·科学技术哲学·

从数据到证据：表征主义和关系主义之争

From Data to Evidence: The Debate Between Representationalism and Relationalism

胡瑞斌 /HU Ruibin

（南开大学哲学院，天津，300350）
(College of Philosophy, Nankai University, Tianjin, 300350)

摘　要：数据在科学研究中扮演着至关重要的角色，其核心价值在于可为相关现象、假设或理论提供证据。对于数据何以发挥证据作用，表征主义解释强调数据本身即具有表征价值，关系主义解释则强调其语境相对性。两者表面上的对立源于对其观点过分强化的误读：强表征主义坚持数据本身具有固定的信息内容，因而具有超越任何探究语境的证据地位；强关系主义则主张数据的证据地位完全依赖于探究语境，甚至只有充当证据的数据才构成数据。然而，表征主义不必然走向强表征主义，关系主义也不必然走向强关系主义。表征主义和关系主义并非矛盾，而是可以彼此兼容：两者侧重于不同的探究阶段、学科领域，而且在数据的表征价值、语境依赖性等论点上可达成一致。

关键词：数据　证据　表征主义　关系主义

Abstract: Data plays a crucial role in scientific research, with its core value lying in its ability to provide evidence for related phenomena, hypotheses, or theories. Regarding how data serves as evidence, representationalism emphasizes that data itself has inherent representational value, while relationalism highlights its contextual relativity. The apparent opposition between the two arises from misinterpretations with too strong standpoints. Strong representationalism insists that data itself contains fixed informational content, thus possessing an evidential status beyond any specific context of inquiry. In contrast, strong relationalism argues that the evidential status of data depends entirely on the context of inquiry, even claiming that only data serving as evidence qualifies as data. However, representationalism does not necessarily lead to the strong version, and neither does relationalism. Representationalism and relationalism are not contradictory but can be compatible with each other. They focus on different stages of inquiry and research fields, and they can reach a consensus on points such as the representational value of data, its context-dependence, as well as its public and social nature.

Key Words: Data; Evidence; Representationalism; Relationalism

中图分类号：N02; C916　DOI: 10.15994/j.1000-0763.2025.05.004　CSTR: 32281.14.jdn.2025.05.004

现代科学研究的一个核心方法为，通过观察、实验或计算机模拟等手段获取数据，分析建模提出假设，从而对目标对象进行描述或表征。随着数据科学的兴起和大数据技术的广泛应用，数据已成为科学知识发现不可或缺的基础。无论是数据驱动的还是理论驱动的科学研究模式，均认可一个根本的前提预设，即数据是有价值的，可以充当相关现象、假设或理论

基金项目：天津市哲学社会科学规划项目“模型表征的视角主义研究”（项目编号：TJZXQN23-001）。
收稿日期：2024 年 12 月 24 日
作者简介：胡瑞斌（1993–）男，河北邢台人，南开大学哲学院讲师，研究方向为科学哲学。Email: huruibin2012@163.com

的证据。这一预设通常是被默认的，而且被认为无可争议，否则数据难以在科学研究中占据如此重要的地位。

相比之下，数据在哲学讨论中长期未获得应有的重视。直至20世纪60年代苏佩斯（P. Suppes）的数据模型研究，特别是80年代末波根（J. Bogen）和伍德沃德（J. Woodward）对数据和现象的区分，数据才开始逐步获得哲学家们越来越多的关注和讨论。在波根和伍德沃德看来，数据是可观察的，现象则是不可观察的，科学理论的解释对象是现象，而非数据，但数据构成了相关现象的证据。[1]问题在于，数据何以充当证据，或者说，从数据到证据需要何种“认识论程序”？

除了一般性地主张数据可充当相关现象的证据，波根和伍德沃德还指出：“认为数据可靠性的评估需要对有关数据的事实进行系统解释，这是完全错误的……数据是否构成某些现象的可靠证据的问题，取决于数据是否可复制、各种混杂因素和其他可能的系统误差来源是否得到适当控制、各种统计论证以及数据分析和简化程序等考虑因素。”[1]简言之，对于数据证据地位的考量无需诉诸理论解释，可根据相对独立的程序加以确定。

这种理论独立性似乎意味着，数据具有“内在”的证据价值，可以脱离具体的探究语境。莱奥内利（S. Leonelli）称这种观点为数据的“表征主义”解释：“数据是具有固定信息内容的表征实体，描述了现实的特定部分，且独立于数据的考量环境……在分析数据时，重要的是揭示它们记录了现实的哪些方面；事实上，其认识论意义源于它们以独立于语境的方式表征现实的这些方面的能力。”[2]“无论在任何时间以何种方式使用数据，都可以对其进行识别，而且可以在不参考相关研究语境的情况下，客观地评估特定数据集包含哪些信息，以及对这些信息的解释正确与否。”[3]

然而，这种表征主义解释隐含的客观主义理解显然是有问题的。一方面，观察、实验都有一定的理论负载性，对数据的可靠性评估的确可以诉诸相对独立的经验程序，但是这并不意味着数据的产生、处理、分析和解释不需要任何理论的介入。例如，在地球和大气科学领域，对数据的解释是一种有价值的策略，“它使理论在现象识别中占据了中心位置”。[4]另一方面，谈及数据的证据地位，本身就意味将数据视作某个现象、假设或理论的证据，为之提供支持或者反驳，因而必然需要基于两者间的关系，而非仅仅基于对数据可靠性的独立评估来理解。

因此，莱奥内利等学者坚持对数据的“关系主义”解释：“数据是一种表达和呈现信息的特定方式……作为证据来源，其科学意义取决于使用它的情况。根据这种观点，数据本身并不具有真值，也不能被视为特定现象的直接表征。相反，数据本质上是可替代的对象，由其可移植性和作为证据的预期有用性所定义。”[2]简单来说，数据本身不具有“内在”的证据价值，不仅数据的证据地位取决于具体的探究语境，而且数据的本质界定，即到底什么是数据，也因具体语境而异：“相同的对象可能作为数据，亦可能不作为数据发挥作用，这取决于它们在科学探究中扮演的角色。”[2]因之，数据的认识论和本体论理解都被赋予了极强的相对主义色彩。

乍一看来，表征主义和关系主义解释之间存在尖锐的观点对立。然而实际上并非如此。对于表征主义和关系主义，均存在一些误读，特别是对其观点的过分强化，我们可以称之为“强表征主义”和“强关系主义”解释：前者认同数据本身具有固定的信息内容，因而具有超越任何探究语境的证据地位；后者则主张数据本身不具有证据价值，其价值完全依赖于探究语境，甚至只有充当证据的数据才构成数据，证据地位是数据之为数据的必要条件。本文旨在论证，正是这些过分强化的解释造成了表征主义和关系主义之间表面上的对立。在对这些误读加以澄清之后，可以看到两者并非矛盾，而是有不同的侧重，可以相互兼容。

一、表征主义解释

数据可以用来表征现实世界中的对象，可

以充当后者的证据，这一论点没有什么争议。有争议的是，应当在何种限度内理解数据的证据价值或表征价值？由于波根和伍德沃德坚称，数据的证据地位考量无需诉诸系统的理论解释，他们通常被视作强表征主义的支持者。然而，这是一种误读。表征主义解释的核心在于，数据具有证据价值或表征价值，但这并不意味着它们完全独立于任何探究语境，最终走向强表征主义。

波根和伍德沃德之所以强调数据证据地位的理论独立性，一方面是因为无法为数据提供系统的理论解释，另一方面则是因为无需提供这种解释。之所以无法提供，是因为数据的产生涉及诸多因素的混杂影响。“这些因果因素数量太多、种类太杂、行为太不规则，任何单一理论都无法顾及所有……这种多样性、异质性和复杂性通常会使数据点无法像现象那样可以获得解释。”[5]另一方面，之所以无需提供，则是因为存在其他有效的方法论策略。例如，控制可能的混杂效应和系统误差，考察实验或探究结果是否具有可重复性，统计分析和数据简化程序方法是否可靠，以及仪器设备的校准和经验调查等等。[6]

当然，仅凭以上经验手段还不足以确立数据的证据地位，还需一个关键的前提预设，将数据和以之为证据的对象关联起来。波根和伍德沃德诉诸的本质上是一种因果关联：数据的产生是现象以及各种因素作用之下的因果过程的产物，数据的证据价值或表征价值就来源于，它们借由特定因果过程记录了所研究对象的特征属性。鉴于数据是所研究的对象以及各种混杂因素的影响之下的产物，我们不可能为这种混杂的影响提供系统的理论解释，那么必然的选择就是诉诸前述的经验策略，处理分析数据提取可靠的“信号”，可以构成所研究对象的证据。

尽管数据并非由世界直接给定，而是我们与世界互动的产物，无法根本摆脱其人造本质，但数据毕竟是由这种互动的因果过程产生的，受到物理世界的强烈约束。作为一种公共记录，数据还可以在一定程度上摆脱个体感官经验的主观性，以相对独立的方式提供有关探究对象的信息。因之，数据被赋予了突出的客观性特征，强表征主义的解释亦颇有市场。例如，有观点认为，“数据根据其内在属性被视为证据，这些属性是固定的、内在的、局部的，并与现实的各个方面存在表征关系，独立于数据的使用语境”。[7]或者认为，“一旦数据产生，无论是否要被继续处理，它都即刻且持续地充当证据角色，在某种程度上支持对象（现象/目标系统）”。[8]

根据莱奥内利的理解，表征主义解释（实为强表征主义）的核心论点包括两个假设：“第一个假设是，数据是世界的表征，因为它们捕捉到了现实的某些特征，从而使它们适于科学研究。这意味着，无论研究者如何使用数据，其信息内容都是固定的，并确定了数据作为知识主张的主要经验基础的身份……第二个假设：数据是检验和验证理论的主要手段，科学方法可以保证其在这方面的可靠性”。（[9]，pp.73–74）一般而言的表征主义解释认可第一个假设的部分观点，即数据本身即具有表征价值，但不必然认同更强的主张，即数据的信息内容是固定的，可以超越探究语境。

莱奥内利反对表征主义解释的一个依据是：“承认世界的物理属性与数据对象的物理属性之间存在因果关系，并不必然就能将数据理解为世界的表征……‘世界之所是’并不是制造数据对象的唯一原因……研究者与其环境之间的互动往往是以大量技术为媒介并经过精心安排的。因此，仪器和环境条件也是决定这些互动产生的对象的物质特征的因果因素之一。”[10]然而，这一论点并不构成对表征主义的有效反驳。

如前文所述，产生数据的因果过程涉及诸多混杂因素的影响，并非所研究对象单独作用的结果，否则根本无需进行数据分析解释以提取相关“信号”。在此意义上，数据的表征价值具有不确定性或“不充分决定性”，但这并不意味着数据本身不具有任何表征价值，毕竟数据依然是特定因果过程的产物，至少部分反映了“世界之所是”。此外，这也不意味着数

据的信息内容是固定的。数据表征的对象或者为之提供证据的对象，是具体探究过程的结果。我们没有任何上帝视角，可以事先知晓所研究的实体、现象或过程究竟是什么样子的，而且在不同的探究活动中，我们也会有意地选择性关注其中的不同对象，因而数据并非只具有固定的表征价值，同样需要考虑具体的探究语境。

实际上，波根和伍德沃德并未完全将探究语境排除在外："数据或任何特定现象探测技术的可靠性，取决于探测过程中的各种要素是否共同发挥作用，使探究者能够正确区分有关现象的各种相互竞争的主张。"[11]在后续的讨论中，伍德沃德更进一步指出："从数据到现象的推理……得出的结论（关于现象的主张）超越了它所依据的证据（数据），或者说具有除证据之外的额外内容……这种推理的特点是，它总是需要在证据之外附加实质性的经验假设。"[12]

波根和伍德沃德甚至不排斥适度的相对主义理解："数据的可靠性是相对于用来产生和解释数据的过程和技术，以及需要区分的可能性而言的。同样的数据，相对于一种解释程序或一组备选主张可能是可靠的，但相对于另一种解释程序或一组备选主张则是不可靠的。"[11]不过他们依然坚持，具体探究语境涉及的数据生产和解释的过程技术、备选主张或者实质性的经验假设虽然不可或缺，但是不等同于对数据的系统性理论解释，我们不可能也不需要为数据提供这样的解释。就此而言，数据的证据地位仍然具有相当的理论独立性，但以上的讨论已经足以表明，这并不等同于莱奥内利等学者所理解的强表征主义解释。

最后，莱奥内利所言的第二个假设同样是一种误读。至少波根和伍德沃德曾明确指出，数据与现象在证据地位方面也存在重要差异：在证据作用方面，数据与现象的区别在于，有关数据的事实通常作为有关现象的事实的证据，而有关现象的事实则"通常是解释这些现象的高层级一般理论的证据"。[1]因此，理论检验通常分为两个阶段："第一阶段，数据产生并获得解释，从而得出关于现象的结论……第二阶段，理论主张与第一阶段得出的有关现象的结论相对抗。"[13]换言之，数据无法直接充当一般理论的证据，需要以现象为中介，而有关现象的事实则是通过数据的处理分析解释而获得的，数据可为之提供证据。对于数据的证据地位考量而言，这是一个更为细化的认识论程序，从数据到现象，再从现象到理论，任何一个阶段都需要具体考察，并非强表征主义刻画的那样，凭借固定的信息内容，数据即可直接发挥证据作用。

二、关系主义解释

从本质上看，证据是针对另外一个对象的证据，因而至少涉及两个要素间的关系：X是支持或者反对Y的证据。就数据而言，数据通常被视作特定现象、假设或理论的证据，因此必然要考虑两者间的关系，而且离不开具体的探究语境。这是第一种一般意义上的关系主义解释。第二种解释则关注数据本身的本质界定，主张应当从事物间关系的角度来回答"什么是数据"这一根本问题，其中核心的"关系"通常被理解为证据关系，由此"是否构成证据"成为回答"什么是数据"的标准依据，这正是前文所言的强关系主义解释。关系主义的支持者并未明确区分这两个层面的解释，然而关系主义不必然走向强关系主义，其与强表征主义的对立也不必然意味着与表征主义的对立。

在第一种一般意义上的关系主义解释下，数据能否发挥证据作用，取决于具体的探究语境，诸多因素都可能被包含在语境之中："使数据得以传播并用作科学证据的条件千变万化，取决于多种因素，从相关研究者使用的语言、概念框架、材料和工具，到可用的物质基础设施、与数据相关的政治和经济价值，以及科学机构和其他机构对数据流通提供支持和管理的程度，不一而足。"（[9]，p.194）基于此，莱奥内利反对强表征主义的解释，强调数据证据价值的开放性："数据本身不具有固定的科学价值，也不能被视为对于特定现象的独立于思维的表征。相反，它们是由在特定探究时刻赋予它们的证据价值所定义的"，在以数据为中心

的研究中，“数据有可能支持科学家尚未提出、甚至可能永远不会提出的主张，简言之，数据具有潜在无穷尽的证据价值”。（[9]，p.70）

这种关系主义解释十分正确地强调了数据证据地位的相对性、多样性、可变性，相同的数据可以被不同的研究者在不同的探究语境下使用，其证据地位和支持强度也会因之而异。莱奥内利称此过程为“数据传播”（data travel），其中包括两个阶段：一是“去语境化”（decontextualization），“让数据失去其在原初研究语境下被赋予的意义”，从而“让数据适应新的研究环境，这是通过尽可能多地剥离数据的限定条件来实现的”；二是“再语境化”（recontextualization），“使用者能够通过查阅元数据评估数据来源，从而评估数据的潜在意义”。（[9]，p.30）正是在此意义上，莱奥内利主张数据不具有固定的表征价值。

然而，如前文所论，数据表征价值的“不充分决定性”，不意味着对其表征价值的彻底否定。同时，这也不意味着走向极端的相对主义立场，认为数据的价值可以彻底脱离原初语境，完全由不同的语境来决定。莱奥内利尽管强调去语境化，但同时也强调了“元数据”（metadata）的重要地位：元数据提供了有关数据来源即数据生产过程的信息，“没有元数据，研究者将无法判断在线数据的可靠性，而可靠性取决于数据是由谁、出于何种原因、在何种环境下提供的”。（[9]，p.29）即便是再语境化阶段，研究者同样需要基于元数据考察数据的可靠性，以确定是否可将之（以及如何）用于新的探究语境。

显然，去语境化和再语境化都无法彻底剥离有关数据生产过程的信息。“经验结果的认识论效用在很大程度上取决于其来源的细节。在认识论上负责任地使用经验结果（如数据）取决于能否获取相关的元数据：数据永远无法与其相关的丰富信息永久分离并保持其认识论效用。”[14]因此，莱奥内利提出的“去语境化”概念蕴含了过分强化的要求，容易引发将之归结为强关系主义的误解。实际上，之所以要让数据广泛传播，而且确实能够传播，正是因为数据本身具有一定的表征价值，在不同的语境下有不同的呈现。即便在某些语境中数据不发挥证据作用，也不能证明数据不具有任何表征价值，而只能说明，该数据集特定方面的表征价值并非某些研究者关注的对象。

值得注意的是，强关系主义不仅将数据的证据地位完全相对化，否定其表征价值，而且也从相反的角度，将证据关系视作数据之为数据的必要前提，从而将关系主义解释渗透进对数据本质的理解。莱奥内利所言的数据传播不仅影响对数据证据地位的考量，而且决定了什么是数据：“数据为以下对象：（1）被视为有关现象的一种或多种主张的潜在证据；（2）其格式和处理方式使其能够在个体或群体之间流通，以便进行分析。……只要符合上述两个要求，任何对象都可以归为数据。什么算作数据取决于谁使用它们、如何使用以及出于什么目的使用。”（[9]，p.78）

可见，数据的证据地位似乎成为数据之为数据的必要前提。这意味着，通常意义上的数据如果在某些语境下不作为证据发挥作用，就不再是数据。然而，这种理解与数据驱动的知识发现逻辑是相悖的：研究者首先掌握了大量数据，然后通过分析处理探究其可作为何种现象或假设的证据，而非首先掌握了证据，然后决定将某对象视为数据。再者，某些数据收集的目的并非验证某个科学主张，而是为了进行数据的详细编目，而且“一组数据也可能因为测量设备损坏或经验不适当的理论假设等原因，变得糟糕或无用”。[15]在这些情形中，我们并没有适当的理由拒绝称通常意义上的数据为“数据”，数据并非没有扮演“数据”角色，而只是没有扮演“证据”角色。

以基因测序为例，无论在何种探究语境下，测序结果都是通常理解的“数据”。如果探究目的是了解某遗传疾病的成因或者对特定生物材料进行鉴定，那么测序结果可发挥证据作用；如果探究目的是了解某物种的基因序列，将之纳入生物基因库，而非检验任何假设，或者样本被污染，检测出现极大偏差，那么测序结果没有发挥证据作用，但是这并不妨碍我们称之

为“数据”。

强关系主义的支持者可能做如下回应：数据定义要求的是其作为“潜在证据”，并不要求其在具体语境下实际发挥证据作用。因此，尽管纳入基因库的测序结果没有实际发挥证据作用，但是这并未排除其在其他探究语境中作为证据的可能性，而且即使样本被污染，检测结果也可能作为其他假设（如“样本被污染”）的证据，因而依然可称之为数据。然而，我们可以设想出无数种可能的假设，经验结果可为之提供证据，于是任何结果都可作为“潜在证据”因而构成“数据”，这就使得强关系主义对于数据的本质定义变得琐碎而无意义。

更关键的问题是，何以确定某经验结果可作为“潜在证据”？存在两种可能的选择。第一种选择是，诉诸其自身的表征价值。正是由于经验结果是特定的因果过程的产物，至少反映了现实世界的部分特征，因而可作为特定现象的证据，但是如此便背弃了强关系主义的核心信条。第二种选择是，主张经验结果的证据地位是探究过程的结果，无需预先为之确定任何价值。这意味着，我们需要关注从经验结果到证据的具体探究过程，关注它们在何种语境下发挥证据作用，进而在此语境下称之为“数据”。然而，这只是对经验结果的重新命名。本质上而言，将证据地位视作数据之为数据的必要条件，实际上是将二者简单等同起来，忽视了从数据到证据同样需要一定的“认识论程序”。特别是在不同的探究语境下，从数据到证据的推理存在可靠性强弱的区分。[16] 这一问题显然无法仅仅通过强关系主义对数据的本质界定提供解答，对于数据究竟如何以及在何种强度上发挥证据作用的解释，才是关键所在。

三、争论的消解

以上讨论已经表明，表征主义不必然走向强表征主义，关系主义也不必然走向强关系主义。表征主义虽然强调数据本身的表征价值，但并非完全不考虑探究语境；同样地，关系主义虽然强调基于探究语境来理解数据，但并非彻底否定其本身的表征价值。实际上，表征主义和关系主义解释之间的矛盾对立源于对其观点过分强化的误读，在对之进行澄清之后，我们可以看到两者在适用的探究阶段、学科领域等有不同的侧重，而且在诸多论点上可达成一致。

首先，表征主义和关系主义解释侧重于不同的探究阶段。根据波根和伍德沃德的观点，从数据到现象的推理可分为两个阶段：数据生产和数据解释。数据生产关注数据得以产生的因果过程，“涉及设定因果相互作用的系统或定位预先存在的系统，这些系统可以产生探究者可以感知和解释的效果，而且使探究者能够区分有关现象的竞争主张”；“数据解释包括使用论证、分析技术和推理模式，对产生的数据进行操作以得出有关现象的结论。这一阶段的典型组成部分是数据分析和简化技术……也可能包括明确使用背景假设或理论假设”。[17]

鉴于数据的表征价值来源于其自身是特定因果过程的产物，数据能否构成特定现象的证据首先便取决于数据生产过程是否可靠。波根和伍德沃德诉诸的方法论策略包括控制混杂效应和系统误差、结果可重复性检验、仪器设备校准等，均侧重于数据生产。“数据生产过程的特征具有至关重要的证据意义，因为它们会影响到在获得不同的现象主张时会产生哪些数据结果，从而影响到关于这些主张的哪些结论会被接受。”[17] 数据解释阶段的数据分析简化操作以及理论假设的介入同样也会影响数据证据地位的考量，但波根和伍德沃德并未对此具体展开论述，他们只是强调解释的变化不会影响生产阶段下的数据本身的特征，反过来讲，这也意味着数据可以支持多样的解释。

相比之下，关系主义更侧重于数据解释。莱奥内利所言的“数据传播”准确捕捉了对数据进行多元解释的必要性：“探索数据作为证据的价值并使之最大化的最佳方式，就是让尽可能多的研究者在其研究语境中以自己的方式使用数据。”（[9]，p.26）特别是在“大数据”时代，大量数据的快速流通使得诸多研究者有机会在其自身的语境中开展研究，同时面对复杂的数据，研究者通常不具备全面系统的视角对

其整体进行处理分析，因而从“去语境化”到“再语境化”的数据传播势在必行，从而最大限度地推进科学知识发现。当然，记录了数据生产信息的“元数据”同样不容忽视，在这一方面表征主义无疑可以提供必要的补充解释。

其次，表征主义和关系主义解释适用的学科领域亦有所区别。表征主义的支持者援引的案例大多来自物理学，如波根和伍德沃德着重分析的欧洲核子研究中心探测弱中性流的案例，而关系主义的支持者如莱奥内利主要援引的案例为以数据为中心的模式生物研究。这些学科领域存在较大的差异：在研究较为分散的生物学领域，“数据是在无数个地点和高度多样化的条件下产生的，数据解释是通过认识和比较数据生产和流通情况的独有特征来确立的”；“而对于数据来源少之又少，大部分科学工作都致力于解释的领域来说，情况可能就不是这样了。高能物理学中粒子加速器的数据生产和处理可能被视为这种情况的例证，因为数据生产是完全集中化的，因而数据流通似乎不涉及数据兼容性和进一步使用的相关性方面的决策”。（[9]，p.191）

因此，表征主义解释更适用于研究范式较为统一且活动集中的学科领域，而关系主义解释则更适用于研究方式多样且活动分散（甚至彼此隔绝）的学科领域。在统一的研究范式下，数据的生产、分析处理和进一步解释都有较为一致的理论框架和方法论选择，同时研究活动的集中开展也使得数据的使用通常不需要跨越不同的探究语境，因而首要的考虑便在于数据是否是由可靠的因果过程产生的，这决定了数据本身是否具有可靠的表征价值。

在缺乏统一的研究范式的学科领域，情况则大不相同。“从基因组学到神经生物学再到气候学，我们拥有大量数据，但理论理解通常有限……这些领域目前的发展水平是‘数据丰富而理论贫乏’。”[12]多元化的理论路径和方法论选择以及高度分散的研究活动使得数据传播具有更强的必要性，研究者除了关注数据生产的可靠性，还需要关注数据是否能够以及如何在新的探究语境下发挥证据作用，因而关系主义的解释更有针对性。数据传播有助于打破研究壁垒，实现数据资源的最大化利用，从而推进知识交流以及新知识的发现。

最后，表征主义和关系主义解释在诸多方面论点一致。关系主义并不完全否定数据本身的表征价值，表征主义也并非完全忽视探究语境，两者在这两方面均可达成共识，只不过各有所偏重。莱奥内利承认，“数据认识论的关系主义进路并不否认数据可以被赋予表征功能。事实上，以数据（某些方面）来表征特定现象，是数据作为特定知识主张的证据发挥作用的必要条件”。（[10]，p.16）波根和伍德沃德也强调证据关系对于探究语境的依赖性：“证据相关性不是某种形式关系的问题，而是经验事实的问题……证据相关性来源于大量与特定领域高度相关的事实……而非高度一般化的、独立于任何领域的形式关系”。[13]表征主义和关系主义相一致，都意图摆脱对于证据关系的纯粹形式化考虑，或者基于逻辑关系的理解，因而都强调证据关系的本质是一种经验关系，需要在具体的探究语境中加以确立。

此外，表征主义和关系主义解释均认可数据的公共性和社会性。莱奥内利指出，“科学主张的确立，包括对数据的可靠性和重要性做出判断，被广泛认为是一种社会活动，需要不止一个人的参与”。（[9]，p.80）波根和伍德沃德也认为，数据具有公共性，而且体现了科学活动的社会性，要使数据发挥其证据作用，“关键是数据生产的方式使其可以获得公开检查，通过复制相关实验和观察结果进行检验，并可由其他研究者随时分类、汇总和分析”。[11]具体到不同的探究语境，数据的这种公共性和社会性便呈现为莱奥内利所言的“数据传播”过程。从数据生产到数据解释，对于数据如何跨越不同的探究语境发挥证据作用，表征主义和关系主义的解释各有侧重，可以彼此兼容而且应当相互补充。

结　语

表征主义和关系主义之争，特别是强表征

主义和强关系主义之争，其背后更深层的争论在于，数据到底是世界的客观记录，还是人为建构的产物。因之这一争论的焦点通常被置于数据的本质理解："两种观点可简单表述为数据是事实的表征，还是单纯体现了事物之间的关系……这两种观点的差异在于：数据本身是否具有'意义'，或者说是否可以表征'事实'"。[18] 然而如前所论，表征主义和关系主义解释在这一方面并不存在根本差异，表面的对立冲突根源于过分强化的误读：强表征主义将数据本身的意义或者其所表征的事实视作客观固定的，而强关系主义则将之视为完全依赖于探究语境。然而，数据既非世界完全客观的记录，亦非纯粹人为建构的产物，而是科学活动中研究者与世界互动的结果。因此，即便在本体论层面，对数据的解释也应当避免极端的客观主义和建构主义立场。更重要的是，无论对数据的本质作何定义，数据与证据之间仍然存在概念性的区分，从数据到证据依然需要一定的"认识论程序"，这才是理解数据何以检验假说、探测现象乃至驱动知识发现的关键所在。

［参考文献］

[1] Bogen, J., Woodward, J. 'Saving the Phenomena'[J]. *The Philosophical Review*, 1988, 97(3): 303–352.

[2] Leonelli, S. 'What Counts as Scientific Data? A Relational Framework'[J]. *Philosophy of Science*, 2015, 82(5): 810–821.

[3] Leonelli, S. 'The Philosophy of Data'[A], Floridi, L. (Ed.) *The Routledge Handbook of Philosophy of Information*[C], London: Routledge, 2016, 191–202.

[4] Lusk, G. 'Saving the Data'[J]. *The British Journal for the Philosophy of Science*, 2021, 72(1): 1–21.

[5] Bogen, J. '"Saving the Phenomena" and Saving the Phenomena'[J]. *Synthese*, 2011, 182(1): 7–22.

[6] Woodward, J. 'Data and Phenomena'[J]. *Synthese*, 1989, 79(3): 410–422.

[7] Canali, S. 'Evaluating Evidential Pluralism in Epidemiology: Mechanistic Evidence in Exposome Research'[J]. *History and Philosophy of the Life Sciences*, 2019, 41(1): 1–17.

[8] 贾玮晗、董春雨. 论数据模型的特征及其认知功能[J]. 自然辩证法通讯, 2024, 46（4）: 23–30.

[9] Leonelli, S. *Data-centric Biology: A Philosophical Study*[M]. Chicago: University of Chicago Press, 2016.

[10] Leonelli, S. 'What Distinguishes Data from Models?'[J]. *European Journal for Philosophy of Science*, 2019, 9(2): 1–27.

[11] Bogen, J., Woodward, J. 'Observations, Theories and the Evolution of the Human Spirit'[J]. *Philosophy of Science*, 1992, 59(4): 590–611.

[12] Woodward, J. 'Data and Phenomena: A Restatement and Defense'[J]. *Synthese*, 2011, 182(1): 165–179.

[13] Bogen, J., Woodward, J. 'Evading the IRS'[A], Jones, M. R., Cartwright, N. (Eds.) *Idealization XII: Correcting the Model*[C], Amsterdam: Rodopi, 2005, 233–267.

[14] Boyd, N. M. 'Evidence Enriched'[J]. *Philosophy of Science*, 2018, 85(3): 403–421.

[15] Ströing, P. 'Scientific Phenomena and Patterns in Data'[D]. Bavaria: Munich. Ludwig Maximilian University of Munich, 2018, 25.

[16] Ströing, P. 'Data, Evidence, and Explanatory Power'[J]. *Philosophy of Science*, 2018, 85(3): 422–441.

[17] Woodward, J. 'Data, Phenomena, and Reliability'[J]. *Philosophy of Science*, 2000, 67(53): S163–S179.

[18] 张贵红. 论数据的本质及其与信息的关系[J]. 哲学分析, 2018, 9（2）: 119–132.

［责任编辑　王巍　谭笑］