

累积证据是可重复性危机的一个解法

Cumulative Evidence is a Solution to the Repeatability Crisis

王浩 /WANG Hao¹ 维克托瑞·道尔可 /DALKO Viktoria^{2, 3}

(1. 综合经济学研究所, 美国昆西, 02171, 美国; 2. 霍特国际商学院金融系, 美国剑桥, 02141, 美国;
3. 哈佛大学继续教育学院, 美国剑桥, 02138, 美国)

(1. Research Institute of Comprehensive Economics, Quincy, MA, 02171, USA;
2. Department of Finance, Hult International Business School, Cambridge, MA, 02141, USA;
3. Division of Continuing Education, Harvard University, Cambridge, MA, 02138, USA)

摘要: 科学研究的根本追求是获得具有客观确定性的科学证据, 而可重复的实验是实现这一目标的重要规范。相对于对照实验, 概念性重复和直接重复之间的异同以及它们各自独立性的优势得到了分析, 结果表明, 当前的可重复性危机本质上是直接重复方面的危机。概念性重复能够产生累积性证据, 并有助于理论构建, 而直接重复仅仅是证实科学事实的存在。因此, 累积性证据更受青睐, 因为它也为当前的可重复性危机提供了解决方案。文中给出了构建能量转化与守恒定律的案例, 其中运用了概念性重复所产生的累积性证据。用简单的数学模型描述累积性证据的形成过程, 建议实验科学家应更注重通过概念重复产生累积证据。

关键词: 可重复性危机 直接重复 概念重复 累积证据

Abstract: The basic pursuit of scientific research is scientific evidence with objective certainty, and repeatable experiments are a major norm to achieve this goal. The similarities and differences as well as the strengths of independence between conceptual repetition and direct repetition are analyzed relative to controlled experiments, revealing that the current repeatability crisis is essentially a crisis of direct repetition. Conceptual repetition can generate cumulative evidence and lead to theoretical construction, while direct replication confirms the existence of scientific facts only. Therefore, cumulative evidence is preferred, because it also provides a solution to the current repeatability crisis. The case of constructing the law of energy conversion and conservation is presented, in which the cumulative evidence generated by conceptual repetition is used. The formation process of cumulative evidence is described in simple mathematical form, and it is suggested that experimental scientists should pay more attention to conceptual repetition to generate cumulative evidence.

Key Words: Repeatability crisis; Direct repeatability; Conceptual repeatability; Cumulative evidence

中图分类号: N031; B81-0 DOI: 10.15994/j.1000-0763.2025.07.007 CSTR: 32281.14.jdn.2025.07.007

一、可重复性危机

可重复性是现代科学研究的三大基石之一,

是形成科学事实的规范和检验假说以通向科学理论的“裁判”。^{[1], [2]}可是进入21世纪后, 实验的可重复性困境越来越引起科学家和公众的注意, 其中影响最大的论文当数约安尼迪斯(John

收稿日期: 2024年7月12日

作者简介: 王浩(1963-)男, 安徽霍山人, 综合经济学研究所高级研究员, 研究方向为科学方法论、医学与科学哲学、经济增长理论、金融监管。Email: HWang@Rice-Institute.org

维克托瑞·道尔可(1963-)女, 匈牙利人, 霍特国际商学院金融系教授、哈佛大学继续教育学院教学教授, 研究方向为健康经济学、金融市场、古典哲学、科学方法论。Email: Viktoria.Dalko@Faculty.Hult.edu

P. Ioannidis) 发表于2005年的论文:“为什么大多数发表的研究结果都是错误的?”。该文本来是针对生物医学领域的实验统计显著性问题的讨论,因为显著性问题与可重复性问题直接相关,所以该文被几乎所有自然科学领域和部分社会科学领域的关于可重复性困难的论文所广为引用。^{[3], [4]}

《自然》(*Nature*) 期刊在2016年对1576个科学家作了问卷调查,问他们认为目前的科学研究中是否存在“可重复性危机”。有90%的回答说是,包括52%的回答说存在严重危机,38%的回答说存在轻度危机。不认为有危机的只有3%。值得注意的是,这1576个科学家涵盖了化学、生物学、物理学、工程、医学、地学和环境科学等主要科学领域。从这些科学家的个人经验发现,物理学实验的可重复性并不比医学实验的可重复性更高,而化学实验的可重复性最低。^[5]同时,公众对科学研究的可重复性常常达不到要求表达了越来越多的忧虑,这些广泛的学术研究和社会舆论引起了美国国会一部分议员的注意,并在2016年起草法案,同年通过了国会批准,次年为时任总统奥巴马签署正式成为法律。该法案是《美国创新和竞争力法案》(*American Innovation and Competitiveness Act*),对科学研究的可重复性问题作出了回应,并指示美国国家科学基金会与美国国家科学、工程和医学院(*National Academies of Sciences, Engineering, and Medicine, NASEM*)合作,评估科学和工程研究的可再现性和可复制性现状,并且提供研究结果和建议,以提高研究的严谨性和透明度。NASEM据此在半年时间里完成了一份共识研究报告,题目是《科学中的可再现性和可复制性》(*Reproducibility and Replicability in Science*),出版于2019年底。该报告经过了严格且独立的同行评审,它代表了NASEM的立场。^[2]

本文的研究计划是在第二节简要分析两份权威学术组织的共识性报告,了解经验研究的可重复性对建立具有客观确定性的科学证据起着关键的作用;第三节针对可控实验分析概念重复(导致累积证据)与直接重复的异同和独

立性强弱,找到当前的可重复性危机的主要原因,并理解之所以上述NASEM报告更倾向于累积证据,是因为它对可重复性危机提供了一种有效解决方案;第四节呈现一个科学史上的累积证据的案例;第五节用简单数学形式描述累积证据的形成过程,强调了概念重复不仅能够确认科学事实,还有助于理论建构;第六节是结语,并且建议实验科学家应从唯直接重复的观念拓宽到既着眼直接重复更重视概念重复(以产生累积证据)的视野。

二、可重复性是科学研究的三大基石之一

在了解NASEM报告之前,有必要了解另一份报告,它是在1990年出版的《面向全体美国人的科学》(*Science for All Americans*),作者是美国科学促进会(*American Association for the Advancement of Science, AAAS*)。该报告源于AAAS在1985年开始的一个全国性科普项目——“2061计划”,目的是让美国的中小学生学习科学的基本知识。AAAS报告中的第一章分为三个部分。第一部分是科学的本性,第二部分是科学的探究,第三部分是科学的事业。本文专注于其中有关可重复性的内容,并把它们概括如后。第一,科学预设了“世界是存在的和世界是可探究的”这两个命题(世界包括自然界和人),这是自然科学各个学科共有的自然观;第二,科学还假设宇宙是各向同性的,这是可重复性要求的自然观基础;第三,证据对科学至关重要,而证据往往由多人提供,可重复性检验是科学家们相信新发现是科学证据(与科学事实对等)的一个实际操作过程;第四,科学探究没有一个通用的方法,逻辑和想象力虽然很重要,但科学证据才是终极“裁判”。其他内容基本上围绕的是探究过程和结果的不确定性与消减方法,以及科学理论的可预测性。^[1]

从AAAS报告可以总结到:科学探究的可重复性、不确定性和可预测性概括了科学知识体系的根本特性,它们三者是现代科学不可或缺的基石。可重复性与可预测性是分别对经验

研究和理论研究的规范式要求。不确定性是描述性事实。针对经验研究,它面临来自研究对象、研究方法、技术手段、人为失误、系统偏倚、研究地点、研究时间等多种来源产生的不确定性。经验研究的目的是消除或限制不确定性从而获得具有一定确定性的证据。

NASEM报告定义了“科学是什么”:“科学是一种探究模式,旨在提出关于世界的问题,通过科学家们的共同努力,从而得出答案并确保它们有充分的根据,并评估这些答案的确定性的程度。”([2], p.27)可以看出,每一项科学研究的终极追求是确定性,而非只是客观性,所以有一定客观确定性的证据对科学是至关重要的。而要使一项探究结果成为科学证据甚至最终成为科学知识,一个重要途径是通过可重复性的检验。与AAAS报告相通的是,NASEM报告也明确地讨论了可重复性、不确定性和可预测性,其中对可重复性的困难进行了深度的分析,指出了满足可重复性要求的变通途径——累积证据。([2], pp.71-83)

总之,AAAS的报告明确给出了科学探究的三个根本特性:可重复性、不确定性、可预测性。NASEM的报告聚焦研究问题和获得有充分根据的答案,强调证据的确定性是由研究可重复性或更一般意义的证据可累积性担保的。可以明确地说,尽管客观而又确定的关系非常稀少,科学家们仍然千方百计地努力获得它们,而研究可重复性或证据可累积性对建立具有客观确定性的科学事实或科学证据起着关键的作用。

三、累积证据、概念重复与直接重复

1. 可控实验与纯粹观察的概念重复

对一般形式的经验研究而言,取得可累积性的科学证据是阶段性目的。经验研究包括可控实验、纯粹观察、跟踪访谈、问卷调查、随机对照试验等形式。为何证据的可累积性比研究的可重复性更普遍的概念?我们以可控实验和纯粹观察为例来分析。康托罗维奇(Aharon Kantorovich)把所有的科学发现分为两种主要

类型:第一类是创生型科学发现(generational discovery),它分为实验式和理论式科学发现;第二类是揭示型科学发现(discovery by exposure),它分为观察式和推理式的科学发现,前者指纯粹的观察;后者是在已有的发现之上通过推理扩展了此发现的范围。^[6]纯粹观察(例如天文观察)属于揭示型发现,因为它不创生任何新事物;可控实验属于创生型发现,因为它介入和扰动自然而创生出崭新的实体供检测和分析。可控实验的本质是把研究对象从自然中“隔离”出来,操控它,然后使用某种作用(如电力、化合物、药物等)来对它施加影响,创造出新的现象,便于测量和分析。^{[7], [8]}因为整个过程为实验者所控制,所以能够提供完整数据供其他科学家重复。

可重复性概念用于可控实验很准确,却很难准确地用于天文观察。因为在天文观察中,观察者无力控制观察对象的产生,更不能改变它,自然无法重复。一个典型例子是引力波的探测。观察者只能被动地等待引力波作用于探测仪器而产生可测量的信号,他们无法控制引力波出现的时间、引力波对仪器的作用是否足够强而能够探测得到、能否在不同地点探测得到来自同一波源的引力波等等。^[9]所以,不同的探测仪器获得的引力波信号是独立的证据,如果证据多于一个,就知道这种证据有可累积性,却不能说该观察有可重复性。然而,当焦点变成研究问题(“引力波是否存在?”),这些观察所获得的独立证据都肯定地回答了同一个问题,在这个共同问题或概念下,也可以说由这些观察得到的结论有可重复性,即有概念重复性。所以,对于天文观察,证据的可累积性与概念的可重复性是同义的。对于问卷调查,问卷的问题对所有参与对象是相同的(例如“你的研究经验说明可重复性危机是严重的吗?”),但调查的对象各不相同,给出同一答案的的问卷属于同一类证据,当这样的问卷不断回收时,证据就有了可累积性。也可以说对调查问题的答案具有重复性,即问卷调查有概念重复性。

2. 可控实验的直接重复和概念重复

下面我们详细分析可控实验。科学家们对可重复性 (repeatability) 的术语使用没有共识。一些文献偏向使用“可再现性”，即 reproducibility。另一些文献青睐使用“可复制性”，即 replicability。^{[5], [10]} 故此本文把焦点放在定义的内容上，而不论其对应的术语。直接重复聚焦于实验中的某种实际操作（如实验方法、计算程序、结果数据等），对应的是物质性的重复；概念重复关注于实验前的研究问题（基于某个理论、假说、推测等），对应的是答案的重复，即概念性的重复。^[11] 因为每个重复的答案虽然独立，它们却指向一致，是同一类证据，所以证据有可累积性，总称累积证据。

直接重复聚焦于原始实验的具体操作，所以重复实验只能在原始实验所在的领域里进行，它一般定义为与原始实验不同的科学家团队在不同的地点和不同的时间而对原始实验作出完全的复制或部分的重现，即直接重复进一步分为完整复制、计算重现、方法重现、结果重现等等。无论有多少个成功的重复实验，它们只是确认原始实验所发现的科学事实或证据，没有得到新的发现；或者说，它们总共只发现和核实了一个证据。概念重复关注原始实验所回答的问题，所以它能够在与原始实验不同的领域甚至完全无关的学科，例如原始实验是物理学实验，重复实验有化学实验、细胞生物学实验、分子生物学实验等等，但它们对同一个研究问题的答案是一致的。直接重复与原始实验的某个部分有实在的联系，因此直接重复受到原始实验的限制，必须选择与后者相同的领域，或者采用相同的实验方法，或者要求得出相同的实验结果等等。因此，直接重复相对于原始实验的自由度是很小的，导致重复条件严苛，独立性低，难以实现。这能够解释为什么可重复性危机发生在几乎所有科学领域，因为绝大多数实验操作者进行的是某种直接重复。

3. 概念重复能够解决直接重复的危机

更进一步说，概念重复与原始实验在任何阶段的操作都没有实在的联系，它们的唯一共同点是回答了同样的研究问题，所以前者相对于原始实验的独立性更大，因为它对学科选择、

方法取舍、样品材料的使用等都有很大的自由度。又因为所有实验各自独立，但得到对共同研究问题的一致答案，每个答案都是一个证据，因此概念重复导致累积证据。相对于直接重复，概念重复条件宽泛，独立性高，较易实现，这些在复杂性高（如某些生物医学实验）或可控性差（如某些高能物理实验）的条件下尤其如此。因此，NASEM对累积证据或概念重复更重视，把它推荐为可重复性危机的解决方案。（[2]，p.3）科学史上有多个由累积证据来证成一个理论或定律的例子，其中最典型者当数能量转换和守恒定律，它跨越了凡是涉及能量的任何学科，下一节会详细描述。

四、案例研究：能量转换与守恒定律

1. 能量转换现象

能量转换与守恒定律（又译为能量转化与守恒定律，本文不作区分）是整个自然科学领域之中最重要和最基础的理论之一，它也是一个涉及多学科、多领域的理论，支撑它的经验证据也来自多个“大学科”，^[12] 这些学科包括力学、热学、电磁学、化学、声学、光学、生物学和核物理学。本文专注能量转换部分，特别是该定律建构成功之前的实验证据或经验技术。

经典力学是最早发展起来的科学领域。虽然早期的科学家们还没有自觉地为建立能量转换与守恒定律产生实验证据，他们的实验研究却催生了机械能守恒的萌芽。斯梯芬 (Simon Stevin) 对滑轮的研究，伽利略对斜面的研究，惠更斯对碰撞的研究，这些研究中的实验部分为机械能守恒提供了证据。莱布尼兹于1686年提出了活力的概念 (mv^2)，并提出了活力守恒原理。1726年提出的伯努利流体力学方程表达了机械能守恒在流体中的应用。

蒸汽机的发明为热能向机械动能转换提供了一个经验证据。1630–1769年，巴本 (Denis Papin)、萨弗里 (Thomas Savery)、纽可门 (Thomas Newcomen)、瓦特等是蒸汽机的发明家或改进人。伦福德 (Count Rumford) 在

1797年进行的炮管镗孔实验使他得到机械能转化为热能的证据。伏打(Alessandro Volta)在1800年发明了伏打电池,获得化学能转化成电能的证据。奥斯特(Hans Christian Oersted)在1820年发现了磁力与电力能够转换。塞贝克(Thomas J. Seebeck)在同年发现了热电效应,是将热能转换成电能的证据。梅洛尼(Macedonio Melloni)约在1830年发现了热能可以向光能转换。^[7]

后来的多位科学家注意到能量转换与能量守恒的规律性关系。法拉第于1831年发现了电磁感应现象。在1845年,他又发现了磁致旋光现象。他揭示了电能、磁能和光能之间的转换,还明确说过“力”(即能量)可以互相转换并守恒。卡诺(Nicolas Léonard Sadi Carnot)在1832年之前,塞甘(Marc Séguin)在1839年,霍尔茨曼(Karl Holtzmann)在1845年,赫恩(Gustave-Adolphe Hirn)在1854年,分别独立地记录了他们对热和机械功能够转换的定量关系的测量。1837-1844年间,莫尔(Karl F. Mohr)、格罗夫(William Grove)、法拉第和李比希认识到有一种单一的“力”(即能量),它可以有许多形式,并且能够转变,却不可能被创造或毁灭。^[13]由于蒸汽机在工业上的广泛应用,多位科学家对热与机械功的相互转化作出了实验和理论研究,其中有埃瓦特(Peter Ewart)、柯尔丁(Ludwig A. Colding)和赫斯(Germain H. Hess),他们在热力学领域构筑了能量转换与守恒定律的基础。^[7]

2. 能量转换与守恒定律的理论构建

现今的科学史把为全面建成能量转换与守恒定律做出主要工作的三位科学家突出了出来,他们是迈尔(Robert Mayer)、焦耳和亥姆霍兹(Hermann von Helmholtz)。^[13]焦耳在1843-1878年间做了400余次实验,主要是为获得热能与机械功的相互转换的热功当量和建立*i*²*R*定律,后者为电能转化成热能提供了定量的证据。迈尔在1842年完整陈述了能量可以转换但不会消灭的规律,并给出了逻辑上的推理。亥姆霍兹在1847年发表了“力的守恒”,他把能量概念从机械运动推广到所有变化过程,从

而推导出普遍性的能量转化与守恒原理。迈尔和亥姆霍兹的理论构建和焦耳的实验确证,把能量转换与守恒定律的“大模样”描绘出来了。1850年克劳修斯(Rudolf Clausius)提出了该定律的数学形式,汤姆森(William Thomson)在1851年改进了该数学方程,再经过数次概念的调整,就得到了现今教科书上的形式。^[7]

在确立了能量转换与守恒定律以后,跨领域的新证据继续产生。例如,赫兹于1887年意外发现了光电效应,提供了光能转化为电能的证据。^{[14], [15]}哈恩(Otto Hahn)在1938年和斯特拉斯曼(Friedrich W. Strassmann)等人一起发现了核裂变现象,这个发现通向了核能转化为热能的证据。^[16]

3. 能量转换与守恒定律是概念重复或累积证据的范例

库恩称能量转换与守恒定律的发现过程是同时性发现(simultaneous discovery)。^[17]默顿认为多重独立发现(multiple independent discovery)能够确证一个想法或发现。^[18]这些都是从科学史角度对累积证据能够确证一个新科学规律或理论的认识。

由上述史实可以知道,在任何两种能量转换所涉及的多个领域,各个实验证据都是相互独立的,实验方法不同、探测仪器不同、样品材料不同,计算方法不同。如果以为直接重复是实验重复的唯一类型,就无法解释为什么能够获得诸多跨领域能量转换的累积证据,即获得“能量可以转换”这个命题的可概括性或可普及性,从而通向对能量转换与守恒定律的理论构建。从这样的案例中,能够理解为什么NASEM倾向于考察一个命题是否在多个跨领域的情境中有累积性的独立证据。([2], p.3)累积证据或概念重复并没有经历过危机,它为直接重复的危机提供了一个有效解法。

五、累积证据产生过程的数学描述

累积证据的形成有一个前提条件:原始实验和后续实验都围绕一个共同的预测性问题。例如,“不同的能量形式能够互相转换吗?”

就是一个具有预测性的问题。这些参与实验可以在同一领域（例如细胞生物学）但使用不同的实验方法和样品物质，或同一学科（例如生物学）但在不同领域（例如细胞生物学和分子生物学），甚至在不同的学科（例如生物学和物理学），后续实验相对于原始实验的独立性随领域或学科间的“距离”越远越强。

假设共同的研究问题是“如果A作用于任何已知物质 M_i 时，该物质的微观结构会产生显著的变化 c_i （ $i=1, 2, 3\cdots$ ）吗？”。这里的A是某种作用（如激光的照射），微观结构指细胞、分子、原子、核内等结构。变化则需要由可测量的指标来表征，如温度的升降、激光拉曼光谱的变化、癌细胞的杀伤率等。

在原始实验中，甲团队使用了样品物质 M_1 ，观察到其微观结构变化的表征 c_1 ，这是一个新发现，即甲团队取得了一个证据。但孤证不证，甲团队必须有至少另一个支持的经验证据，才能够初步地宣布他们获得的新发现是科学事实，可以用作科学证据，构建假说或检验理论。

根据共同的研究问题和使用相同的作用A，乙团队在不同于甲团队的领域里进行了第二个独立的实验，使用了样品物质 M_2 ，乙团队获得了另一个新发现 c_2 。乙团队肯定地回答了共同的研究问题，得到了独立于原始实验的第二个证据。

也是根据共同的研究问题和使用相同的作用A，丙团队在不同于甲团队或乙团队的学科里进行了独立的实验，使用的样品物质是 M_3 ，得到了一个新证据 c_3 。这是肯定地回答了共同问题的第三个证据。

丁团队也是根据共同的研究问题和使用相同的作用A，在与丙团队相同学科但不同的领域里做了第四个实验，使用了样品物质 M_4 ，获得了全新的证据 c_4 。这个证据也肯定地回答了共同的问题，是第四个证据。

戊团队也是根据共同的研究问题和使用相同的作用A，在与上述实验都不同的学科里进行了第五个可控实验，样品物质是 M_5 ，得到了一个肯定地回答了共同问题的证据 c_5 ，也是第

五个证据。

需要指出的是，上述五个团队分别属于五个不同的科研机构，他们的实验地点各不相同，实验时间也各不相同，实验方法互不相干，检测手段各自无关，分别获得了一个肯定地回答了共同的研究问题的证据，这五个证据（ c_1 、 c_2 、 c_3 、 c_4 、 c_5 ）就合称为相对于共同问题的累积证据。所以，可以确证甲团队的发现是一项新的科学事实。它也能够作为新的科学证据，用来检验有关的假说或理论或用作其他用途。这个共同问题有一定的普遍性，它能够成为一个科学假说的关键部分，有潜力通向理论建构。

从上述的例子可以看到，五个累积证据都一致地符合共同问题中的预测内容。这显示一种客观确定性（接近于客观规律性），它包含客观性，但不只是客观性。上述讨论用简化的形式，支持了NASEM关于“科学是什么”的定义，即科学家们追求的是有一定客观确定性的证据。（[2]，pp.27-33）

由于共同的研究问题中含有“任何已知物质”，表示作用A可能跨越以物质为研究对象的所有学科或领域。所以这个研究问题有可能具有相当广泛的普遍性，只有生成多个跨学科或跨领域的累积证据才能较清楚地显示该普遍性，而直接重复做不到这一点。又因为概念重复使用不同的实验设置来测试一个预测性假设或原始实验的结论，常常得到跨领域的累积证据，所以，概念重复的功能不仅是确认原始实验发现了一个科学事实，而且还有助于开发假说、模型或理论。^[11]

这些都说明了为什么NASEM更偏向累积证据。现今很多科学家只因为直接重复没有成功，就彻底否定原始实验的全部潜在意义，这也可能是目前的重复性危机的症结之一。

六、讨论与结语

当下的可重复性危机是直接重复性的危机，不是概念重复性的危机。因此，概念重复性，也就是累积证据，是解决可重复性危机的一个有效方案。

在解法之外,对可控实验等经验研究的目标问题需要更全面的了解,免得陷入以直接重复性的失败而否定整个研究的“牛角尖”。回顾 NASEM 报告中的第5页:“只注重个别研究的可重复性,并不是确保科学知识可靠性的有效方法。相反,对所研究主题的累积证据加以探究,评估其总体效应的规模和可概括性,才更是一种对科学知识充满信心的有用的方法”。([2], p.5) 该报告的第17页写道:“科学的目标是了解一组科学研究的整体效果或推论,而不是严格确定任何一项研究是否重复了其他任何一项”。([2], p.17)

NASEM 报告的两段话,进一步显示了累积证据在科学研究中的重要性。科学家和关心科学的人对可重复性的讨论,实质上关注的是研究主张的真实确定性。把直接重复当作判断原始实验及其所含研究主张的真实确定性的唯一判据,是不少科学家的习惯思维,是不够科学的。需要增加关注研究主张或问题和能否依此产生累积证据的更加开阔的视角。

作为化解或减轻可重复性危机的可操作性方案,本文建议原始实验的团队需要全面了解他们的实验所回答的问题,特别是了解这个问题是否蕴含一定的普遍性。在他们发表的实验论文里,除了给出直接重复所需要的全部条件、数据和事实,还需要对原始实验的研究问题和实验结果的推论可能含有的普遍性作充分的讨论,并且给出在其他领域甚至不同学科产生累积证据的建议实验的设计原则。也就是说,把关于原始实验结果的推论当作一个具有预测性的假说,建议围绕此推论产生累积证据的实验。这样,能够在直接重复之外,增加产生累积证据的可能性,除了发掘原始实验所含研究主张的普遍性潜力外,还能够避免重复性危机。

[参考文献]

- [1] American Association for the Advancement of Science. 'Science for All Americans'[R]. New York: Oxford University Press, 1990, 1-13.
- [2] National Academies of Science, Engineering, and Medicine. 'Reproducibility and Replicability in Science'[R]. Washington D. C.: Academic Press, 2019.
- [3] Ioannidis, J. P. 'Why Most Published Research Findings Are False?'[J]. *PLoS Medicine*, 2005, 2(8): e124.
- [4] Leek, J. T., Jager, L. R. 'Is Most Published Research Really False?'[J]. *Annual Review of Statistics and Its Application*, 2017, 4: 109-122.
- [5] Baker, M. '1,500 Scientists Lift the Lid on Reproducibility'[J]. *Nature News*, 2016, 533(7604): 452-454.
- [6] Kantorovich, A. *Scientific Discovery: Logic and Tinkering*[M]. Albany, NY: SUNY Press, 1993, 29-35.
- [7] 郭奕玲、沈慧君. 物理学史[M]. 北京: 清华大学出版社, 2005, 41-57.
- [8] 吴彤. 复归科学实践: 一种科学哲学的新反思[M]. 北京: 清华大学出版社, 2010, 58-60.
- [9] 刘见、王刚、胡一鸣等. 首例引力波探测事件 GW150914 与引力波天文学[J]. *科学通报*, 2016, 61(14): 1502-1524.
- [10] Goodman, S. N., Fanelli, D., Ioannidis, J. P. 'What Does Research Reproducibility Mean?'[J]. *Science Translational Medicine*, 2016, 8(341): 1-6.
- [11] Schmidt, S. 'Shall We Really Do It Again? The Powerful Concept of Replication is Neglected in the Social Sciences'[J]. *Review of General Psychology*, 2009, 13(2): 90-100.
- [12] 韦中桑、刘静. 焦耳在能量转化与守恒定律建立过程中的实验创新[J]. *物理通报*, 2023, 42(1): 135-138.
- [13] Cahan, D. 'The Awarding of the Copley Medal and the "Discovery" of the Law of Conservation of Energy: Joule, Mayer and Helmholtz Revisited'[J]. *Notes and Records of the Royal Society*, 2012, 66: 125-139.
- [14] Hertz, H. 'Ueber Einen Einfluss des Ultravioletten Lichtes auf die Electriche Entladung'[J]. *Annalen der Physik*, 1887, 267(8): 983-1000.
- [15] 钱长炎. 赫兹对光电效应的研究及其历史意义[J]. *自然杂志*, 2003, 25(2): 117-122.
- [16] 左维. 迟来的原子核裂变[J]. *物理*, 2015, 44(6): 381-382.
- [17] Kuhn, T. S. 'Energy Conservation as an Example of Simultaneous Discovery'[A], Clagett, M. (Ed.) *Critical Problems in the History of Science*[C], Madison, WI: University of Wisconsin Press, 1959, 321-356.
- [18] Merton, R. K. 'Resistance to the Systematic Study of Multiple Discoveries in Science'[J]. *European Journal of Sociology*, 1963, 4(2): 237-282.