

脑、心、世界：预测编码、新康德主义与超越论的观念论

[丹] D. 扎哈维/文 李恒威 康文煌/译

提 要：最近一些神经科学家和哲学家认为预测编码进路支持了一种激进的神经表征主义，按其所言，我们有意识体验的内容是一种神经建构，一种由脑生成的模拟。这种说法与 19 世纪中期德国新康德主义者发展的思想极为相似。而一些新康德主义者最终对他们所处的表征主义框架的说服力和内在一致性产生怀疑。我在本文中首先论证当代预测编码的支持者需要认真对待其中的一些关注点。之后我会转向现象学。我们会看到，胡塞尔支持超越论的观念论的动机部分是由于他拒斥表征主义和现象主义，以及他试图捍卫体验世界的客观性。这让我们面临一个有趣的问题，即哪种立场最能顺应我们对实在论的自然倾向：当代神经表征主义还是胡塞尔的超越论的观念论？

关键词：预测加工；新康德主义；超越论的观念论

中图分类号：B534

文献标识码：A

当代认知科学和心智哲学能与与哲学史的更密切的交互中获益吗？对这个问题的回答应该是肯定的，而且理由很多。对心智 - 世界关系的深刻反思和分析不是新东西，忽视传统资源会适得其反。这样做可能会让你错过一些重要见解，而这些见解可能要在几十年或几百年后才被重新发现。在下文中，为了例证这点，我会论证目前流行于认知神经科学中的预测编码^①框架与 19 世纪中期德国新康德主义者发展的思想之间存在一些有趣的相似点。然而，这种比较的要点将不仅是历史方面的。我们会看到，一些新康德主义者最终提出某些问题，并对他们所处的表征主义框架的说服力和内在一致性怀有疑虑。我会论证，当代预测编码的支持者需要认真对待其中的一些关注点。之后我会转向现象学。我们会看到，E. 胡塞尔 (E. Husserl) 支持超越论的观念论的动机部分是由于他拒斥表征主义和现象主义以及他试图捍卫体验世界的客观性。这让我们面临一个有趣的问题。哪种立场最能容纳我们对实在论的自然倾向：当代神经表征主义还是胡塞尔的超越论的观念论？

^① 预测编码是信息传输中一种高效的编码策略，在层级中传输预测与输入的差异信息，而不是直接地传输全部输入，以此降低传输量。其广泛使用于电信领域，也被发现是神经系统的编码方式。之后发展的脑与心智的预测加工理论综合了预测编码、统计学中的变分贝叶斯推断、自由能原理等。“预测编码进路”与“预测加工”有时被不严格地交换使用。——译者注

一、预测编码

最近，包括 C. 弗里斯 (C. Frith)、T. 梅青格尔 (T. Metzinger)、J. 霍威 (J. Hohwy) 等在内的一批神经科学家和哲学家对不同形式的预测编码进路 (*predictive coding approach*) [有时也称为预测加工理论 (*predictive processing theory*) 或预测误差最小化理论 (*prediction error minimization theory*)] 进行了辩护，并认为它支持了一种激进的“神经表征主义” (*neuro-representationalism*)；之所以激进，是因为它不是简单地主张我们以神经表征为中介来通达外部世界，而是更明确地声称体验世界本就是一种表征建构 (*construct*)。其主张的核心是，脑并不处理接受到的所有信息，而是将其资源集中在意外输入上。不过，为了使处理惊异 (*surprisal*) 的代价最小化，脑总是尽力预期感官将接收到什么信号。为了尽可能高效地做到这一点，脑建构了那些输入的可能原因的内部模型。这些模型让脑能更好地预测可能输入，然后将这些预测不断地与实际传入的感官输入进行比较。当出现预测误差，即如果预测与实际输入之间出现了大的差异，模型会得到修正和完善^①。根据这种观点，脑被看作是假设检验器官。然而，脑能操作的唯一数据就是被刺激的感官的内部结果。内部结果的外部原因仍然是未知。可以说，脑能通达的所有证据因此都只是内部可用的证据。如霍威所说，我们永远不可能爬出脑，将我们的表征和预测与外部事态进行直接的比较^②。我们有意识体验的内容因此必定要被当作神经建构，一种脑产生的模拟。弗里斯说道“我的感知不是关于世界的，而是关于我脑中的世界模型的。”^③我们所见、所闻、所触、所嗅以及其他所感知之物都局限在脑中，但却被投射到外界并得以外化，以至于我们在正常生活里并没有将它们认作是一种建构，而是错误地将它们当成实在本身^④。颜色常常充当这种相对无辜的起点：

夕阳的杏粉色不是傍晚天空的属性；它是你的脑所创造的傍晚天空的内部模型的属性。傍晚天空是无色的。世界中根本不存在有色物体。[……] 你的眼前只是一片电磁辐射的海洋，各种不同波长的疯狂混合物。^⑤

但这种适用于颜色的说法也适用于其它熟悉的事物。直接的知觉对象实际上是心智建构。视觉上显现的玫瑰、被触摸到的冰块、被听到的旋律等都是脑产生的表征，都内在于和包含于脑中。

鉴于我们从未直接接触外部事态——毕竟外部事态隐匿于表征之幕后面——我们应该拒绝关于心智与世界之间存在无缝紧密耦合的所有主张。霍威说到内部与外部之间严格而绝对的划分，以及将脑与其边界外的一切事物隔离开来的“证据界线”^⑥。这甚至包括我们自己的身体，身体对我们来说也是隐匿的，就像“距离遥远的感官输入的原因，诸如那些退去

^{①③} C. Frith, *Making Up the Mind: How the Brain Creates Our Mental Worlds*, Blackwell, 2007, pp. 126—127; p. 132.

^{②⑥} J. Hohwy, “The Self-evidencing Brain”, *Noûs*, 50 (2), 2016, p. 265.

^{④⑤} T. Metzinger, *The Ego Tunnel*, Basic Books, 2009, pp. 6—7; p. 20.

的星系”^①。

从认识论上说，二元论在此占主导地位。但是，正如霍威承认的那样，尽管他们提供的理论解释“蕴含了怀疑论”^②，尽管我们（脑中）的表征过滤器“阻止我们看到世界的本来面目”^③，但这一点不该被视为重大隐患。梅青格尔向我们保证说“外部世界确实存在，知识和行动在因果上将我们与它联系起来”^④。以工具主义思路，弗里斯摒弃我们永远无法知道自己的内部模型是否真的与外部世界匹配的忧虑。他认为关键的是预测和控制是可能的，即模型是有效的。至于模型是否真实地描述了实在则无关紧要^⑤。

二、新康德主义

有趣的是，最近一些文章经常引用 H. v. 赫尔姆霍茨 (H. v. Helmholtz) 的著作作为历史参照^⑥。这主要是因为人们认为赫尔姆霍茨是如下观点的一个早期捍卫者，即感知的任务是推理出感觉的原因，而感知在某种程度上等价于一种假设检验^⑦。此外，赫尔姆霍茨还是个新康德主义者，他主张自然科学已经证明了康德的一些主要见解^⑧。赫尔姆霍茨在生理学家 J. 穆勒 (J. Müller) 的著作中发现了这一主张的重要灵感来源，穆勒宣称外部原因的属性不会以保真且精确的方式通过神经传导到意识。实际上，有如此多的中间步骤和转换会发生在外部原因与被体验到的结果之间的路途中，以至于可以安全地排除这两者之间任何的相似或类似之处。当我们被刀割伤，我们感到的不是外部之物，而是内部疼痛。同样，当我们看见某物时，我们的感知不仅取决于外部原因，而且更重要的是取决于我们的生理构造。穆勒写到：

通过感官中介而被感觉中枢所感知的事物实际上只是我们神经状况的变化或属性；但是想象力和理性已准备好要把外部影响所产生的神经状态的变化解释为外部物体本身的属性。^⑨

赫尔姆霍茨认可这个推理，并同样主张由于外部事物的信息在经过神经系统时已经面目全非到识别不出了，所以严格说来，我们最终感知的是内部结果，而非外部原因：

按当前的理解，[科学] 考查的结果是，感官确实向我们提供了在它们上面产生的外部作用的信息，但却把那些作用以完全不同的样式传递到我们的意识，以至于感官知

^{①②⑦} J. Hohwy, "The Self-evidencing Brain", *Noûs*, 50 (2), 2016, p. 275; p. 265; p. 77.

^{③④} T. Metzinger, *The Ego Tunnel*, Basic Books, 2009, p. 9; p. 23.

^⑤ C. Frith, *Making Up the Mind: How the Brain Creates Our Mental Worlds*, Blackwell, 2007, p. 136.

^⑥ Cf. C. Frith, *Making Up the Mind: How the Brain Creates Our Mental Worlds*, Blackwell, 2007, p. 41, p. 102; A. Clark, "Whatever Next? Predictive Brains, Situated Agents, and the Future of Cognitive Science", *Behavioral and Brain Sciences*, 36 (3), 2013, pp. 181—204; J. Hohwy, *The Predictive Mind*, Oxford University Press, 2013, p. 5; J. A. Hobson & K. J. Friston, "Consciousness, Dreams, and Inference: The Cartesian Theatre Revisited", *Journal of Consciousness Studies*, 21 (1—2), 6—32, 2014, p. 8.

^⑧ Cf. L. D. Kock, "Hermann von Helmholtz's Empirico-Transcendentalism Reconsidered: Construction and Constitution in Helmholtz's Psychology of the Object", *Science in Context*, 27 (4), 2014, pp. 709—744; pp. 20—32.

^⑨ J. Müller, *Elements of Physiology II*, W. Baly trans., Taylor and Walton, 1842, p. 1059.

觉的特征与其说取决于被感知对象的属性，不如说取决于我们用来接收信息的器官的属性。^①

我认为感觉只是物体作用的一种迹象。迹象的本质仅仅属于这种属性，即相同的物体总是给出相同的迹象。此外迹象与产生迹象的物体之间在类型上没有任何相似性，就像言语单词与我们借此指称的物体之间很少有相似性一样。^②

最后，赫尔姆霍茨把穆勒的理论及其给出的证据当作对康德在《纯粹理性批判》中提出的如下基本主张的科学证实，即在某种程度上“我们认知的不是物自身，而只是显象 (appearance)”^③，他还认为当代科学基于生理学证据正在达到康德通过先验^④考量 (a priori considerations) 所达到的同样见解。我们的知识所关涉的是在我们心中表征的那种实在，而不是作为如其所是的独立于心智的实在，后者仍是不可知的。尽管呈现给我们的永远只是内部感觉，然而我们能推断出必定存在外部世界，因为必定有“神经激发的原因；毕竟没有原因就不会有结果”^⑤。通过（无意识的）最佳解释推理过程，内部感觉由此指涉到作为其假定原因的外部物体。

让我们暂时回到康德。最近有人宣称康德在其涉及感知和认知的著作中预示了预测编码理论的一些核心方面，而且认为，预测编码理论由于坚持感知的主动性和假设驱动的特征，甚至可被看作是康德的哥白尼革命^⑥的升级版。^⑦ 不过，康德理论早就受到批评。早期对《纯粹理性批判》的一个极有影响的反对意见是 F. H. 雅可比 (F. H. Jacobi) 1787 年提出的。雅可比认为康德对物自身的诉诸和涉及违反了自己的批判体系。按照康德观点，我们不可能具有物自身的知识，但我们却假定有理由地断定物自身存在。尽管对康德来说因果性是一个知性范畴，只适用于现象界，但康德仍然把它赋予物自身；尽管物自身既不存在于空间中也不存在于时间

①② H. v. Helmholtz, *Science and Culture: Popular and Philosophical Lectures*, Chicago University Press, 1995 [1853/1892], p. 13; p. 408.

③ I. Kant, *Critique of Pure Reason*, P. Guyer and A. W. Wood trans., Cambridge University Press, 1998 [1787], B p. xxvi.

④ a priori 也译为“先天”“验前”，此处译为“先验”。与之相关的 transcendental 译为“超越论的”。——译者注

⑤ H. V. Helmholtz, *Das Sehen des Menschen*, Leopold Voss, 1855, p. 41.

⑥ L. R. Swanson, “The Predictive Processing Paradigm Has Roots in Kant”, *Frontiers in Systems Neuroscience*, 10 (79), 1—13, 2016, pp. 1, 4. doi: 10.3389/fnsys.2016.00079.

⑦ 斯万森 (Swanson) 进而主张预测编码理论能被“看作是康德超越论心理学历史进展中的一个主要进步” [Swanson, L. R., “The Predictive Processing Paradigm has Roots in Kant”, *Frontiers in Systems Neuroscience*, 2016, 10 (79), pp. 1—13. doi: 10.3389/fnsys.2016.00079], 尽管他也承认前者的演化的、计算的和神经科学的进路“以康德无法想象的方式超越了康德的见解” [Swanson, L. R., “The Predictive Processing Paradigm has Roots in Kant”, *Frontiers in Systems Neuroscience*, 2016, 10 (79), 1—13. doi: 10.3389/fnsys.2016.00079]。不过，我们可能想，预测编码的自然主义是否兼容于康德的超越论框架。对自然化康德的可能性的批判，参见艾利森的文章 [Allison, H. E., “On Naturalizing Kant’s Transcendental Psychology”, *Dialectica*, 1995, 49 (2—4), pp. 335—356]。对于自然主义和超越论哲学能否被调和，参见笔者相关文章 [Zahavi, D., “Phenomenology and the Project of Naturalization”, *Phenomenology and the Cognitive Sciences*, 2004, 3 (4), pp. 331—347; Zahavi, D., “Naturalized Phenomenology: A Desideratum or a Category Mistake?” *Royal Institute of Philosophy Supplement*, 2013, 72, pp. 23—42]。

中（两者都是纯粹的直观形式），但却被认为能影响我们并引发我们的表征。雅可比总结到：

不管与康德哲学精神如何背道而驰，大家倾向于说物体在感官上产生印象，并以此方式产生表征，如果没有这种预设，很难看出康德哲学如何为自己找到一个入门，并对其教义做任何形式的表述。[……]若没有这种预设，我找不到进入康德体系的路，但这种预设让康德体系自相矛盾，以至于无法让我信服。^①

费希特（J. G. Fichte）也同意雅可比的批判，他在1796年初的一篇文章中赞同雅各比对康德的抱怨^②。费希特认为物自身的概念对康德体系有害，其他德国观念论者如黑格尔也这样认为，黑格尔进而认为物自身与为我们之物（the thing for us）之间的区分是我们制造出来的区分，一种为我们的区分。黑格尔在《逻辑学·哲学全书·第一部分》第44节写到：

当一个人抽掉所有向意识显现的东西，抽掉所有感性规定和思想规定，剩下的就是物自身所表达的东西。显而易见，余留下的是完全的抽象物，某种完全的空，某种只作为“超出”而被规定的东西[……]。同样简单的反思是，这种余留物本身仅仅是思想的产物，确切地说，是已经发展到纯粹抽象阶段的思想的产物。^③

有趣的是，我们在生理学的新康德主义者中也发现涉及物自身地位的相似发展。这点在F. A. 兰格（Friedrich Albert Lange）的不朽巨著《物质主义的历史及其当前重要性的批判》（*Geschichte des Materialismus und Kritik seiner Bedeutung in der Gegenwart*）中尤为明显。兰格最初赞同赫尔姆霍茨的观点，即对感官生理学的科学研究部分证实了康德的基本观点，并可以被解释为对康德主义的某种形式的更正和改进^④。感官只向我们提供事物在内部引发的作用，而不是让我们通达外部的物自身。相应地，我们所体验的世界必须被看作是我们构造的产物。但兰格指出，如果我们这样思考的话，就会有一些经常被忽视的含义。第一，这也显然适用于我们的身体器官。我们的身体、感官、神经乃至脑本身，都是经验世界中的成分，因此也不过是不可知之物的不可靠的形象和迹象。兰格写到：

我们所相信的用来观看的眼睛本身也是我们观念的产物；而当我们发现视觉图像是眼睛结构的产物时，我们永远不要忘记，眼睛连同其生理结构、视网膜神经连同脑以及所有那些我们已经发现作为思想原因的生理结构都仅仅是观念，这些观念事实上形成了一个自洽的世界，一个指向着自身之外的观念世界。^⑤

① F. H. Jacobi, "On Transcendental Idealism", in B. Sassen ed., *Kant's Early Critics: The Empiricist Critique of The Theoretical Philosophy*, Cambridge University Press, 2000 [1787], pp. 169—175.

② J. G. Fichte, *Early Philosophical Writings*, D. Breazeale trans., Cornell University Press, p. 325.

③ G. W. F. Hegel, *Encyclopedia of The Philosophical Sciences in Basic Outline, Part I: Science of Logic*, K. Brinkmann and D. O. Dahlstrom trans., Cambridge University Press, 2010 [1830], p. 89.

④⑤ F. A. Lange, *The History of Materialism and Criticism of Its Present Importance*, 3 vols., vol. 3, E. C. Thomas trans., Kegan Paul, 1925 [1865], pp. 202—203; p. 224.

然而，最终不只是我们的感觉，我们那些依赖于生理构造和组织的概念也是如此。按照兰格的观点，产生感觉的内部机制也负责产生物质概念和外部实在概念^①。就像鱼儿在池塘中游弋且无法超越其界限，我们生活在概念和表征的领域中。即便谈到物自身时，我们也没有超出自己的领域^②。自然科学已经取得了巨大进步，没有理由怀疑它们的成就，但人们最终还是无法摆脱这些根本的认识论问题，而兰格认为，此时人们会认识到物质主义者的实在论是错误的。对感官的生理研究貌似是在提供对知识获取过程的一种彻底的物质主义的解释，但实际上却瓦解了我们对物质的、自为存在的物体的信念，这就是为何物质主义经足够彻底的思考后暴露出自身是某种形式的观念主义^③。

兰格最后对物自身模棱两可。有时他似乎认为，因为显象存在，并且我们需要解释显象，所以我们能根据最佳解释推理去设定物自身为显象的外部原因。有时他又说原因概念的有效运用只限定于经验界限内，我们因此不能超出经验范围去使用因果解释。有时他又认为物自身概念虽然是可理解的，但很可能是个空概念，且我们无法知道它是否真有所指^④。

虽然兰格一直含糊地提及物自身，被穆勒的学生教导过的 E. 马赫 (E. Mach) 在几年后迈出了最后一步。马赫在 1880 年代出版的著作中认为感官的生理学研究为知识理论提供基础，但他认为可以安全地取消物自身概念。经验对象不是超出经验的世界的（失真）表征，所有存在物不过是感觉的复合物^⑤，这观点后来被解释为要么是中立一元论的一种，要么是强健版的现象主义。

三、脑的独特地位

在做了这些回顾后，让我们回到当代神经表征主义，我们可以合理地给它贴上“新新康德主义” (neo-neo-Kantians) 的标签^⑥。要吸取的一个教训是，一旦踏上表征主义之旅，就很难维持外部世界的存在了。所以完全不清楚梅青格尔何以自信地宣称体验世界是脑产生的错觉，而物理学所描述的世界，即电磁波的世界，是真实不虚的世界。科学家如何超越自己的内部世界模拟呢？科学家是如何达成这个认识论成就的呢？我们如何知道科学理论（作为体验世界的解释）真的把握了外部实在？科学理论为什么不可以只是一种精致的认知外推，像其他所有东西一样仍然是脑产生的并且内在于我们的构造呢？

在此，从演化的角度来考虑也许很有诱惑力。人类栖息于一个预先存在的自然世界中。为了生存，我们必须有能力获得关于实在世界的真实知识，因为只有如此我们方能采取有利于提升生存的行动。换言之，除非能让我们的内在表征匹配并追踪外部实在，否则我们的认知装置就不可能演化成现在的样子，不可能经受住选择压力。但这思路很难令人信服。我们不能诉诸演化理论来建立表征的可信度，因为前者显然要预设后者。

^{①②③} F. A. Lange, *The History of Materialism and Criticism of Its Present Importance*, 3 vols., vol. 3, E. C. Thomas trans., Kegan Paul, 1925 [1865], p. 204; p. 226; p. 223.

^④ S. Edgar, "The Limits of Experience and Explanation: F. A. Lange and Ernst Mach on Things in Themselves", *British Journal for the History of Philosophy*, 21 (1), 2013, pp. 100—121.

^⑤ E. Mach, "The Economical Nature of Physical Inquiry", in *Popular Scientific Lectures*, Th. J. McCormack trans., Open Court, 1895 [1882], pp. 200—201.

^⑥ M. L. Anderson & T. Chemero, "The Problem with Brain GUTs: Conflation of Different Senses of 'Prediction' Threatens Metaphysical Disaster", *Behavioral and Brain Sciences*, 36, 2013, pp. 204—205.

另一个策略指向主体间性的作用。科学不是由个人发展的，而是集体合作努力的结果。科学理论是科学界随时间推移建构起来的，在此意义上确实可以说它们超越了个体对世界的模拟。然而，这个论点的问题也是显而易见的。鉴于神经中心主义的非具身（disembodied）框架，很难认为主体间性是理所当然的。其他主体的存在与外部事物的存在一样可疑。

对于这一点，有人或许会说，我们应该直面我们的认知是有限的，关于外部世界的绝对可靠的知识是不可企及的，我们自始至终都在使用溯因推理，科学发现仍是我们的最佳选择。然而，这种争论方式漏掉了重点。当前被质疑的不是我们是否是不可靠的认知者，或科学是否为有价值的事业，而是正在讨论的这个立场何以能这么自信地信奉科学实在论。

然而，最终说来这还不是主要的担忧。正如我们在兰格那学到的，还有个更棘手的问题可能破坏整个格局。我们最初为何要考虑知觉的对象可能是内部产生的建构，而不是真实的时空对象呢？因为这是我们对脑的神经科学研究所认为的。但如果要严肃对待这个理论，它就必须是一致的。不能认为只有视觉上显现的锤子、橘子和护照才是脑产生的虚拟现实的一部分，这必须同样适用于在视觉上显现的脑，不论是在做脑外科手术时“直接”感知的那种脑，还是在进行脑扫描时看到的那种脑。毕竟，我们关于脑的经验知识（以及神经生理学）必定是由知觉提供的，如果我们不信任诸感官的判决（deliverance），那么这必定同样适用于诸感官对脑的判决。换一种稍微不同的说法，很难理解人们基于神经科学的发现何以提出对感知体验的普遍怀疑，因为神经科学的发现至少在某种程度上是以感知的有效性为前提的。简言之，主要挑战不在于我们如何能在认识上走出脑，而是我们一开始如何能够进入脑。我们究竟如何知道确实存在一个脑？无论如何，为了领会最初貌似合理性，上述神经表征主义者的说明必定是不完善的。神经表征主义要求我们放弃朴素的实在论，放弃对日常体验对象的客观存在的信念，但却敷衍了事地相信脑的客观存在。斯拉比（Slaby）和海林格（Heilinger）正确地指出，既然其整个理论是围绕着脑的工作方式建构起来的，那么这个模型必须预设脑这个世间物体免受怀疑论的忧虑，并预设我们确实如其所是地观察和描述了脑。但是，

如果作为科学发现的脑在超越的意义上是“实在的”，那么难于让人信服的是：我们就止步于此，并声称在我们感知的所有事物中，只有脑这个唯一对象是“真正实在的”，如其本来的样子被感知到，而不仅仅是表征。^①

根据讨论中的说法，体验世界是脑产生的表征建构。但这说法面临一个明显的两难。它或者把脑也当作体验世界的一部分，即把脑也看作表征建构，而如果这样，这种解释就是循环的并且在解释上是空洞。毕竟，当解释表征是如何产生时，这个解释又必须诉诸另一个表征。另一个选项是认为脑不是表征建构。但这个理论如何有资格持有这种观点，不仅这一点是不清楚的，而且脑为何能是唯一例外这一点也是不清楚的。

如果我们在认知上确实局限于最近神经科学所提议的道路，那么很难理解讨论中的这个立场如何能得到融贯一致的表述，更别说到得到辩护了。这是一种非常经典的反驳，熟悉康德传统的人立马就会想到这个反驳。也许神经表征主义者已经有现成的解答。然而，据我所

^① J. Slaby & J.-C. Heilinger, "Lost in Phenospace: Questioning the Claims of Popular Neurophilosophy", *Metodo*, 1 (2), 2013, pp. 83—100.

知，他们还未直面这个挑战。

在神经科学模型与其所遭遇的理论解释之间做出区分是非常重要的。K. 弗里斯顿 (K. Friston) 的理论工作^①构成了最近预测编码讨论的重要灵感来源，且被誉为“将在未来主导心智科学和脑科学的理论”^②。但是尽管霍威认为这个理论蕴含了认知系统与环境之间存在严格边界，并因此排除了一些涉及心智的延展 (extended)、具身 (embodied) 和生成 (enactive) 特征的假说^③，A. 克拉克 (A. Clark) 却认为预测加工兼容并支持情境、具身和分布式 (distributed) 认知进路，且预测误差最小化理论并没有在心智与世界之间引入让人心烦的边界^④。克拉克说道，与其说是我们所感知到的是内在表征，不如说是脑的复杂的亚人格 (sub-personal) 加工流允许一种紧密的心智 - 世界关联，并且让我们能在知觉上对世界本身开放^⑤。不过，另一些人批评克拉克对表征主义的批判还不够彻底，有人最近更是直接地试图把弗里斯顿的核心观点与生成主义 (enactivism) 结合起来^⑥。我不在这里对弗里斯顿工作的正确解释是什么表明我的立场，我对预测编码范式是否必然承诺神经表征主义也不表明立场。但有一点很明显，上面所述的担忧主要是针对表征主义解释，而对非表征主义的替代解释则不会造成太多困扰。^⑦

四、现象学

神经表征主义者向我们呈现的这种自然化知识论部分是由科学发现驱动的，因此显得有坚实基础。但它并没有很好地担保世界的实在性。我们多数人想成为实在论者。但哪种理论能最好地容纳我们的实在论直觉并不总是那么容易判定的。实际上，刚才提到的神经表征主义似乎更符合康德自己所定义的经验主义观念论 (empirical idealism)：“观念论认为我们直接体验的只是我们自己的存在，而我们只能推断外部事物的存在（这种从结果到原因的

① K. Friston, “The Free-energy Principle: A Unified Brain Theory?”, *Nature Reviews, Neuroscience*, 11 (2), 2010, pp. 127—138.

②③ J. Hohwy, “The Self-evidencing Brain”, *Noûs*, 50 (2), 2016, p. 259; p. 259.

④ A. Clark, “Whatever Next? Predictive Brains, Situated Agents, and the Future of Cognitive Science”, *Behavioral and Brain Sciences*, 36 (3), 2013, pp. 181—204; A. Clark, “Busting Out: Predictive Brains, Embodied Minds, and the Puzzle of the Evidentiary Veil”, *Noûs*, 2016. doi: 10.1111/nous.12140.

⑤ A. Clark, “Whatever Next? Predictive Brains, Situated Agents, and the Future of Cognitive Science”, *Behavioral and Brain Sciences*, 36 (3), 2013, pp. 181—204.

⑥ J. Bruineberg, J. Kiverstein & E. Rietveld, “The Anticipating Brain Is Not a Scientist: The Free-energy Principle From an Ecological-enactive Perspective”, *Synthese*, 2016, doi: 10.1007/s11229 - 016 - 1239 - 1. S. Gallagher & M. Allen, “Active Inference, Enactivism and the Hermeneutics of Social Cognition”, *Synthese*, 2016. doi: 10.1007/s11229 - 016 - 1269 - 8. 这些争论的原因不只是人们倾向于在相当不同的意义上使用“直接”、“间接”、“无中介”等词汇，还在于人们倾向于混淆对这些词汇的认识的、体验的和因果的理解。因果中介的存在并不必然导致知觉在体验上和/或认识论上是间接的。同样地，虽然我们对世界中时空对象的感知是由各种亚人格机制和无意识的认知过程所支持和因因而可能的，但这一事实并不一定导致我们不能如其所是地看到对象本身。确切的说，人们可以把认知加工看作是使我们能够最初体验这些对象 (Cf. Hopp, W. , *Perception and Knowledge: A Phenomenological Account*, Cambridge University Press, 2011, p. 163)。

⑦ 让我补充一点，神经表征主义显然也可以在预测编码框架之外被找到。一个突出的例子参见格鲁什的相关论述 (Cf. Grush, R. , “In Defense of Some ‘Cartesian’ Assumptions Concerning the Brain and its Operation”, *Biology and Philosophy*, 18, 2003, pp. 53—93)。

推断实际上是不确定的)”^①。

如普特南曾论证过的，正是传统上那些被指责为观念论的哲学家——即康德主义者、实用主义者和现象学家——而不是形而上学实在论者，才真正尊重和推崇我们对实在论的自然倾向^②。康德的著名观点认为，他自己的超越论的观念论 (*transcendental idealism*) 不仅兼容于经验主义的实在论 (*empirical realism*)，而且是后者的前提^③。胡塞尔捍卫同样的立场，不过胡塞尔对超越论的观念论的理解非常不同于康德，比如胡塞尔反对物自身。

我在此不会提供胡塞尔超越论的观念论的全面阐释^④，但值得指出的是，胡塞尔很熟悉生理学的新康德主义，其转向超越论的观念论的部分动机是因为反对表征主义和现象主义，并努力保卫体验世界的客观性^⑤。

胡塞尔认为，那个无论是在感知中、在日常关切中、还是在科学分析中向我们显现的世界是唯一实在的世界。宣称在这个世界外还有一个超越所有现象、所有体验和理论证据的幕后世界，并把这幕后世界当作真正的实在，这对胡塞尔来说是种空洞、悖谬的命题。

对这些观点的清晰阐述可以在胡塞尔的《观念 I》中找到，他在该书中讨论并反驳了那种认为感知上显现的东西只是某种主观的东西、是隐匿原因的虚幻描画或纯粹迹象的想法。这种表征主义未能尊重感知意识与符号和图像意识之间的范畴区别。我们可以思考某物，并由此意指它，甚至在其不在场时，但遇到感知意向性时，情况就不同了。感知意向性不是让我们面对事物的肖像或图像（除非是在看相片或图画时），而是让我们面对事物本身的亲身在场。当我们说某物以感知方式显现时，这不应当被理解为在感知上被给予的是其他事物的图像或迹象。相比之下，在符号和图像的意向性中，我们感知的是能意指或描画其它事物的图像或符号。在此，我们被指向的不是我们所感知的，而是通过它指向其他东西。符号和图像意向性因而以感知意向性为前提，而不能解释感知意向性。

正如胡塞尔在 1915 《现象学问题择要》 (*Ausgewählte phänomenologische Probleme*) 中的讲座课程指出的那样，说我知觉到的对象在我的意识之外，似乎没有比这更自然的了。当我的体验——不论是感知或其它类型的意向活动——把对象呈现给我，我必须询问这是如何可能发生的，而答案似乎简单明了：是通过某些表征中介。我所意识的对象在意识外，但我在自己的意识中找到了这些对象的表征（图像和迹象），而且正是这些内部之物让我能意识到

① I. Kant, *Notes and Fragments*, C. Bowman, P. Guyer & F. Rauscher trans., Cambridge University Press, 2005 [1785—1789], p. 294.

② H. Putnam, *The Many Faces of Realism*, LaSalle, Open Court, 1987, p. 12.

③ Cf. H. E. Allison, “From Transcendental Realism to Transcendental Idealism: The Nature and Significance of Kant’s ‘Transcendental Turn’”, S. Gardner & M. Grist eds., *The Transcendental Turn*, Oxford University Press, 2015, pp. 20—34.

④ Cf. D. Zahavi, “Phenomenology and Metaphysics”, In D. Zahavi, S. Heinamaa & H. Ruin eds., *Metaphysics, Facticity, Interpretation*, Kluwer Academic Publishers, 2003, pp. 3—22; D. Zahavi, “Internalism, Externalism, and Transcendental Idealism”, *Synthese*, 160 (3), 2008, pp. 355—374; D. Zahavi, “Husserl and the ‘Absolute’”, C. Ierna, H. Jacobs & F. Mattens eds., *Philosophy, Phenomenology, Sciences: Essays in Commemoration of Husserl*, Springer, 2010, pp. 71—92; D. Zahavi, *Husserl’s Legacy: Phenomenology, Metaphysics, and Transcendental Idealism*, Oxford University Press, 2017.

⑤ E. Husserl, *Transzendentaler Idealismus. Texte aus dem Nachlass (1908—1921)*, Husserliana 36, R. Rollinger ed., Kluwer Academic Publishers, 2003, p. xiii.

外部之物。不过，胡塞尔接着说这种理论非常荒谬。这种理论把意识当作储存着与外物相似的表征的箱子，但却忘了问我们如何能知道表征或错误表征 [(mis) representations] 确实是对外部事物的表征或错误表征：

自我 (ego) 不是注视这些图像的箱中的小人，并且还偶尔离开箱子去比较外部对象与内部对象，等等。对于这种观察图像的自我来说，图像本身就成了外部事物；而内部图像又需要它自身匹配的内部图像，如此以至无穷。^①

表征主义招致了众所周知的怀疑论：为什么对一个事物（一个内部对象）的觉知会促成对另一个相当不同的事物（一个外部对象）的觉知？而且，我们怎能知道内部可通达的事物确实对应了某个外部事物？然而，根据胡塞尔的反表征主义观点，心智和与世界——感知与实在——之间的契合和关联不是外在或巧合的“意识（心智过程）与实在事物绝不是并列对照的东西，和平地各居一边，只是偶尔相互关联”^②。这种说法在胡塞尔的作品中随处可见。多年以后，他在《笛卡尔式的沉思》一书中写道，认为意识与真正的存在仅仅是外在的关联是荒谬的，事实上它们本质上是相互依赖和统一的^③。胡塞尔的观念论不是还原式的观念论。胡塞尔不是那种把世界还原为感觉复合物的现象主义者。他的对手不是二元论，而是那些认为实在决然地独立于心智的激进客观主义者。否认激进的客观主义，否认“真实存在物的宇宙”在“所有可能意识、可能知识、可能证据的领域之外”^④，并不是说实在完全存在于心智中，也不是说实在是心智内的建构，而是说，实在必定是可显现的，必定在原则上是可被意识通达和获得的。我想，萨特在 1939 年的文本中对胡塞尔意向性理论的解释相当正确：

相对于经验批判和新康德主义的消化哲学，相对于所有“心理主义”，胡塞尔坚持认为人们不可能把事物消融在意识中。你看这个树，你确实看到了。但你是在树所在的空间中看到它：树在路边，在雾中，在炎热中孤独而干枯，在离地中海海岸八英里的地方。树不能进入你的意识，因为树与意识的本性不同。[……] 但胡塞尔也不是实在论者：在它那一小块干枯的土地上的这棵树并不是一个随后才与我交互的绝对的东西。意识和世界是一下子被给予的：世界虽然本质上在意识之外，但世界本质上又与意识相关。^⑤

按此解释，胡塞尔是 4E (embodied , embedded , enactive , extended) 认知进路——而不

① E. Husserl , *Transzendentaler Idealismus. Texte aus dem Nachlass (1908—1921)* , Husserliana 36 , R. Rollinger ed. , Kluwer Academic Publishers , 2003 , p. 106.

② E. Husserl , *Ideas Pertaining to a Pure Phenomenology and to a Phenomenological Philosophy , First Book , General Introduction to a Pure Phenomenology* , F. Kersten trans. , Martinus Nijhoff , 1982 [1913] , p. 111.

③④ E. Husserl , *Cartesian Meditations: An Introduction to Phenomenology* , D. Cairns trans. , Martinus Nijhoff , 1960 [1929/1950] , p. 84; p. 84.

⑤ J. -P. Sartre , “Intentionality: A fundamental Idea of Husserl’s Phenomenology” , *Journal of the British Society for Phenomenology* , 1 (2) , 4—5 , 1970 [1939] , p. 4.

是神经表征主义所推行的那种当代形式的笛卡尔式的内在主义——的天然同盟。当 F. J. 瓦雷拉 (F. J. Varela)、E. 汤普生 (E. Thompson) 和 E. 罗施 (E. Rosch) 在《具身心智》(*The Embodied Mind*)^① 中试图寻找中道——在那种把认知当作是对预先存在的外部世界 (实在论) 的恢复的斯库拉式认知 (Scylla) 式认知与另一种把认知当作对预先存在的内部世界 (观念论) 的向外投射的卡律布狄斯式认知 (Charybdis) 之间的中道——时, 以及当他们认为他们陈述的目标是超越内部之于外部的表征主义框架时, 他们清楚地表达了一种与胡塞尔立场有众多亲缘关系的生成主义。

最后——并且在此语境中最重要的是——正是被感知的日常对象 (而不是其它东西) 才是“物理学家按照物理学方法所探索和科学地确定的”^② 东西。我在星空中观察到的行星、我饮用的水、我欣赏的花等日常对象, 也正是自然科学家正在研究的对象, 并力图尽可能以精确和客观的方式确定它们的真正本质:

他 [物理学家] 所观察的、对之进行试验的、持续观看、放在手中、放到天平或熔炉中的物理事物: 那个物理事物, 而非别的, 成为物理学谓词 (诸如重量、温度、电阻等) 的主语。^③

简言之, 对胡塞尔来说, 这里不存在两种存在论上不同的东西, 显现的 (心内的) 对象和物理的 (心外的) 对象。确切地说, 只存在一个显现的 (心外的) 对象, 它具有范畴上不同但相互兼容的感觉规定和理论规定。这也是为什么科学发现与日常经验、科学图景与显现图景两者之间并不存在相互矛盾的原因。两者可以按各自的标准都为真。更一般地说, 感知的世界与科学的世界之间的区别不是那种为我世界与世界本身之间的区别 (分别落入现象学与科学的领域)。它是世界的两种显现方式之间的区别。科学世界不是一个自治的世界, 不是一个在显现世界背后或之下的世界。科学所研究的世界与日常经验的世界是同一个世界, 即显现的实在, 但现在它被科学的方式丰富和扩大了。如海德格尔指出的, 这也是为什么如下说法在现象学上是荒谬的:

在显象表征了和呈递了 (expresses) [其他之物] 的意义上, 把现象说成是某种背后之物的现象 (的说法是荒谬的)。现象不是某种背后藏有其他事物的东西。更准确地说, 人们根本不可能询问现象背后的东西, 因为现象所给出就是事物本身。^④

对胡塞尔来说, 物质的本质以知觉显现的方式被认识。将真正真实的实在 (really real reality) 定义为我们体验的未知原因, 并且认为被研究的东西只是一个完全不同的隐匿之物的迹象, 而其真实本质必定是未知的, 且永远不能按其自身规定被把握, 这种想法本身对胡

① F. J. Varela, E. Thompson & E. Rosch, *The Embodied Mind: Cognitive Science and Human Experience*, The MIT Press, 1991.

②③ E. Husserl, *Ideas Pertaining to a Pure Phenomenology and to a Phenomenological Philosophy. First Book. General Introduction to a Pure Phenomenology*, F. Kersten trans., Martinus Nijhoff, 1982 [1913], p. 119; pp. 120—121.

④ M. Heidegger, *History of the Concept of Time*, T. Kisiel trans., Indiana University Press, 1985 [1925], p. 86.

塞尔来说不啻是一种神话^①。与其把客观实在定义为事物本身所是，与其区分事物向我们显现的存在方式与事物决然的（simpliciter）存在方式，并以此认定对后者的研究才是真正重要的，胡塞尔则敦促我们面对如下事实：我们对客观性的通达以及客观性的本质必然涉及主观性和主体间性两者。事实上，与其作为客观性的对立面，与其构成科学知识的阻碍，主体（间）性对胡塞尔来说是客观性和科学的必要的能行条件。胡塞尔接受客观性和实在的现世概念，从而驳斥怀疑论，因为怀疑论主张世界向我们显现的样子兼容于实际上完全不同的（另外）世界。

如果胡塞尔的超越论哲学方案在今天看起来有些遥远和远离主流哲学，那主要是由于最近几十年我们目睹了自然主义的猛烈冲击。然而，值得注意的是，相比胡塞尔的超越论的观念论，我们在讨论神经表征主义时所例示的某些形式的自然主义实际上离常识实在论更远。

五、结论

M. R. 贝内特（M. R. Bennett）和 P. M. S. 哈克（P. M. S. Hacker）曾经指出，很多神经科学家苛责哲学的所谓失败，指责哲学在 2500 年的历史中没达成任何有科学价值的东西，但却对自己的大部分思想框架中蕴含有多大程度的哲学遗产浑然不觉。贝内特和哈克强调了 17 世纪认识论和形上学的持续影响^②。不过，正如我在上面所论证的，我们也不应忽视最近复活的一些起源于 19 世纪中期的观念。然而，无论哪种情况，关键在于忽视传统是不明智的。那样我们不仅会有错失重要见解的风险，还会有重复过去错误的风险。^③

（D. Zahavi, “Brain, Mind, World: Predictive Coding, Neo-Kantianism, and Transcendental Idealism”, *Husserl Studies*, 34 (1), 2017, pp. 47—61 译者单位：浙江大学哲学系 责任编辑：马新晶）

① E. Husserl, *Ideas Pertaining to a Pure Phenomenology and to a Phenomenological Philosophy, First Book, General Introduction to a Pure Phenomenology*, F. Kersten trans., Martinus Nijhoff, 1982 [1913], p. 122.

② M. R. Bennett & P. M. S. Hacker, *Philosophical Foundations of Neuroscience*, Blackwell, 2003, p. 134, p. 407.

③ 我很感激 L. 德·扣克（Liesbet de Kock）和 J. 迈克尔（John Michael）的有益评论。

means that the world picture of modern science denies the pre-scientific and everyday experience , which is the basis of Aristotle's doctrine of substance.

Scanlon's Justification on the Equality of Human Moral Status

Li Yaming

The justification for the equality of human being is a task brought to us by the *Universal Declaration of Human Rights*. Also , it is a necessary approach for understanding and resolving the disturbing moral dilemmas of our time. However , this justification has been proved to be unbelievable difficult. Each approach shows serious defects. Scanlon 's Contractualism defined moral states as an interpersonal relationship of mutual recognition , thus avoided all contradictions that had caused by other kinds of justifications. It proved that there is no hierarchy in moral status , and the realm of moral status accommodates all members of human being even those who lack of typical human capacities. By deliberating why human being have equal moral status , it also clarified the source of moral duty and the source of normativity from a special perspective , provided us with reliable principles to deal with moral difficulties today.

Experimental Semantics and Theories of Reference

Ren Yuan

Experimental Semantics put forward serious criticism to Kripke 's theory of reference and its methodological foundation. We reconstruct the line of argument of experiment semantics and carefully examine its three core claims , that is , thesis of cultural diversity of referential intuitions , thesis of cultural diversity of semantic reference of names and thesis of referential pluralism. It is argued that all these claims are defeasible. But we also use empirical evidences of naming practice to explain away the apparent diversity of referential facts in Chinese and Western traditions.

Brain , Mind , World: Predictive Coding , Neo-Kantianism , and Transcendental Idealism

Dan Zahavi

Recently , a number of neuroscientists and philosophers have taken the so-called predictive coding approach to support a form of radical neuro-representationalism , according to which the content of our conscious experiences is a neural construct , a brain-generated simulation. There is remarkable similarity between this account and ideas found in and developed by German neo-Kantians in the mid-nineteenth century. Some of the neo-Kantians eventually came to have doubts about the cogency and internal consistency of the representationalist framework they were operating within. In this paper , I will first argue that some of these concerns ought to be taken seriously by contemporary proponents of predictive coding. After having done so , I will turn to phenomenology. As we shall see , Husserl's endorsement of transcendental idealism was partially motivated by his rejection of representationalism and phenomenism and by his attempt to safeguard the objectivity of the world of experience. This confronts us with an intriguing question. Which position is best able to accommodate our natural inclination for realism: Contemporary neuro-representationalism or Husserl's transcendental idealism?