与泛心论在逻辑上相容。我们需要一种更为强有力的论证用以排除构成性罗素主义泛心论的“逻辑空间”。下文将论证,神经结构与意识现象结构同构的证据能为心灵的物理主义提供有力的经验支撑,这需要明确物理主义作为一种形而上学在经验世界中的蕴涵。

二、物理主义与心理—神经同构性

对于在心身问题上秉持物理主义立场的哲学家来说,最自然的选择就是认为心灵同一于或者形而上地依赖于大脑(中枢神经系统)的状态(过程)。换言之,心灵无外乎大脑(中枢神经系统)的状态(过程)。对于“无外乎”,有两种合理的解读:一是殊型同一(token identity)的情形,二是“构成性底定”(constitutive grounding)的情形^②。后者意味着神经结构与心智之间存在着非因果的解释性关系,只要相应的神经结构S1、S2……Sn存在,意识状态P的存在就是“本体论上的免费午餐”(ontological free lunch)。这就好比一场学术组会无外乎不同研究生的报告、讨论以及导师的点评。现在回到心身关系的问题,大部分的物理主义者(还原的和非还原的)显然接受心理状态(过程)殊型同一于或“构成性地底定于”中枢神经系统的状态(过程)。笔者在此将这种立场称为“神经物理主义”(neurophysicalism)。现将“神经物理主义”定义如下:

神经物理主义:任何心理状态都殊型同一于或“构成性地奠基于”中枢神经系统的结构状态;任何心理过程都殊型同一于或“构成性地奠基于”中枢神经系统的动力学过程。^③

在阐述了“神经物理主义”基本观点的基础上,现在来考察这种关于心灵的形而上学论题在经验世界中的蕴涵(implication)。“神经物理主义”虽然并不要求心理—神经状态(过程)殊型之间具有严格意义上的数的同一(numerical identity),但是下面这一点自然是无可争议的:那就是在“神经物理主义”的拥护者看来,在大脑神经元网络中的空间关系等静态关系和神经元发放等动态活动之外并不存在任何额外的“心理实体”(mental entities)。因此基于“神经物理主义”,任何意识状态都殊型同一于或“构成性地奠基于”其对应的“神经关联物”。根据这一点,一个直观的看法当然就是:大脑中的神经结构状态(动力学过程)必定是与相应的意识状态(过程)是相像或相似的;更确切地说,任何意识状态都必定与其相应的“神经关联物”(neural correlates)在结构上是相似的。这一点被称为“心理神经同构性”(psychoneural

① 一种合理的形而上学说明是:对特定性质做出个体化的是其质的方面,基于质的方面该性质得以拥有相应的倾向方面。参见D. Smith, “Quid Quidditism Est?”, *Erkenntnis*, 2016, 81: 237–257。

② Philip Goff, *Consciousness and Fundamental Reality*, Oxford: Oxford University Press, 2017, pp. 42–44.

③ 这个表述方案之所以附加了“结构”和“动力学”(即静态和动态关系)的限定,是为了在逻辑上将“神经物理主义”与“构成式的泛心论”(constitutive panpsychism)或“构成式的泛原心论”(constitutive panprotopsyism)划清界限。在上述两种观点看来,意识显然也是神经系统的状态或过程,但是要“构成地”解释意识,仅仅诉诸神经系统构成物之间的(静态或动态)关系是不够的,在此之外还要“求助于”微观物理层面上绝对的非关系性质。

isomorphism)。正如雷文索所说:“同构性的问题是说如果意识确实‘栖居于’大脑之中,那么在大脑中必定存在着确实与意识相像或者相似的东西——这就是说,意识本身。”^①

如何理解“心理神经同构性”呢?姑且抛开心身问题,精确地说,同构性是一种“双射的同态性”(bijective homomorphism)。所谓“同态性”就是两个对象或“域界”(domains)之间的函数或映射,这两个对象或“域界”之间具有部分相同的结构^②。当然,结构关系的相似性是具有程度区别的。然而,同构性是“同态性”的一种特殊情形,它要求两个对象(或域界)之间结构关系的映射是完全的、一对一的(one-one)。换言之,二者之间的关系结构是精确相似的(exactly similar)。就心理与神经状态(过程)来说,就是基于特定的描述,从心理状态(过程)的结构特征到相应“神经关联”的结构特征之上存在“一对一的映射”(one-one mapping)。可以说,同构性是相似性最强的形式。因此,倘若两个对象的结构是同构的,它们必定是相似的。因此,心理—神经结构的同构性无疑蕴涵了心理—神经结构的相似性。

“心理神经同构性”为何是“神经物理主义”的经验蕴涵呢?这是因为,根据“神经物理主义”的基本立场,任何心理状态(过程)都不过是中枢神经系统的结构状态和动力学过程,尽管未必是严格意义上的数的同一,但是在神经系统的静态和动态结构关系之外不存在任何额外的“心灵”,因此二者在结构关系上具有“一对一”映射是一个自然而然的结论。如雷文索在论及现象意识问题时所言,“当涉及组织的层级时,构成或同一于现象层面的是神经生物层面。因而现象意识与脑中的某一特定组织层级之间必定存在着同构性,这不过是因为它们就是同一个事物”,因此区分意识与神经活动的同构性论断“隐含了某种残余的二元论,它将主观经验与大脑分割开来”^③。我们还可以从解释的角度理解“心理神经同构性”^④,因为不论是还原还是非还原的“神经物理主义”都承认,一切心理现象都能够在神经生物结构和动力学机制的层面上获得解释,而结构上的同构性(至少是相似性)是解释成立的前提条件,否则二者之间的关联无疑是神秘的。

这里有必要对什么是“现象意识的结构特征”作出一些说明。在某些心灵哲学家看来,现象意识之所以对神经生理的、功能的以及计算的还原解释提出了严峻挑战,是因为现象意识具有基于第一人称视角的主观性方面和非关系的质的方面,而任何神经生物、功能和计算的机制都是客观的(基于第三人称视角)以及结构性(关系性)的^⑤。但是,这并不意味着现象意识就不具有相应的结构特征。心灵哲学家查尔默斯(David Chalmers)对此有清晰的阐述:“我们的现象性具有丰富的和特定的结构:它是统一的、边界性的,并被区分为不同的方面,但是却具有涉及许多方面的基础同质性,并且似乎具有一个单一的经验主体。”^⑥通过查尔默斯的表述,我们可以很直观地把握现象意识的结构特征。从静态的方面举例来说,现象意识具有“统一性”,也就是基于各种感觉模态(sensory modality)的感受质(如看到蓝色的质和听见蝉鸣的质)都是一个整体性的“现象场域”的构成组分;在这个统一的“现象场域”中,不同的感受质

①③ Antti Revonsuo, “Prospects for a Scientific Research Program on Consciousness”, in T. Metzinger eds., *Neural Correlates of Consciousness: Empirical and Conceptual Questions*, Cambridge: The MIT Press, 2000, p. 67, p. 67.

② Alfredo Vernazzani, “Psychoneural Isomorphism: From Metaphysics to Robustness”, in Galzavari F. and Viola M., (Eds.), *Neural Mechanism: New Challenges in the Philosophy of Neuroscience*, Switzerland, Springer, 2020: pp. 283-310.

④ Alva Noe and Evan Thompson, “Are There Neural Correlates of Consciousness?”, *Journal of Consciousness Studies*, 2004, 11, No. 1: p. 5.

⑤ Uriah Kriegel, *Subjective Consciousness: A Self-Representational Theory*, Oxford: Oxford University Press, 2009, p. 53.

⑥ David Chalmers, *The Character of Consciousness*, Oxford: Oxford University Press, 2010, p. 136.

之间还具有特定的“现象空间关系”(比如一块红色视觉余像处在一块蓝色余像的“左侧”)。从现象意识的动态结构方面来看,日常清醒状态下的知觉经验具有特定的恒常性、稳定性和确定性,比如当我们看到一朵红色花朵时,会产生相应红色经验的感受质,这种红色的质在我们的“现象场域”中显然是稳定、恒常的,并不会莫名其妙地消失、变化或者被替代;但是在快速眼动睡眠的梦境中,现象意识的结构特征显然不具有日常知觉经验所具有的恒常和稳定性,“现象场域”中各种质的内容会发生莫名其妙的消失、转移和置换,因此梦境在动态的“现象结构”上表现出非恒常、不确定的离奇怪异的特征^①。在接受“心理神经同构性”论断的“神经物理主义”者看来,在中枢神经系统的特定层次上必定存在着与上述“现象结构”同构的神经生物结构(比如在大脑中存在着统一性、整体性的神经活动状态)。需要强调的一点是,“现象结构”与脑中神经生物结构的同构并不一定是空间性的(比如大脑物理空间中局部结构的改变未必带来“现象空间”中质的局部改变,“现象空间”质的改变有可能是全局性的)。

三、神经科学的经验案例与物理主义的解释优势

在说明了“神经物理主义”基本蕴含的基础上,我们需要理解形而上学理论对经验资料的解释。在这个问题上,笔者支持一种温和的方法论自然主义立场,认为和科学假说或理论一样,形而上学理论同样要接受经验证据的检验^②。当然。有别于对科学假说或理论的检验,哲学家并不需要明确地阐述相应的检验条件^③。长久以来,根据科学哲学界的基本共识,若要评估一个假说或理论对相关经验现象的解释力,我们需要考察相应现象的发生是否符合特定假说或理论的预期;如果答案是肯定的,则说明该假说或理论对相关经验现象具有较强的解释力,反之则说明缺乏解释力。科学哲学家索伯(Elliott Sober)对这一点提供了相应的形式化说明:

“惊讶原则”(the surprise principle)描述了在何种情况下,相较于假说(H2)观察O强烈地偏好假说(H1)。对此有两点要求:

(1)如果H1为真,你将预期到O为真;(2)如果H2为真,你将预期到O为假。

这就是说,(1)如果H1为真,O将不会令人感到惊讶;(2)如果H2为真,O将是令人惊讶的。^④

上述说明指明了最佳解释推理的要义,这也是对早先的假说—演绎(hypothetico-deductive)方法的进一步充实和完善。因此,对于经验资料O,若相信H1的观察者不感到惊讶,但相信H2的观察者感到吃惊,那么相比于H2,H1无疑具有对O更强的经验解释力。最佳解释推理的首要原则当然是评估不同假说(理论)对经验资料的解释力,在此基础上进一步考虑假说或理论的精简性以及美观性的特征。

神经科学为意识结构与神经结构的同构提供了明确证据。内省明确地告诉我们,视觉体验中的知觉对象(percept)是一个现象空间中连续的整体,形状、颜色和其他各种特征都融贯并统一地结合在一起。比如当我们看到一只奔跑的白兔时,在我们的现象空间中呈现的是一个具有白色、特定形状以及相应运动特征的统一的整体,并不是各种视觉现象特征“零散”、不连续地分布于现象空间中。换言之,我

① Antti Revonsuo, *Consciousness: The Science of Subjectivity*, Oxford: Psychology Press, 2009, pp. 244–245.

② Robert Almeder, *Harmless Naturalism*, Maryland: Rowman & Littlefield Publishers, 1998: pp. 166–167.

③ 举例来说,假设神经科学家提出了某种关于视觉体验“神经关联物”的经验假说,那么他必定需要明确地给出这个假说的检验条件,比如哪些具体的大脑结构(例如纹状皮层)与视觉体验相关联。相反,对于作为形而上学理论的“神经物理主义”来说,哲学家仅需提出某些笼统的经验可能性,比如大脑中存在着与意识现象结构同构的神经生物结构,但是并不需要明确地指出这些神经结构涉及哪些皮层区域或者与哪些频率的神经元发放活动相关。

④ Elliott Sober, *Core questions in philosophy*, 6th edition, London: Pearson Education, 2012, p. 30.

们的知觉对象具有现象空间意义上的整体性和连续性。然而,从大脑皮层解剖学结构的角度讲,关于颜色、形状等不同特征的视觉信息是由大脑不同的皮层区域处理的。也就是说,单从解剖学的角度讲,对不同感官信号进行处理的神经结构是空间上分散的。但我们的知觉对象却是一个现象空间意义上连贯的整体。这就是当代神经科学所广泛讨论的“特征绑定”(Feature Binding)问题。基于“心理神经同构性”原则,我们自然将预期到对应现象空间层面连续性和整体性的知觉对象,在大脑中也有着物理空间上连续和整体性的神经结构。神经科学家凭借双眼拮抗(binocular rivalry)实验并且运用脑磁图(MEG)发现了这种在大脑的物理空间中整体、连续性的神经结构^①。

研究表明,相同频率神经元振荡(发生于大脑皮层的众多不同区域)的发生贯穿了受试者视觉体验的全部阶段,然而只有当视觉内容(整体性的知觉对象)有意识地呈现时,不同皮层区域神经元的同频率振荡才会出现精确的同步化。也就是说,在受试者有意识地经验知觉对象时,不同脑区的神经元出现了同步化的振荡。内省揭示的视觉知觉对象都是由各种现象特征组合成的融贯整体,基于“心理神经同构性”原则,体验者大脑中必定存在着整体性的神经活动。现在已有大量的神经科学实验印证了这一点,那就是整体性视觉对象经验的“神经关联物”是大脑皮层不同区域同步化的神经元振荡活动。尽管不同视觉信息的处理由解剖学上不同的皮层区域分管,但是当它们同步性地活动时,也就具有了物理空间上的连续性和整体性,这与视觉经验相应的连续性和整体性现象空间结构无疑是同构的。而且,从动态的角度讲,知觉对象的出现和消失这种往复不止的节律对应了大尺度神经元同步化活动的形成与消失。这种大尺度的神经元同步化活动一般会维持数秒,这个时间尺度与新的知觉内容(对象)在体验者现象场域中占据优势地位的时间是吻合的。此外,神经科学研究还找到了许多其他支持神经—心理同构性的证据,这里不予赘述^②。

通过上文对“神经物理主义”经验蕴含的分析,我们会很自然地预期到现象意识状态或过程的结构与其“神经关联物”的结构具有同构性,当然前提是我们需要接受体验者的内省报告是相对可靠的(即便不是不可错的)。在这里需要明确一点,那就是在大脑中找到了与现象意识结构上同构的神经动力活动结构,并不意味着我们已经明确了解释现象意识的神经机制,只是说如果现象意识具有相关的神经解释机制,我们会自然地预期现象意识的结构特征关联着某些神经生物层面上同构的结构特征。我们可以再通过形式化的方式更为细致地说明这一点:

前提(1)如果“神经物理主义”为真,那么在神经结构或动力学的层面上必定存在着现象意识的特定解释机制(“神经物理主义”逻辑上的基本蕴含);

前提(2)如果神经结构或动力学层面存在着现象意识的解释机制,我们将预期现象意识“神经关联物”的结构特征同构于现象意识的结构特征;

前提(3)神经科学的经验研究提供了大量“神经—现象同构性”的证据;

前提(4)其他替代性形而上学方案对“神经—现象同构性”现象提供的解释劣于“神经物理主义”提供的解释;

① Diego Cosmolli et al., “Waves of Consciousness: Ongoing Cortical Patterns During Binocular Rivalry”, *Neuroimage*, 2004, 23(1): 128-140.

② 例如神经科学对梦境、致幻剂状态的研究表明,神经网络物理不确定性(熵)的增高伴随着意识体验非恒常性、多变性的增强,这是神经—心理同构性的进一步证据。Robin Carhart-Harris, et al., “The entropic brain: a theory of conscious states informed by neuroimaging research with psychedelic drugs”, *Frontiers in Human Neuroscience*, 2014, 8: pp. 1-2.

结论：“神经物理主义”为“神经—现象同构性”现象提供了最佳的形而上学解释。^①

四、罗素主义泛心论的解释困境

本节将为上述论证中的前提(4)进行辩护。基本思路是,现象结构与神经动力结构的同构尽管能够在罗素主义泛心论框架下获得解释,但是这会带来明显的代价:因为这将使得微观物理结构不再具有解释相关性,即使引入额外的假设也只会增加理论的复杂度,这进一步削减了泛心论的解释优势。

根据罗素主义泛心论的基本主张,宏观现象意识的本质就是不计其数的微观现象性质 $P(1) \cdots P(n)$ 处在特定的组合关系 R 当中。换言之,在构成性泛心论看来,现象意识的本质既涉及宏观层面的神经动力结构,又涉及微观层面的物理结构,因此泛心论承诺了微观物理结构与宏观神经动力结构共同的解释相关性。但是对于解释现象结构与神经动力结构的同构而言,只有宏观神经动力结构发挥了解释效用,微观物理结构并未带来解释的收益。这是因为,微观物理结构并未在现象场域中有任何呈现。从这个意义上讲,由于泛心论承诺了微观现象属性及其组织关系,但是它们并未在宏观层面承担解释作用,因此相比起“神经物理主义”,泛心论在解释中显然处于劣势。正如洛克伍德(Michael Lockwood)所言,现象意识在结构上完全不具有微观物理结构的细密“颗粒性”:

相比任何可能的生理关联,内省觉知的现象对象在结构的精细性上都相差甚远。举例来说,对于某个现象上无瑕的听觉经验,比如小提琴音符的听觉体验,它的生理基质大概是高度结构化地(更不用说杂乱地)发生在听觉皮层数十亿计神经元中电位变化的串联,这一系列变化都是由穿过细胞膜的钠离子和钾离子的转移以及突触“化学汤”中递质物质分子的转移所调节的。这些微观结构上的不连续性和非同质性何以被掩盖,使得它们能够以特定的方式产生出完美无瑕的听觉现象性……我们如何在哲学上理解这种现象的“粗粒性”(coarse-graining)?^②

由此看来,倘若微观物理结构与宏观神经动力结构都具有解释相关性的话,那么现象意识应当同时呈现出宏观层面的神经动力结构与微观层面的物理结构。但这并不符合事实,在这里,“神经物理主义”的解释优越性明确凸显。这是因为,根据“神经物理主义”,现象意识不过是宏观层面的神经动力结构,神经—现象结构的同构是“神经物理主义”自然的经验预言。相反,罗素主义泛心论还额外地承诺了微观物理结构对意识的构成,然而这种额外的形而上学设定并未在宏观层面发挥任何解释效用,更像是可有可无的“摆设”。从这个意义上讲,“神经物理主义”明显更有精简性的优势。

当然,意识的“粗粒性”并不构成对泛心论的有力反驳。泛心论者可以坚称,微观物理结构的确参与了意识的构成,但是它们未必一定要在宏观的现象场域中呈现。对此,泛心论者可以主张,微观物理结构的确以某种方式构成了意识的结构,但其无法被内省所通达^③。从这个意义上讲,内省活动也就并未揭示出意识结构的全部本质,微观物理结构是意识的“隐结构”。从逻辑上讲,承诺意识的“隐结构”能够

① 即使不考虑“逻辑上必然”或“形而上学必然”的蕴含这种过度严格的标准,我们仍然有理由认为探寻结构上的同构性是在经验层面为意识或广义上的心理现象寻求解释机制的有用启发法。Alfredo Vernazzani, “Psychoneural Isomorphism: From Metaphysics to Robustness”, in F. Galzavari and M. Viola ed., *Neural Mechanism: New Challenges in the Philosophy of Neuroscience*, Switzerland: Springer, 2020, pp. 283–310.

② Michael Lockwood, “The Grain Problem”, in T. Alter and Y. Nagasawa ed., *Consciousness in the Physical World*, Oxford: Oxford University Press, 2015, p. 147.

③ 比如布洛克(Ned Block)就特别讨论了所谓“现象溢出”(phenomenal overflow)问题。Ned Block, “Consciousness, Accessibility, and the Mesh Between Psychology and Neuroscience”, *Behavioral and Brain Sciences*, 2007, 30: 481–548.

赋予微观物理结构解释上的效用,但这绝不是好的解释。

一方面,这个对策进一步预设了难以在经验层面验证的“事体”和过程(即“隐结构”以及“隐结构”构成或产生宏观神经动力结构的过程),因此进一步增加了理论预设的复杂性,但并未带来经验解释力的实质性增强。从根本上讲,这个对策仅仅是理论对经验现象的事后容纳而绝非预言。这是因为,如果关于某个现象的陈述是特定理论的经验预言,那么这个理论要么演绎地蕴涵关于该现象的陈述(比如万有引力定律蕴涵了关于苹果落地现象的陈述),要么该理论为真(相对于该理论为假)使得该现象的出现更加可信(probable)。很显然,构成性泛心论并不在逻辑上蕴涵神经—现象同构性以及意识“粗粒性”的同时出现;同时,假定构成性泛心论为真(相对于其为假),神经—现象同构性和意识“粗粒性”的同时出现也并不具有更高的可信度。从这个意义上讲,这种微观、宏观结构共同构成意识的主张无异于引入上帝的干预或者“前定和谐”去解决心身二元论在神经—心理同构性现象面前的困境。对于这种经过修订的构成性泛心论来说,微观结构并未发挥任何经验解释的效力,反而明显增加了理论的复杂性。

另一方面,接受上述对策意味着放弃泛心论者普遍接受的“现象透明性”(phenomenal transparency)原则,也就是说内省活动揭示了现象意识的全部本质^①。根据“现象透明性”原则,在内省揭示的主观感受中,不存在实在和现象的区别,内省呈现出的直观现象特征就是主观经验的本质。由于构成性泛心论断定微观现象性质及其组织结构构成了现象意识的本质,基于“现象透明性”原则,内省必定要呈现出这种细密、复杂的微观结构。倘若承认微观结构是内省不可通达的“隐结构”,这无疑意味着放弃“现象透明性”原则。“现象透明性”论题尽管不是泛心论逻辑上必然的承诺,但是如果抛弃这个论断,泛心论将失去一个重要甚至基础性的理论动机。那就是,如果内省无法完全揭示出意识的本质特征,甚至内省呈现出的仅仅是某种认知显象,那么诉诸基础性的微观现象性质去解释现象意识还有何意义呢?泛心论者似乎很难接受这一点。

反驳者或许会说,构成性泛心论的上述对策尽管牺牲了理论的精简性,但是引入微观现象性质和结构仍然是必要的,因为“神经物理主义”始终面临着神经结构与感受质之间的解释鸿沟。的确,纯粹的神经结构和动力学活动似乎难以解释似乎是质的以及第一人称的主观经验。但是,在这个问题上构成性泛心论同样面临着相应的解释鸿沟。因为,主观经验特定质的特征以及第一人称视角性似乎是不可还原的。比如,疼痛体验的质的特征似乎无法还原为其他体验(比如颜色体验)质的特征。因此从内省的角度讲,特定体验的质的特征是绝对意义上初始的(primitive),这意味着这种质的特征不可能由其他质的特征构成。对于这一点,有论者这样引用并补充了休谟的论述:

蓝与绿是不同的简单观念,但是相比起蓝与猩红它们之间更为相似;然而它们完美的简单性排除了一切分离与区分的可能性。对于特定的声音、味道和气味来说情况同样如此。在总体的显象与对照之下它们都具有无限的相似性,但是没有任何一种是相同的一般情形……它们在自身简单性(simplicity)的意义上彼此相似。但是从它们本质(其排除了一切的组成)的角度来讲,它们彼此相似的那种情形与其余的情形既无法区分也无法分离。^②

休谟的上述论断试图说明,即便现象质A与现象质B具有相似性,但是A却不可能还原为B或者由B构成,主观体验的每一种质的特征都是简单且原始的。由于罗素主义的构成性泛心论承诺了奠基基础物理倾向的现象质,这种现象质必定是十分基础的,与日常经验的感受质相差甚远。因此,从内省的角度看,即便彼此具有较高相似度的不同颜色感受质都不可能相互还原或构成,更何况迥异于宏观感受

① Philip Goff, *Consciousness and Fundamental Reality*, pp. 107–108.

② Lok-Chi Chan, “Russellian physicalism and its dilemma”, *Philosophical Studies*, 2020, 178: 2053–2054.

质的微观现象质? 因此,构成性泛心论同样面临微观现象质与宏观现象质之间的解释鸿沟。在此情形下,“神经物理主义”的理论精简性无疑有更大的优势。当然,泛心论者为了弥合微观与宏观的解释鸿沟可以引入突现论的主张,但这样一来必定要承诺某种突现作用力的存在,这无疑违背了物理因果闭合性原则。泛心论者提出的各种对策无助于增强经验解释力,还损失了理论的精简性。从解释力和理论精简性的角度讲,“神经物理主义”具有显著优势。

总而言之,本文剖析了物理主义因果闭合性论证的局限性,指出其无法在逻辑上排除构成性的罗素主义泛心论。而基于神经—心理同构性现象的最佳解释推理论证则为物理主义提供了更为有力的辩护,这无疑是物理主义者的福音。但这也带来了一些尚待解决的问题。例如,在某些特殊的意识状态下(比如致幻剂诱发的神秘体验),意识结构与神经结构出现了错配^①,物理主义者如何解释这些反例? 另外,本文的分析聚焦于原子论的构成性泛心论,相对于构成主义的宇宙论泛心论^②(cosmopsychism,认为个体意识由超宏观即宇宙层面的现象性质构成),物理主义是否仍然具有解释优势? 这些问题都有赖于进一步的理论与经验研究来回答。

Causal Closure of the Physical, Russellian Panpsychism and Psycho-Neural Isomorphism

Wang Shuaixiang

Abstract: Physicalist account of the mind is widely embraced because of causal closure of the physical world. Physicalists, resorting to empirical researches from physics and neurophysiology, contend that any causal interactions in the physical world are causal interactions characterized by fundamental forces. That such a characterization of physical causation cannot draw a logical distinction between physicalism and Russellian panpsychism implies that the argument from causal closure of the physical does not decisively lend support to physicalism. This article aims to respond to such a challenge in light of empirical researches from neuroscience. Abundant neuroscientific evidences illustrate that the phenomenal structures of conscious experiences are isomorphic to the physical structures of corresponding neural correlates. In virtue of the fact that physicalism furnishes a better explanation for the phenomenon of psycho-neuro isomorphism than that provided by Russellian panpsychism, there is good reason to regard physicalism as a plausible metaphysics for the mind-body problem.

Keywords: causal closure of the physical; Russellian panpsychism; psycho-neural Isomorphism; inference to the best explanation

【责任编辑:全广秀;责任校对:全广秀,张慕华】

^① Bernardo Kastrup, “What Neuroimaging of the Psychedelic State Tells Us about the Mind-Body Problem”, *Journal of Cognition and Neuroethics*, 2016, 4 (2): 1-9.

^② I. Shani, “Cosmopsychism: A holistic approach to the metaphysics of experience”, *Philosophical Papers*, 2015, 44: 389-437.