

## 科学研究方法可重复性提升路径及对策建议

叶喜艳<sup>1,2)</sup> 刘蔚<sup>1,2)</sup> 侯春梅<sup>1,2)</sup> 刘宇<sup>3)</sup>  
贺郝钰<sup>1,2)</sup> 马瀚青<sup>1,2)</sup>\*

收稿日期:2023-12-10

修回日期:2024-04-16

- 1) 中国科学院西北生态环境资源研究院, 甘肃省兰州市城关区天水中路8号 730000  
2) 甘肃省知识计算与决策智能重点实验室, 甘肃省兰州市城关区天水中路8号 730000  
3) 清华大学出版社有限公司期刊中心, 北京市海淀区双清路30号 100084

**摘要** 【目的】分析科学研究方法可重复性提升路径,为我国科学研究方法的开放共享与可重复性建设提供借鉴。【方法】主要采用文献调研法、案例分析法、对比分析法和专家咨询法等开展可重复性提升路径案例研究。【结果】可以从研究方法共享的内容、结构化标准化特征、审查程序、检测评估、研究资源认证及其可参与度和可扩展性等方面提升研究方法的可重复性。【结论】建议我国科技期刊开设研究方法专栏专刊,或创办方法类期刊,共享研究方法的所有细节,推进研究方法共享内容的结构化和标准化。建议科技期刊加强与第三方平台的合作,对研究方法进行标准程序审查,加强元数据的自动化检测评估,并进行研究资源认证,同时增强研究方法的可参与度和可扩展性,以促进研究方法的共享与协作,提升科学研究方法的可重复性。此外,建议积极筹备我国可重复性研究机构,启动相应科研项目,逐步完善相关政策和评价机制,让更多的期刊和知识平台积极参与可重复性建设的发展。

**关键词** 研究方法;可重复性;开放共享;检测评估;开放科学

DOI: 10.11946/cjstp.202312100986

随着开放科学的快速发展,科学研究方法的开放共享与管理服务已成为开放科学体系中的重要内容,科学研究方法的可重复性也备受关注。可重复性或者可再现性,即“研究者使用与原研究者相同的材料复制先前研究结果的能力”<sup>[1]</sup>对科学进步至关重要<sup>[2]</sup>。据估计,只有不到50%的临床研究是可重复的<sup>[3-6]</sup>,这种广泛又深入的问题被科学家称为可重复性危机<sup>[7]</sup>。可重复性危机出现的领域从最初的心理學,逐步扩展到实验经济学、癌症医学、生命科学,甚至是人工智能<sup>[8]</sup>、化学<sup>[9]</sup>等领域,可重复性危机已发展为很多学科领域面临的共同问题。

面对可重复性危机,科学哲学家Romero<sup>[10]</sup>提出科学的社会结构应该鼓励一部分人做重复性工作,科学的奖励系统既要奖励新颖发现,也要奖励做出重复性工作的研究者。有科学家建议团队负责人负责检查核实并在论文投稿的补充材料中提供详细的实验信息<sup>[11]</sup>,以提升论文成果的可重复性。也有学

者认为,交叉领域的交流沟通对实验结果的可重复性有一定影响<sup>[12]</sup>,在交叉领域的科学合作中,研究者通过合作来整合专长与技能<sup>[13]</sup>,在这个过程中,也可能存在可重复性危机。2016年,《Nature》上一项针对科研人员的大规模调查发现,超过70%的研究人员不能重复其他科研人员的实验成果<sup>[7]</sup>,也有研究指出高达60%的心理學领域的研究在二次检验中被报告为不可重复<sup>[14]</sup>。心理學和心理语言学领域报道了几个众所周知的失败的可重复性试验<sup>[15-20]</sup>,这种失败试验反映了心理语言学和相关领域不可重复问题的严重程度较高。可重复性也被广泛认为是可靠性评估的一个重要方面<sup>[21]</sup>。然而,如果一项研究未能重复,通常很难准确地找出重复失败的原因。这是因为失败可能由所调查研究问题之外的许多其他因素驱动,如所研究的人群或语言的差异、因变量的自然变异性、实验室设置不同、设备和方案不同可能一起导致非常不同的结果。可重复性危机普遍

**基金项目:** 2023—2024年度中国科学技术期刊编辑学会基金项目“科学数据关联研究方法开放共享与出版研究”(cessp-2023-B-05);中国科学院自然科学期刊编辑研究会研究课题“期刊论文研究方法开放共享与可重复性研究”(YJH202325)。

**作者简介:** 叶喜艳(ORCID: 0000-0001-6466-0428), 硕士, 编辑, E-mail: yexy@lzb.ac.cn; 刘蔚, 博士, 研究员; 侯春梅, 硕士, 编审; 刘宇, 硕士, 编辑; 贺郝钰, 硕士, 副编审。

\*通信作者: 马瀚青(ORCID: 0000-0003-2852-7683), 博士, 编辑, E-mail: mahq@llas.ac.cn。

存在,已成为所有科研相关利益者共同面对的重大问题<sup>[22]</sup>,影响着科学自身的发展以及公众对科学的信任<sup>[23]</sup>。

已有很多期刊关注和出版科学实验和研究方法。比如Nature出版社在于2004年推出*Nature Methods*后,不断扩大以研究方法为中心的期刊组合,如*Nature Protocols*,以及之后的*Protocol Exchange*。爱思唯尔推出了*STAR Protocols*、*Methods*和*Methods X*等。其中,*Methods*侧重于实验生物学和医学科学中迅速发展的技术<sup>[24]</sup>,*Nature Methods*主要发表生命科学领域基础研究的新方法与新技术<sup>[25]</sup>,而*Methods X*接受所有学科的投稿<sup>[26]</sup>,发表研究过程中使用的详细的研究方法,尤其关注创新点,展示如何针对研究课题,对原本宽泛的研究方法进行具体的分析与调整,使其更加具有可操作性并提升研究效率,让研究方法更易被发现、被重现、被引用。因此,越来越多的期刊、资助机构和研究人员开始意识到应共享的不仅仅是代码和数据。在20世纪90年代末,各主体转向使用专门的数据库和代码库,如GitHub、Dryad、Figshare和开放科学框架,而研究方法库一直处于相对缺乏的状态,近年来相关的平台数据库也出现在学术界的视野中,如protocols.io<sup>[27]</sup>、SciScore<sup>[28]</sup>,并且已有较多的研究型期刊与之合作,如*PloS ONE*、*eLife*等,其在投稿指南中鼓励作者将科学实验或者详细的研究方法的链接加入论文。

综上所述,可重复性危机一直是所有科研利益相关者共同面对的重大问题,影响着科学自身的发展以及公众对科学的信任。而随着数据密集型科研范式的兴起和开放科学的发展,科学研究方法的开

放共享与可重复性研究已成为开放科学体系下的重要内容。因此,本文在深度调研国内外文献、管理服务平台和总结前人研究的基础上,进行可重复性提升路径案例分析,以期为我国科学研究方法的开放共享与可重复性提升服务提供借鉴,进一步促进开放科学的发展。

## 1 研究方法与资料来源

主要采用网络和文献调研法、案例分析法、对比分析法和专家咨询法等开展研究。其中,主要利用网络和文献调研法获取研究方法可重复性提升的相关基础资料,了解研究方法期刊发表文章的内容特征和研究方法共享管理平台的平台功能以及提升可重复性的措施。在此基础上,进一步利用案例分析法和对比分析法介绍多个研究方法典型案例平台提升可重复性的实践路径,探索适用于我国可重复性建设的路径方案。采用专家咨询法和线上平台功能演示等有针对性地解决研究中存在的问题,提升研究结果的可借鉴性和可操作性。

研究涉及较多国际平台,主要基于这些平台的公开信息获取研究资料。在调研国内研究方法相关管理平台的时候,发现专门针对科学研究方案方法的管理平台相对较少,更多平台以数据为中心,在与之关联的数据论文或研究论文中介绍研究方法,也即传统的研究方法介绍模式,鲜有专门的科学研究方法共享管理平台。因而主要调研了与研究方法可重复性提升相关的国际平台,以期为我国科学研究可重复性建设提供借鉴和参考。主要研究资料如表1所示。

表1 主要研究对象和研究资料信息

序号	调研对象	网址	相关介绍
1	protocols.io	<a href="https://www.protocols.io/">https://www.protocols.io/</a>	protocols.io是一个用于创建和分享可重复科研方法的安全平台,除了提供实验方案托管服务,该平台还能让研究人员在上面撰写和分享最新的研究方法,以支持更大限度和更高效的合作与创新
2	Codecheck	<a href="https://codecheck.org.uk/">https://codecheck.org.uk/</a>	Codecheck是一个代码审查平台,利用代码审查标准工作流程、操作指南、工具等,支持代码审查员对支撑科学论文的计算机程序进行审查和评估,是提升代码可重复性的主要解决方案之一
3	F1000Research	<a href="https://f1000research.com/">https://f1000research.com/</a>	F1000Research是一个面向所有学科领域研究人员的开放研究出版平台。通过快速发表、透明的发表后同行评议以及Scopus和PubMed索引,提升科研人员的影响力
4	SciScore	<a href="https://sciscore.com/">https://sciscore.com/</a>	SciScore主要针对生物医学领域的研究进行自动化检测评估。SciScore根据论文方法部分包含的受试者性别、知情情况、随机性信息,以及所有研究资源的目录号和研究资源识别码(Research Resource Identifier, RRID)的陈述预期设定,对所有论文进行1~10分的评分
5	RRID	<a href="https://scicrunch.org/resources">https://scicrunch.org/resources</a>	RRID为研究资源持久且唯一的标识符,用于促进对研究资源的识别和跟踪,SciCrunch、NIF(Neuroscience Information Framework)、dkNET等均与其有合作关系
6	MDAR(Materials, Design, Analysis, and Reporting)	<a href="https://sciscore.com/reports/MDAR-Report.php">https://sciscore.com/reports/MDAR-Report.php</a>	MDAR是一份检查清单,给出了对研究方法应该包含的结构化内容的详细要求,主要包括材料、设计、分析3个部分,每一部分又包括诸多元数据要素

续表

序号	调研对象	网址	相关介绍
7	STAR(Structured, Transparent, Accessible Reporting)	<a href="https://www.cell.com/star-methods">https://www.cell.com/star-methods</a>	STAR通过直观、一致的框架促进严谨和稳健的研究,该框架无缝集成到科学信息流中,使作者更容易生成报告,让读者更容易复制研究
8	STAR Protocols	<a href="https://star-protocols.cell.com/search">https://star-protocols.cell.com/search</a>	STAR Protocols是Cell Press的开放获取、同行评议期刊。提供来自生命、健康、地球和物理科学各个领域的结构化、透明、可访问且可重复的分步实验和计算过程。研究方案主要包括摘要、准备工作、关键资源表、材料和设备、分步方法的详细信息、预期结果、局限性、故障调试、参考文献和其他信息
9	SciCrunch	<a href="https://scicrunch.org/">https://scicrunch.org/</a>	SciCrunch是数据共享和展示平台,旨在消除不同科研群体传统门户网站的孤岛现象,促进相互交流,鼓励利用其他人所做的工作并分享他们的专业知识。SciCrunch通过提高现有科学资源的可发现性、可访问性、可视性、实用性和可操作性来增加其价值,研究人员无须对其组件或信息模型进行广泛的重新设计。与更通用的搜索引擎不同,SciCrunch提供集中更多资源的深入访问链接,并为用户量身定制搜索策略
10	dkNET	<a href="https://dknet.org">https://dknet.org</a>	dkNET是一个搜索门户,可帮助研究人员找到与其研究相关的研究资源,并及时了解新工具、服务和指令,以支持稳健且可重复的科学研究。研究资源包括试剂、生物体、软件工具、数据库和服务。dkNET是RRID项目的主要支持者,其目的是确保此类资源在生物医学文献中得到正确标识
11	NIF	<a href="https://neuinfo.org/">https://neuinfo.org/</a>	NIF是美国国立卫生研究院(National Institutes of Health, NIH)蓝图联盟的一项倡议,该联盟将支持神经科学研究的16个NIH研究所、中心和办公室聚集到一个协作框架中,以协调他们正在开展的工作并规划新的项目。NIF提供发现门户、注册表、数据共享服务等
12	Resource Identification Initiative	<a href="https://force11.org/group/resource-identification-initiative">https://force11.org/group/resource-identification-initiative</a>	是FORCE11推出的一个计划,通过使用RRID,促进以机器可读的格式对生物医学领域关键资源的一致引用
13	FORCE11	<a href="https://force11.org/info/about-force11/">https://force11.org/info/about-force11/</a>	FORCE11源于2011年8月在德国达格斯图尔举行的FORCE研讨会,是一个由学者、图书馆员、档案管理员、出版商和研究资助者组成的社区,通过有效利用信息技术来改变现代学术交流,促进知识创新和共享
14	NIH通过提升严谨性和透明性提升可重复性的指导方针	<a href="https://grants.nih.gov/policy/reproducibility/guidance.htm">https://grants.nih.gov/policy/reproducibility/guidance.htm</a>	致力于在各种资助项目的支持下,促进所有科学领域严谨、透明的研究
15	TOP(Transparency and Openness Promotion) Guidelines	<a href="https://www.cos.io/initiatives/top-guidelines">https://www.cos.io/initiatives/top-guidelines</a>	TOP Guidelines是出版和资助领域公认的标准。2015年在Science上发布的TOP Guidelines包括8个模块化标准,每个标准都有3个逐渐严格的级别。期刊选择希望实施的标准,并为每个标准选择实施级别。标准涉及数据引用、材料和代码透明度、预登记、复制等
16	SpringerProtocols	<a href="https://www.springernature.com/gp/librarians/products/databases-solutions/springerprotocols">https://www.springernature.com/gp/librarians/products/databases-solutions/springerprotocols</a>	SpringerProtocols是现代研究实验室的宝贵资源,是生物医学和生命科学领域最大的可重复实验室方案数据库,用户可以在其中找到适合其实验室设置的正确方案。SpringerProtocols为研究人员提供了长达30年的成型的分步实验方案,该方案可立即在实验室中使用。基于分子生物学方法的传承和其他优质资源的内容,研究人员可以确信他们选择的方案是最可靠和最准确的
17	Springer Nature Experiments	<a href="https://www.springernature.com/gp/librarians/products/databases-solutions/experiments">https://www.springernature.com/gp/librarians/products/databases-solutions/experiments</a>	Springer Nature Experiments是Springer Nature开发的一种新的研究解决方案,旨在将研究人员与最相关、最重要的协议和方法联系起来,以支持他们的研究项目。汇集了经证明的可用协议和研究方法,内容主要来自Nature Protocols、SpringerProtocols、Nature Methods和Nature Reviews Methods Primers
18	IEEE DataPort	<a href="https://iee-dataport.org/why-ieee-dataport">https://iee-dataport.org/why-ieee-dataport</a>	IEEE DataPort是一个研究数据平台,旨在存储有价值的研究数据,并使全球研究人员能够访问科学数据。①存储:将研究数据无限期地安全存储到由Amazon Web Services提供支持的单个可信存储库中。②访问:搜索、下载和分析数以千计的数据集,以支持可重复性和创新研究。③管理:制定数据管理计划,跟踪浏览量和引文,并将研究数据与论文和其他工作联系起来。④连接:邀请研究者在数据竞赛中分析数据,联系数据集所有者,并与同行建立联系

## 2 结果分析

### 2.1 研究方法共享内容分析

研究方法是学术论文结构化内容的一部分,但很难在学术论文中被充分地说明,研究方法的开放共享程度也显然不足。研究表明,对现有工作认识的缺乏或负面结果会导致不必要的重复<sup>[29]</sup>,造成不

必要的损失。因此,为了弥补传统以论文形式共享科学研究的不足,共享底层的研究方法和数据显得尤为迫切。如图1所示,科学研究的很多成果是在实验室内部得出的,包括详细的实验记录、数据、代码、模型等,而主要以传统的静态的论文形式进行共享。根据共享的论文并不能获悉实验室内部的研究方法,这不利于研究方法的再利用和可重复性研究。

Buckheit 等<sup>[30]</sup>指出:“科学出版物中那些关于计算科学的论文并不是真正的学术研究成果,它只相当于学术研究的一个广告。真正的学术研究成果是完整的软件开发环境和生成的完整指令集。”因此,开放研究方法的所有细节是提升可重复性、促进科学研究的重要路径。

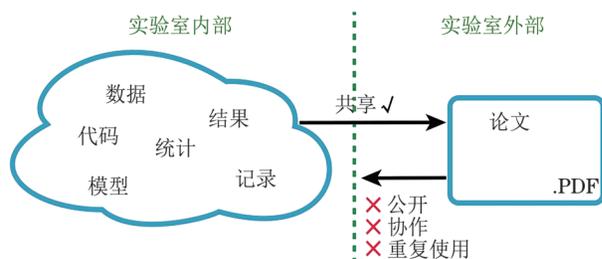


图1 可重复性研究中的逆向问题<sup>[31]</sup>

主要以 protocols.io 平台为例来说明共享研究方法细节的有效路径。protocols.io 是一个用于创建和分享可重复研究方法的安全平台,除了提供实验方案托管服务,该平台还允许研究人员撰写和分享最新的研究方法,以支持更大限度和更高效的合作与创新。protocols.io 允许科研人员创建开放实验社区<sup>[32]</sup>,社区可以整合期刊、评论、科学研究资源(研究方法、研究数据、案例库等)等内容。在具体的研

究方法中, protocols.io 共享所有与之相关的信息,包括研究计划、实验标准、实验记录日期、负责人联系方式、实验室位置等,研究者可根据这些信息找到所需的研究资料,从而高效提升研究方法的可重复性。

## 2.2 研究方法共享内容的结构化、标准化情况分析

在研究方法共享内容的结构化、标准化方面, SciScore 平台是比较典型的案例。该平台针对生命科学领域的期刊论文,给出 MDAR 检查清单<sup>[33]</sup>,列出了对研究方法应该包含的结构化内容的详细要求。如表 2 所示:材料(Materials)部分是保障研究方法可重复的基础,要求给出抗体、细胞系、实验动物、植物和微生物、人类参与者等客体的详细信息,并建立这些关键资源之间的链接;设计(Design)部分主要包括研究方案、实验方案、实验研究设计、样本定义与实验室内复制情况、伦理道德等步骤说明,是具体实验操作的详细指南;分析(Analysis)部分要求给出数据损耗情况、统计情况、数据和代码可用性声明以及遵守的社区标准,主要涉及对实验数据的分析说明以及质量检验,是对实验结果的客观评估。MDAR 制定了核查生命科学研究的最低要求,旨在帮助作者、编辑和其他相关方提高相关研究报告的透明度,进一步为研究方法的可重复提供基础保障。

表2 MDAR结构化项目、内容以及包含的元数据要素与描述

MDAR结构化项目	项目内容	元数据要素与描述
材料部分	主要包含与生物试剂和独特样本相关的信息(如有关人类参与的信息等),以及一些关键资源的信息(如 RRID 等)	①抗体: 供应商名称、目录号、RRID 等信息 ②细胞系: 菌株、登记号、RRID 等信息 ③实验动物: 物种、基因状态、供应商名称、RRID 等信息 ④植物和微生物: 相关的属性描述等信息 ⑤人类参与者: 包括机构声明、知情同意书、参与者的基本信息
设计部分	主要包含与研究实验设计相关的信息,包括方法和统计数据	①研究方案: 尽量提供临床试验研究的试验注册号或 DOI ②实验方案: 尽量提供 DOI(数据库登录号、URL 也可)以获取详细的实验程序 ③实验研究设计: 说明样本量确定、随机性、知情情况、纳入和排除标准等 ④样本定义与实验室内复制情况: 复制发生的次数、复制类型 ⑤伦理道德: 请说明有关道德批准的详细信息(如果适用),包括机构审查委员会的批准、动物护理和使用机构委员会批准、现场样品许可证批准等
分析部分	主要包含与统计分析、代码和研究数据相关的信息	①数据损耗情况: 如果有任何数据被排除在分析之外,请报告 ②统计情况: 描述分析中的每项统计测试结果,并进行 <i>t</i> 检验以计算统计差异 ③数据可用性: 数据可用性声明、公开可用的数据标识符(DOI 或 URL)、可重复使用的公开数据标识符(DOI 或 URL) ④代码可用性: 代码可用性声明、可识别的代码标识符(可被 GitHub、Google Code 和 Bitbucket 识别) ⑤遵守的社区标准: 说明已遵循的指南和清单

此外,由 Cell Press 创建的 STAR 主要包括结构化(Structured)、透明(Transparent)、可获取(Accessible)、报告(Reporting)4个部分,所有 Cell Press 研究文章都具有直观、统一的框架。STAR 提供了一个关键资源的表格框架,旨在提高科学研究的严谨性和可重复性。STAR 确保研究人员报告足够的信息,以便使用相同的关键资源复制研究结果<sup>[34]</sup>。如表 3 所示,

STAR 主要的特点是要求提交的稿件具有标准的结构和顺序,要求列出全面、一目了然的关键资源表,要求在一个专门的部分提供可直接链接到主要资料的完整的方法信息(不限长度)并可在线完整访问,除此之外与其他可重复性倡议保持一致<sup>[35]</sup>。

## 2.3 研究方法程序审查分析

标准程序审查是确保研究方法可重复的有效措

表3 STAR包含要素、内涵以及相关描述

STAR要素	要素内涵	相关描述
结构化	注重细节,包括结构、内容和功能	具有可充分公布的标准、直观的结构;具有全面、一目了然的关键资源表;所有方法信息集成在一个专门的部分
透明	透明、全面和准确的方法可以促进研究人员的工作,应保证方法的可信度	所有Cell Press研究文章都具有直观、统一的框架;稳健且透明的方法可以让其他人在原有工作的基础上继续开展和传播研究
可获取	通过严谨、易于理解、详细且易于遵循的方法,让科研数据变得更容易获取	轻松访问有关试剂和资源的信息;具有完整的方法信息,并可直接链接到主要资料;促进更有效的资源、数据和知识共享
报告	强有力的方法和准确的报告是科学快速发展的基础,STAR致力于帮助研究人员优化科学研究	与作者密切合作,给予研究方法应有的关注;与利益相关者积极合作,以提高方法和报告的标准;继续改进和创新,以增强方法的可重复性

施。代码、数据、模型等是研究方法的核心要素,是标准程序审查的核心。主要以Codecheck为例,阐释代码标准程序审查的具体步骤,以为研究方法的程序审查提供借鉴。

Codecheck是一个代码审查平台,利用代码审查标准工作流程、操作指南、工具等,支持代码审查员对支撑科学论文的计算机程序进行审查和评估,

以提升代码可重复性。从图2可以看出,Codecheck工作流程主要涉及三方人员:(1)提供待审查代码的论文作者;(2)发表作者论文的期刊编辑部;(3)代码审查员。代码审查员只负责审查代码本身,而不负责审查由代码产生的论文内容。Codecheck对代码和数据进行标准程序审查,以提升其可重复性。

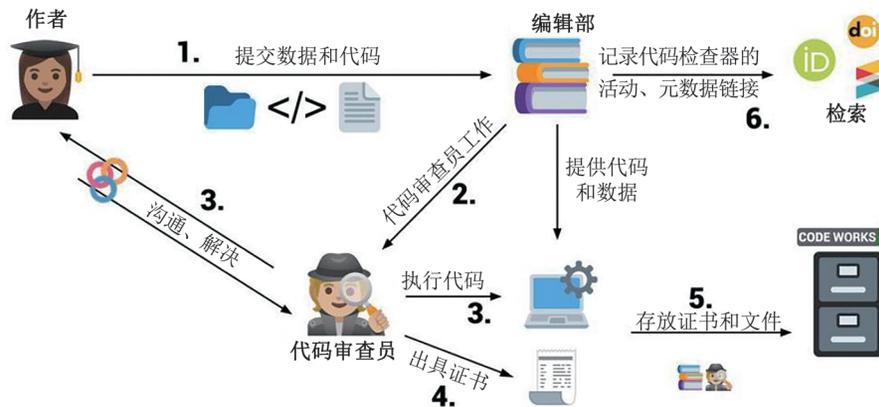


图2 Codecheck基本工作流程<sup>[36]</sup>

#### 2.4 研究方法元数据检测评估分析

对研究方法元数据描述进行检测评估是提升研究方法可靠性和可重复性的有效措施。研究方法元数据描述受学科影响,没有固定的要素,因此需要按学科来制定需要检测的元数据要素。主要以SciScore平台为例说明平台工具如何对学术论文中研究方法的元数据描述进行自动化检测评估,并给出评估报告和评估分值的实现路径。

SciScore是对研究方法进行评审的自动化平台工具,旨在以标准化、易于理解的格式查找并呈现散布在论文中的元数据信息,协助审稿专家对学术质量的评审,但其并不能替代专家评审。目前,SciScore主要针对生物医学领域的研究进行自动化检测评估,主要根据论文研究方法部分包含的受试者性别、知情情况、受试者随机性,以及所有研究资源的目录号和RRID的陈述预期设定,对所有论文进

行1~10分的评分。如果SciScore没有找到元数据要素,例如知情情况,将显示“未检测到”并扣分。如果找到了如“使用雌性老鼠”等包含性别的标准句子,则该句子将在报告中呈现,且分数将会增加。如果SciScore认为某个元数据不适用于该项研究,则会从评估报告中删除对应的分数<sup>[37]</sup>。SciScore还检查多个唯一标识符的准确性,包括稿件中的RRID。完成检测的论文将会得到一份带有编号的SciScore报告,包含评估分值和具体的检测情况。SciScore通过对标准化元数据要素的检测评估,确保数据保持开放和免费使用,强调科学研究的可重复性,也致力于提升公众对科学研究的信任。

#### 2.5 研究方法资源认证分析

将研究方法作为单独的研究资源进行开发利用,需要赋予其作为学术成果的研究资源资格属性,进行统一的标准化认证。比如借鉴学术论文的认证

方法,给研究方法分配独一无二的DOI,或者建立国际认可的研究资源认证体系,通过认证的研究资源将会被分配一个标识符,如给研究方法分配RRID。RRID为研究资源持久且唯一的标识符,用于促进对研究资源的识别和跟踪,SciCrunch、NIF和dkNET均与其有合作关系。RRID门户网站集成许多数据存储库,如MGI、Addgene、MMRRC、Cellosaurus。如果作者使用其他学者的资源,可以用RRID标注引用;如果作者自己创建了新的资源,可利用RRID将其存放在数据库中的具体位置。RRID可独立于研究资源本身而持续存在,提供了统一的引用和获取方式,被科学界许多参与可重复性建设的学者采用。可见,对研究方法进行资源认证,有利于其进行广泛传播和再利用,是提升透明度、可重复性和识别度的有效实用的方法<sup>[36]</sup>。

## 2.6 研究方法可参与度和可扩展性分析

提升研究方法的参与度、可扩展性是提升其可重复性的有效路径。由前文可知,在开放科学生态系统中,研究方法共享平台已经在内容共享、程序审查、评估检测、研究资源认证、结构化标准化等方面取得重大成效。而在对上述内容的调研中发现,每一种研究方法可重复性提升路径都需要多方参与。“研究方法可重复”本身就要求研究方法开放共享并且可以获得再利用,其他研究人员再利用是验证研究方法可重复的基本前提。因此,在研究方法共享的过程中,通过设置交互窗口,提供开放版本编辑、评论、历史记录保存等可扩展功能,可提升学界研究人员的参与度以及研究方法本身的可扩展性。

2023年7月,世界领先的实验方案出版商Springer

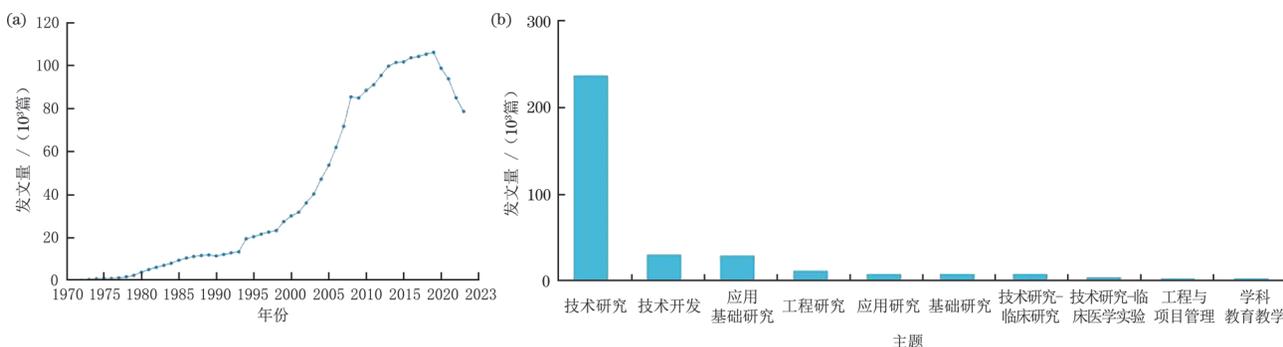


图3 中文期刊相关文献统计结果  
(a) 发文量年度趋势; (b) 研究主题分布

可重复性研究是开放科学的重要内容,然而这方面的研究却没有得到足够的重视,甚至在某些研究领域被忽视。国内缺乏专门从事相关研究的组织机构,政策引导也相对欠缺,研究者的初步成果

Nature收购了用于开发和共享可重复研究方法的安全平台protocols.io<sup>[32]</sup>。protocols.io作为协议和方法领域的创新者,在传统协议和方法工作流程之外为科学合作提供了一个安全的平台,为研究人员提供了在平台内编写和共享方法的功能,促进了多人的高效协作。protocols.io首席执行官兼联合创始人Lenny Teytelman认为,与Springer Nature合作将进一步促进对研究方法的讨论、修订、优化和重用,进一步扩展其产品,使得科学家及其研究从动态交互式协议方案中进一步受益<sup>[32]</sup>。此外,IEEE DataPort通过举办数据竞赛邀请研究者分析数据、联系作者、提供反馈、主动分享等,使得研究数据得到有效利用。

## 3 讨论

在对国内平台的调研中未发现典型的研究方法管理平台,更多平台以数据为中心,在与之关联的数据论文或研究论文中介绍研究方法。在中国知网中对相关中文文献进行检索和不完全统计(截至2024年4月16日),发现:侧重方法的且还在正常出版的期刊数量很少,可以检索到的不足10种,而包含方法的论文集或者摘要集十分常见;篇名中包含“方法”的论文数量在2019年以前一直呈明显增加的趋势(自2020年以来有所下降,可能与英文期刊的发展和稿源外流有关),但其主要关注技术研究与技术开发,其他领域的方法研究较少,如图3所示。虽然技术类方法论文数量比较多,但对应的方法类期刊数量未见明显增多。因此,目前我国科技期刊对研究方法及其再利用还不够重视。

很难得到认可。因此,如何加强可重复性研究是亟待解决的问题。有学者从出版流程与出版导向两个维度梳理分析了国际科技期刊出版商、学/协会等利益主体面对可重复性危机时提出的保障措施

及实践进展,提出加快论文关联数据汇交、创新学术出版模式、完善科技评价体系等措施<sup>[38]</sup>,以应对可重复性危机;也有学者针对论文数据可重复性危机制定了解决方案,使论文与其背后的数据、代码等科研要素有效关联,以支持研究结果的验证<sup>[39]</sup>。这与本研究的目的一致,即促进开放科学背景下的可重复性建设。本文主要以研究方法为切入点,从研究方法开放共享的角度,提出提升研究方法可重复性的路径措施,并结合国内现状给出优化建议。

## 4 建议与启示

### 4.1 共享研究方法的所有细节

目前,国内专门针对研究方法的管理平台较少,更多平台以数据为中心,在与之关联的数据论文或研究论文中介绍研究方法,这就使得研究方法很难被连贯详细地集中介绍。因此,建议相关平台尽可能详细地共享研究方法的所有细节,有效促进研究方法的再利用和可重复性研究。建议我国科技期刊开设研究方法专栏专刊,创办方法类期刊,为研究方法可重复性研究提供发表平台。

### 4.2 促进研究方法共享内容的结构化与标准化

促进研究方法共享内容的结构化、标准化有助于提高研究方法的可重复性、透明度、可解释性和质量控制水平,同时加速知识传播和应用,推动科学研究的进步和发展。建议国内出版平台推出研究方法开放共享平台,提出结构化、标准化方案,给出研究方法的结构化内容以及标准化的程序步骤。同一学科领域方法类期刊给出研究方法内容撰写的结构化、标准化模板,以使研究方法在行业内更容易可重复利用,更规范地在不同研究项目之间共享协作。

### 4.3 进行研究方法标准程序审查

标准程序审查是确保科学研究方法正确和可靠的重要步骤。建议我国科技期刊平台与研究方法审查平台进行合作,对方法类文章进行结构化和标准化程序审查,确保研究方法符合模板规范,符合科学原则和研究目标,对不符合规范的方法类文章进行预警提示,并退回作者修改,以提升科学研究的可靠性和研究人员对科学成果的信任度。

### 4.4 对研究方法元数据进行自动化检测评估

建议科技期刊平台对研究方法元数据进行自动化监测评估,以标准化、易于理解的格式查找并呈现散布在论文中的研究方法元数据信息,对研究方法

元数据进行自动化检测评估,并给出分值。一方面,可确保文章中研究方法基本信息的完整性;另一方面,可帮助研究人员快速评估研究方法的质量,判断其是否具有可重复性。

### 4.5 对研究方法进行研究资源认证

对研究资源的标准化认证有利于其广泛传播、再利用、被引用。因此,要将研究方法作为单独的研究资源进行开发利用,需要赋予其作为学术成果的研究资源属性,以便在使用过程中按统一的格式进行重复利用。建议国内出版平台在推出研究方法管理平台的基础上,基于国际认可的研究资源认证体系对研究方法进行研究资源认证,分配唯一标识符,以认可研究者的学术成果,并将其纳入评价体系。

### 4.6 提升研究方法的参与度和可扩展性

建议科技期刊出版平台增强方法类文章的参与度和可扩展性,以有效提升研究方法的可重复性。可以通过提供详细的方法描述和操作指南、建立在线评论社区与协作平台、建立反馈和改进机制、整合研究资源等途径吸引更多的参与者,提升围绕研究方法的其他资源的扩展性能,从而促进研究方法的共享。

### 4.7 加强中国可重复性研究相关平台建设

可重复性研究是开放科学的重要内容,然而这方面的研究却没有得到足够的重视,甚至在某些研究领域被忽视。国内缺乏专门从事相关研究的组织机构,政策引导也相对欠缺,研究者的初步成果很难得到评价机制的认可。因此,建议积极筹备组建相关研究机构,启动相应的科研项目,逐步完善相关政策和评价机制,让更多的期刊和知识平台积极参与可重复性研究的发展。

## 5 结语

本研究调研研究方法相关数字平台,重点调研了 protocols.io、Codecheck、SciScore、RRID、Springer Nature Experiments、IEEE DataPort 等门户网站,开展研究方法可重复性提升路径研究,主要从研究方法共享的内容、结构化标准化特征、程序审查、检测评估、研究资源认证及其可参与度和可扩展性等方面进行分析,并有针对性地提出建议。由于调研的学科领域主要为生命科学或生物医学领域,而研究方法元数据要素受学科影响较大,本研究选取的案例难免具有局限性。下一步计划在研究方法可重复性提升路径研究的基础上,进行可重复性评价指标的

研究。另外,可重复性与学术诚信密切相关,因此也将继续关注其与学术诚信的关系研究。

**致谢** protocols.io平台的战略计划总监Dr. Emma Ganley在线一对一演示了protocols.io平台的部分功能,并提供了相关的培训,给文章的撰写提供了重要的思路,在此表示诚挚的感谢!

## 参考文献

- [ 1 ] Goodman S N, Fanelli D, Ioannidis J P A. What does research reproducibility mean? [ J ]. *Science Translational Medicine*, 2016, 8( 341 ): 341ps12.
- [ 2 ] Hughes B T, Niemann A, Tritz D, et al. Transparent and reproducible research practices in the surgical literature[ J ]. *Journal of Surgical Research*, 2022, 274: 116–124.
- [ 3 ] Downing S M. Reliability: On the reproducibility of assessment data[ J ]. *Medical Education*, 2004, 38( 9 ): 1006–1012.
- [ 4 ] Chalmers I, Glasziou P. Avoidable waste in the production and reporting of research evidence[ J ]. *The Lancet*, 2009, 374( 9683 ): 86–89.
- [ 5 ] Nosek B A, Ebersole C R, DeHaven A C, et al. The preregistration revolution[ J ]. *Proceedings of the National Academy of Sciences of the United States of America*, 2018, 115( 11 ): 2600–2606.
- [ 6 ] Kerr N L. HARKing: Hypothesizing after the results are known[ J ]. *Personality and Social Psychology Review*, 1998, 2( 3 ): 196–217.
- [ 7 ] Baker M. 1, 500 scientists lift the lid on reproducibility[ J ]. *Nature*, 2016, 533( 7604 ): 452–454.
- [ 8 ] Leichtmann B, Nitsch V, Mara M. Crisis ahead? Why human-robot interaction user studies may have replicability problems and directions for improvement[ J ]. *Frontiers in Robotics and AI*, 2022, 9: 838116.
- [ 9 ] Biagioli M. Before and after photoshop: Recursive fraud in the age of digital reproducibility[ J ]. *Angewandte Chemie International Edition*, 2019, 58( 46 ): 16334–16335.
- [ 10 ] Romero F. The division of replication labor[ J ]. *Philosophy of Science*, 2020, 87( 5 ): 1014–1025.
- [ 11 ] Bergman R G, Danheiser R L. Reproducibility in chemical research[ J ]. *Angewandte Chemie International Edition*, 2016, 55( 41 ): 12548–12549.
- [ 12 ] Hensel W M. Double trouble? The communication dimension of the reproducibility crisis in experimental psychology and neuroscience[ J ]. *European Journal for Philosophy of Science*, 2020, 10( 3 ): 44.
- [ 13 ] 朱晶. 认知劳动分工视角下的科学合作与集体知识[ J ]. *哲学动态*, 2020( 3 ): 111–118.
- [ 14 ] Open Science Collaboration. Estimating the reproducibility of psychological science[ J ]. *Science*, 2015, 349( 6251 ): eaac4716.
- [ 15 ] Hagger M S, Chatzisarantis N L D, Alberts H, et al. A multilab preregistered replication of the ego-depletion effect[ J ]. *Perspectives on Psychological Science*, 2016, 11( 4 ): 546–573.
- [ 16 ] Jäger L A, Mertzen D, van Dyke J A, et al. Interference patterns in subject-verb agreement and reflexives revisited: A large-sample study[ J ]. *Journal of Memory and Language*, 2020, 111: 104063.
- [ 17 ] Klein R A, Vianello M, Hasselman F, et al. Many labs 2: Investigating variation in replicability across samples and settings[ J ]. *Advances in Methods and Practices in Psychological Science*, 2018, 1: 443–490.
- [ 18 ] Nieuwland M S, Politzer-Ahles S, Heyselaar E, et al. Large-scale replication study reveals a limit on probabilistic prediction in language comprehension[ J ]. *eLife*, 2018, 7: 33468.
- [ 19 ] Vasishth S, Mertzen D, Jäger L A, et al. The statistical significance filter leads to overoptimistic expectations of replicability[ J ]. *Journal of Memory and Language*, 2018, 103: 151–175.
- [ 20 ] Stack C M H, James A N, Watson D G. A failure to replicate rapid syntactic adaptation in comprehension[ J ]. *Memory & Cognition*, 2018, 46( 6 ): 864–877.
- [ 21 ] Laurinavichyute A, Yadav H, Vasishth S. Share the code, not just the data: A case study of the reproducibility of articles published in the *Journal of Memory and Language* under the open data policy[ J ]. *Journal of Memory and Language*, 2022, 125: 104332.
- [ 22 ] 孔祥辉. 欧美研究型大学图书馆可重复性专题服务调查研究[ J ]. *图书馆工作与研究*, 2022( 11 ): 37–45.
- [ 23 ] 朱晶. 科学实践与合作网络中的可重复性问题[ J ]. *自然辩证法研究*, 2022, 38( 11 ): 30–36.
- [ 24 ] *Methods*: About the journal[ EB/OL ]. [ 2023–05–27 ]. <https://www.sciencedirect.com/journal/methods>.
- [ 25 ] *Nature Methods*: About the journal[ EB/OL ]. [ 2023–05–27 ]. <https://www.nature.com/nmeth/aims>.
- [ 26 ] *Methods X*: About the journal[ EB/OL ]. [ 2023–05–27 ]. <https://www.sciencedirect.com/journal/methodsx>.
- [ 27 ] Bring structure to your research: A secure platform for developing and sharing reproducible methods[ EB/OL ]. [ 2023–05–27 ]. <https://www.protocols.io/features>.
- [ 28 ] SciScore™ is the best methods review tool for scientific articles[ EB/OL ]. [ 2023–11–27 ]. <https://sciscore.com/>.
- [ 29 ] Cost-benefit analysis for FAIR research data—Cost of not having FAIR research data[ EB/OL ]. [ 2023–05–27 ]. <https://data.europa.eu/doi/10.2777/02999>.
- [ 30 ] Buckheit J B, Donoho D L. WaveLab and reproducible research[ M ]//Antoniadis A, Oppenheim G. Wavelets and Statistics. New York: Springer, 1995: 55–81.
- [ 31 ] Nüst D, Eglen S J. Codecheck: An open science initiative for the independent execution of computations underlying research articles during peer review to improve reproducibility[ J ]. *F1000Research*, 2021, 10: 253.
- [ 32 ] Springer Nature continues open research drive with acquisition of protocols.io[ EB/OL ]. [ 2023–05–27 ]. <https://group.springernature.com/gp/group/media/press-releases/publisher-acquires-protocolsio-driving-open-research/25844774>.

- [ 33 ] MDAR-report[ EB/OL ]. [ 2023-05-27 ]. <https://sciscore.com/reports/MDAR-Report.php>.
- [ 34 ] STAR-report[ EB/OL ]. [ 2023-05-27 ]. <https://sciscore.com/reports/STAR-Report.php>.
- [ 35 ] STAR-methods[ EB/OL ]. [ 2023-05-27 ]. <https://www.cell.com/star-methods>.
- [ 36 ] About RRID[ EB/OL ]. [ 2023-05-27 ]. <https://dknet.org/about/rrid>.
- [ 37 ] SciScore report[ EB/OL ]. [ 2023-05-27 ]. <https://www.scicrunch.com/sciscorereport-faq>.
- [ 38 ] 陈序文, 姚长青, 雷雪. 学术出版视角下科研成果可重复性保障机制研究 [ J ]. 中国科技期刊研究, 2023, 34( 11 ): 1451-1457.
- [ 39 ] 孟美任, 彭希琚. 论文数据可重复性问题剖析及一种解决方案 [ J ]. 中国科技期刊研究, 2021, 32( 3 ): 397-402.

#### 作者贡献声明:

叶喜艳: 策划选题, 构思框架, 撰写和修订论文;  
刘蔚, 侯春梅: 指导研究;  
刘宇: 调研国内研究现状, 完成调研报告;  
贺郝钰: 修订论文;  
马瀚青: 完善研究思路, 修订论文。

## Paths and countermeasures for improving reproducibility of scientific research methods

YE Xiyang<sup>1,2)</sup>, LIU Wei<sup>1,2)</sup>, HOU Chunmei<sup>1,2)</sup>, LIU Yu<sup>3)</sup>, HE Haoyu<sup>1,2)</sup>, MA Hanqing<sup>1,2)\*</sup>

- 1) Northwest Institute of Eco-Environment and Resources, Chinese Academy of Sciences, 8 Middle Tianshui Road, Chengguan District, Lanzhou 730000, China
- 2) Key Laboratory of Knowledge Computing and Intelligent Decision, 8 Middle Tianshui Road, Chengguan District, Lanzhou 730000, China
- 3) Journal Publishing Center of Tsinghua University Press Co., Ltd., 30 Shuangqing Road, Haidian District, Beijing 100084, China

**Abstract:** [ **Purposes** ] This study aims to analyze the paths of improving the reproducibility of scientific research methods and provide a reference for the construction of open sharing and reproducibility improvement of scientific research methods in China. [ **Methods** ] We mainly used literature research, case analysis, comparative analysis, and expert consultation to conduct case studies on the paths of improving reproducibility. [ **Findings** ] The reproducibility of research methods can be improved in terms of shared content, structured and normalized features, review procedures, detection and evaluation, research resource certification, participation, extendibility, etc. [ **Conclusions** ] It is recommended that Chinese scientific journals establish column special issues on research methods or establish methodological journals to share all the details of research methods and promote the structuring and normalization of shared content of research methods. It is recommended that scientific journals should strengthen their cooperation with third-party platforms, conduct standard procedural reviews of research methods, enhance automatic detection and evaluation of metadata, and conduct research resource certification. At the same time, scientific journals should also enhance the participation and extendibility of research methods to promote the sharing and collaboration of research methods and improve the reproducibility of scientific research methods. In addition, it is suggested that China should actively prepare reproducibility research institutions, initiate corresponding scientific research projects, gradually improve relevant policies and evaluation mechanisms, and encourage more journals and knowledge platforms to actively participate in the development of reproducibility construction.

**Keywords:** Research method; Reproducibility; Open sharing; Detection and evaluation; Open science

(本文责编: 杨恒)