

全学科的撤稿原因分析

——基于心理学领域

康存义^{1,2} 王子衿^{1,2} 王文轩^{2,3}

1. 北京师范大学珠海分校, 珠海;

2. 心理学画廊计划课题组, 平衡科学与交叉学科标准化研究交互体系研究所, 北京;

3. 埃森哲大中华区柔性工作室, 上海

摘要 | 目的: 高质量的科学研究结果需要基于可追溯的科学研究方法, 一旦科学研究方法出现问题, 科学研究的结果就不再可靠。当前心理学研究存在一系列由主观和客观原因造成的方法不规范的问题, 这破坏了实证研究的可靠性, 并进一步引发文章的勘误和撤稿。本文从心理学的研究过程出发, 为避免不规范现象提供研究建议, 以期讨论不规范的学术研究对学术界的影响。方法: 研究收集了Retraction Watch Database中领域字段为心理学的1034条关注、勘误和撤稿声明, 从文献的撤稿原因、期刊来源、出版商来源、发表时间、撤稿时间、发表与撤稿的时滞、学科交叉、声明类型、文献类型、发表国家、作者规模等维度对心理学领域撤稿文章特点展开基于前述已构建的心理学研究框架的讨论。结果: 研究发现, 在1943年到2021年发表文献的撤稿声明中, 心理学领域最普遍的撤稿原因是有关数据的顾虑和问题、抄袭和重复发表; 撤稿文献涵盖领域常与心理学共现的有教育学、社会学、医学下的精神病学、商学下的管理学等; 撤稿涉及最多的期刊为 *Perceptual and Motor Skills* 和 *Psychological Reports*, 最多的会议为上海2011年举办的ICEE, 最多的出版商为IEEE, 其次分别为SAGE、Wiley, 最多的国家(或地区)为中国、美国、俄罗斯和英国; 撤稿时滞最大为78年, 最小为0天即当日撤稿, 平均时滞为5.6年。结论: 应加强心理学研究的规范化, 从源头尽可能地避免学术原因导致的撤稿现象, 这有利于心理学领域的稳定发展。

关键词 | 心理学研究框架; 撤稿; 科学计量

Copyright © 2023 by author (s) and SciScan Publishing Limited

This article is licensed under a [Creative Commons Attribution-NonCommercial 4.0 International License](https://creativecommons.org/licenses/by-nc/4.0/). <https://creativecommons.org/licenses/by-nc/4.0/>



1 引言

1.1 撤稿文章的影响

学术不端导致的撤稿行为会对领域内后续开展的研究产生众多不利影响。期刊发表撤稿声明, 是对

问题文章造成影响的补救措施,但问题文章引发的不利影响难以根除。例如,论文撤稿后依然会被继续引用,即使是由于数据造假这一学术不端原因被撤稿,大多文献在对撤稿论文的引用中均为正面(Kochan and Budd, 1992),这体现出被撤稿文章依然对后续研究产生误导,由于撤稿具有时间上的滞后,在撤稿前的这一段时间,也具备被其他研究引用的可能,在撤稿发生后,影响力并不会完全消失,依然有一定的被引量。同时,非剽窃或出版商错误的撤稿会影响到本领域研究的活跃度,相关研究被引量、发文量、资金资助程度均会降低(Furman et al., 2012),从而造成研究者回避或转移到其他方向的现象,也会牵连到无辜的相关研究者,降低科研基金获取的可能。

学术界论文产量飞速增长,即使撤回文章只占总发文量的0.02%左右,论文撤稿总量也已经达到了巨大的规模(Wager and Williams, 2011),撤稿研究也成为重要的研究问题。Retraction Watch Database于2010年成立,致力于对世界范围内撤稿文章的收录,以便减少撤稿文章对后续研究的影响。但在此之前,大部分撤稿领域的研究主要通过Web of Science、Scopus等文献数据库和期刊官网以“retraction”相关撤稿关键词进行领域针对性检索(Budd, 1999; Cokol, 2008),但由于不同数据库收录文献的数据架构存在一定的差异,导致横向比对过程中存在诸多局限。撤稿数据库成立后,大大节省了撤稿领域研究者的工作量,出现了较多以数据库为主要原始数据来源的研究(王景周, 2021; 邓支青、吴任力, 2021),同时也为读者了解撤稿文献的详细信息提供了查询平台。

1.2 撤稿指标及代表文章

学者文章撤稿的相关指标较多,基于期刊影响因子、电子邮箱、国家经济制度差异、总发文量、被引量、领域差异等因素均有学者做出了分析(本研究涉及33个表格,均可在补充材料<https://k1ax0er2lc.feishu.cn/docx/P5P9dwFKxouoSHxTgFFc54yKnld>获得)。

学术不端事件由来已久,1756年,第一篇有记载的科学论文被撤回(Oransky, 2015),即便伽利略、牛顿、伯努利等知名科学家也时有不端(或疑似)行为被曝出。心理学领域同样存在许多被撤稿的研究论文。早在1943年,心理学家伯特(Burt)通过伪造数据用以论证人类智力的先天遗传比例达到75%。1990年,Retraction Watch Database纳入了目前发表最早(1987)的心理学领域撤稿文章。随后,施塔珀尔(Stapel)、艾森克(Eysenck)分别于2011年、2020年被曝大规模数据造假,对心理学领域造成了重大影响。

1.3 心理学研究的一般思路

心理学作为依托定量和定性两大社会科学基本范式、研究人类心理现象和行为的学科,自冯特1879年创建心理学第一个实验室开始,已经有140多年的历史。随着心理学研究领域广度和深度的不断拓展,研究要面对的抽象概念大量增加,而由于抽象概念无法直接观测的特性,需要学者针对抽象概念进行操作性定义,随后对抽象概念进行分类、维度划分、实验范式的开发等诸多操作,该流程中的每个环节,都会对抽象概念的代表性产生削减,研究过程中上游环节的失误或不准确,会对下游带来巨大的危害,这一特点最终体现在发表的论文结构上。本研究着重关注撤稿原因所对应的文章结构对心理学研究进行探讨,如图1所示。

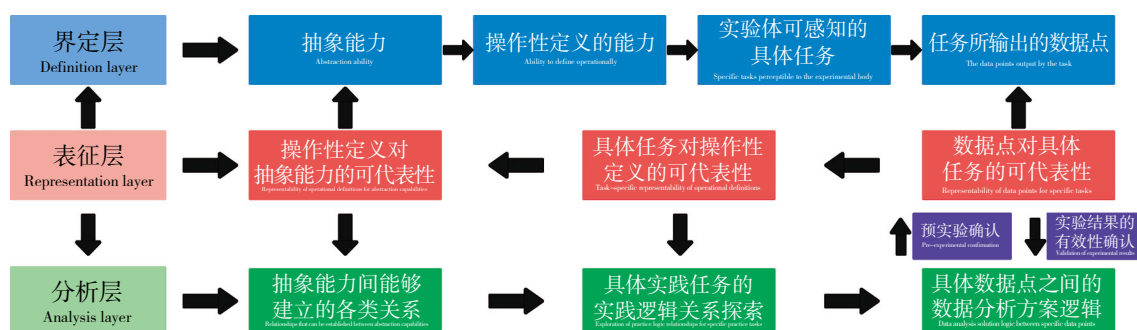


图 1 心理学研究开展的一般框架

Figure 1 General framework for conducting psychological research

本研究提出了一个心理学研究的一般框架，心理学研究可以分为界定层、表征层和分析层三个部分。研究过程将界定层作为标准的心理学研究开展流程，通过表征层的逆向检验和分析层的实际数据验证过程，提升心理学研究的科学性。（1）界定层。界定层针对心理学某一抽象能力展开研究，但由于抽象能力无法直接进行测量，需要对其进行操作性定义，以便进一步确定人可以感知到的具体实验任务，借助开发好的实验任务进行施测，获取供研究分析的数据点。（2）表征层。表征层通过验证界定层环节的有效性，即需要保证收集到的数据点对具体任务的代表性，以支撑具体任务对操作性定义的代表性，最终保证操作性定义对抽象能力的代表性。（3）分析层。同时，表征层作为连接界定层和分析层的桥梁，其有效性也需要通过分析层检验。分析层作为表征层的下游，对表征层有效性的验证需要研究者展开预实验，对研究进行调整完善后再次施测，根据预实验结果的有效性进行确认。其中，数据点对具体任务的代表性，需要通过建立有效性的具体数据点之间数据分析方案逻辑来实现；具体任务对操作性定义的代表性，需要通过探索具体实践任务的实践逻辑关系实现；操作性定义对抽象能力的代表性，需要通过探索抽象能力的各类关系实现。

2 方法

2.1 撤稿原因

2.1.1 以往研究对撤稿原因的分类

撤稿有多种分类标准（Azoulay, 2012; Bilbrey et al., 2014），众多学者或机构团体均对撤稿原因做出过详细的划分（Budd et al., 1998; Nath et al., 2006），但以上的分类都体现出以下特点：（1）引发撤稿的不当行为较难界定是否为主观恶意的学术不端，还是无心的错误，因此只将其归为难以界定；（2）以往的研究对撤稿原因的分类，虽然部分涉及了原因与文章结构的对应，但未明确根据此维度对原因进行分析；（3）以往学者大多从计量出发，侧重于数据挖掘导向对研究的末端，即在期刊上发表文章后撤稿的直接原因进行分析，较少针对在研究实验设计初期、文章撰写前规避撤稿风险的讨论；（4）尚无依据心理学一般研究框架所建立的一套撤稿原因分类标准。因此，本研究在心理学研究框架的基础上，将撤稿原因进行分类并一一对应，为心理学研究者更科学地展开研究提供建议与帮助，减少心理学领域的撤稿现象。

2.1.2 心理学撤稿原因分类

截至 2022 年 2 月 1 日,伊万·奥兰斯基 (Ivan Oransky) 和亚当·马库斯 (Adam Marcus) 创建的 Retraction Watch 网站上罗列出了 104 条撤稿原因,本研究对撤稿原因的分析有三个阶段,第一阶段参考吴任力等 (2020) 的分类方式,将撤稿原因分为未知或信息不全、沟通问题、程序问题、法律与伦理问题、实验材料污染、研究错误、学术不端、研究疑虑与争议、研究客观性与真实性、出版原因、重复问题、研究抄袭、不可靠问题、研究外原因及其他难以分类的原因,如图 2 所示。

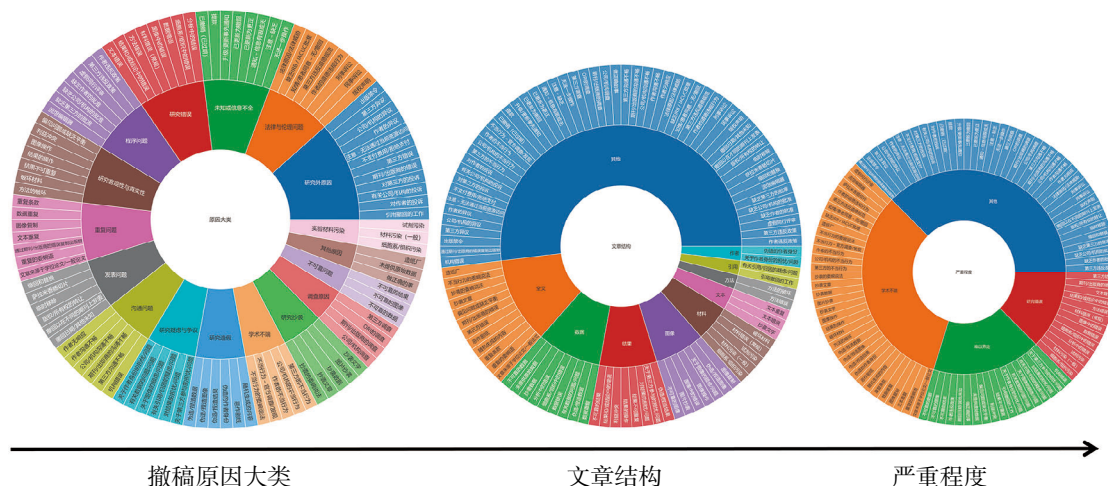


图 2 三种撤稿原因划分方式

Figure 2 Three types of reasons for withdrawal

在第二阶段,对有关学术研究直接相关的撤稿原因进一步依据心理学研究写作过程展开划分,设置全文、引用、作者、数据、图像、结果、材料、方法。在第三阶段,将第一阶段整理的撤稿原因依据严重程度进一步划分为学术不端、研究错误、难以界定及其他,如图 2 所示。

2.2 检索策略

研究数据通过撤稿观察数据库 (Retraction Watch Database) 获得,限定学科为社会科学分类下的心理学 (Psychology),检索日期为 2022 年 2 月 1 日,得到 1943/5/1—2021/8/11 之间发表的 1034 条声明结果。其中撤稿文章 889 条目,关注文章 100 条目,勘误 43 条目,2 条未说明 (经人工检查后,1 条为撤稿文章条目,1 条为关注条目)。2 条未说明 (经人工检查后,1 条为撤稿文章条目,1 条为关注条目)。由于研究立足于心理学研究的规范化,因此研究对象包括声明的撤稿、关注和勘误三个类型,而非大多数先前研究只纳入撤稿条目进行分析。

2.3 撤稿数据字段

研究提取了文献题录的标题、领域、期刊、出版商、机构、撤稿观察报道链接、撤稿原因、作者、原始文章发表日期、原始文章 PubMedID、原始文章 DOI、撤稿或其他声明日期、撤稿或其他声明 PubMedID、撤稿或其他声明 DOI、文章类型、声明类型、国家、期刊是否收费,字段及描述。

针对撤稿原因归属心理学模块的变量编码,由 3 位心理学专业研究者进行编码,可以清晰界定为学术部分为 1,可以清晰界定为非学术部分为 0,难以界定为 0.5;可以清晰界定为无意错误、有意学术不端分别编码为 1、2,难以界定编码为 3,其他编码为 4;针对心理学研究框架,分别对应研究的 1 ~ 6 个步骤;针对心理学文献报告框架,1 ~ 4 分别对应引言、方法、结果和讨论部分。

2.4 分析方法

研究首先对纳入的数据借助 Jamovi (1.6.18) 进行描述性统计,得到字段的频次、占比、年份变化情况;对于标题长度、国家及作者数量、文章类型、声明类型等变量,展开相关分析;领域、作者、国家、撤稿原因字段,利用 Bibexcel (2016.2.20) 软件生成共现矩阵,输出共现网络数据文件,导入 Vosviewer (1.6.17) 软件制作共现图谱,并对图谱进行聚类着色、调整布局。在对纳入的大批量数据进行定量分析后,根据分析的结果筛选出有代表性的撤稿文献展开深入的定性讨论。

3 研究结果

在 1034 条撤稿声明中,包含 890 条撤稿,101 条关注,43 条勘误。发表时间为 1943—2021 年,如图 3 所示。其中,原始论文发表数量在 2010 年和 2011 年达到顶峰,分别为 130 和 132 篇,2015 年达到第二个高峰 103 篇,此后发表的撤稿文献量开始有所下降。撤稿声明的发表时间为 1990—2021 年,2011 年左右达到第一个顶峰,2020 年前后达到第二个顶峰。这两个高峰分别为施塔珀尔和艾森克为核心的撤稿事件。原始论文发表数量和撤稿声明量,均在 2021 年左右出现明显的下滑趋势,需要考虑研究发表与撤稿的时间滞后、数据库更新的滞后等原因。

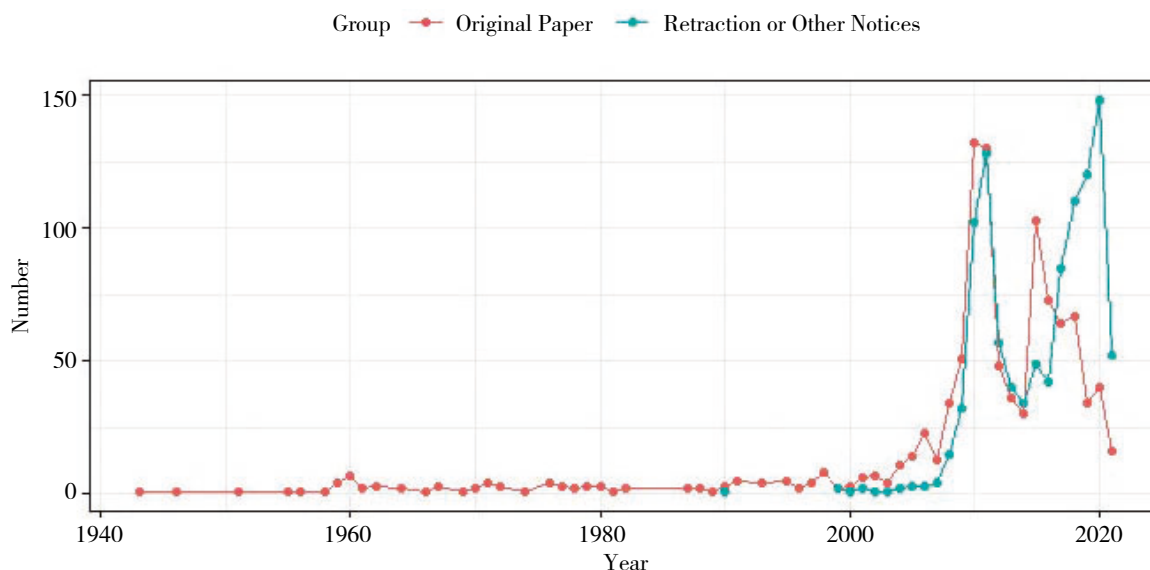


图 3 原始论文及声明量的发表年份变化

Figure 3 Changes in the number of original papers and statements

图 4 为发表撤稿时滞,单位为年,不满一年按一年统计。文章发表到撤回的平均时滞为 5.6 年,这

一时滞较大于其他研究（Dinh et al., 2019, 3.9 年；Budd et al., 1998, 2.2 年）。当区分声明类型后，结果发现撤稿研究时滞为 2.6 年，关注时滞为 32.8 年，勘误时滞为 3.4 年，此时撤稿时滞所代表的自我净化能力较强。

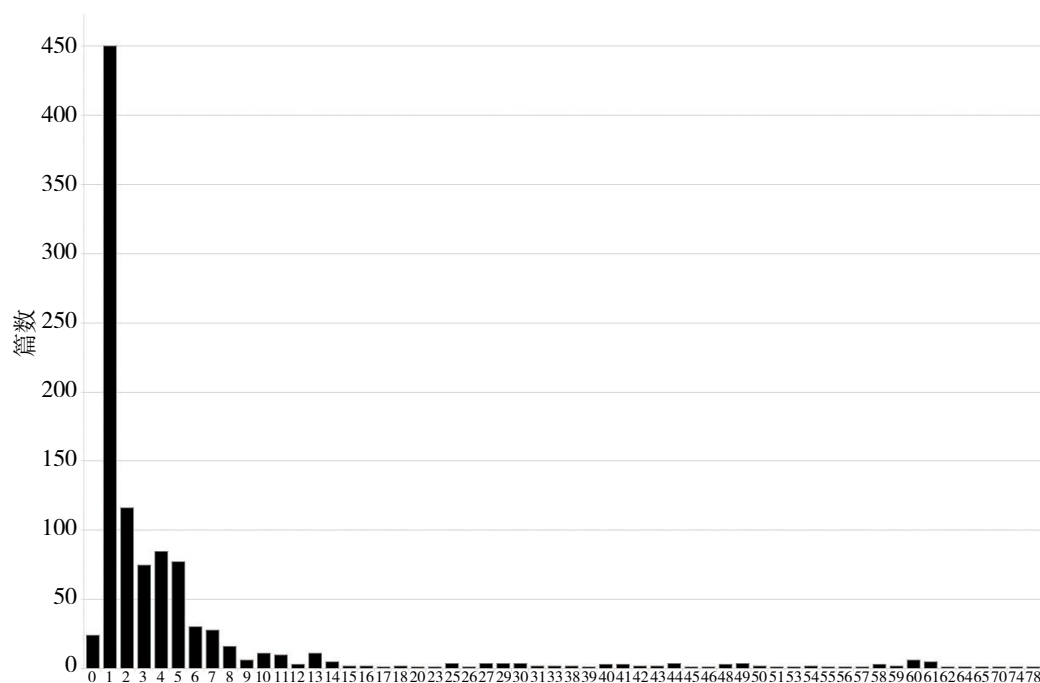


图 4 发表撤稿时滞 / 年

Figure 4 Time lag-year between publishing and retracting

大多数撤稿声明在研究发表 1 年内发生，占比达到 45.8%（其中，24 篇发表当日撤稿），该结果与 2017 年巴伊兰（Bar-Ilan）等人 1 年时滞占比最大的结果一致。撤稿与发表最大跨度为 78 年，时滞时间在 5 年内的文章撤稿量占全部撤稿数量的 77.7%，发表与撤稿在同一天研究有 24 篇，位于每月 1 号的有 6 篇。并且，时滞在 1 年以内的研究中，发表在 2010—2011 年间的占比 45%，如图 5 所示。

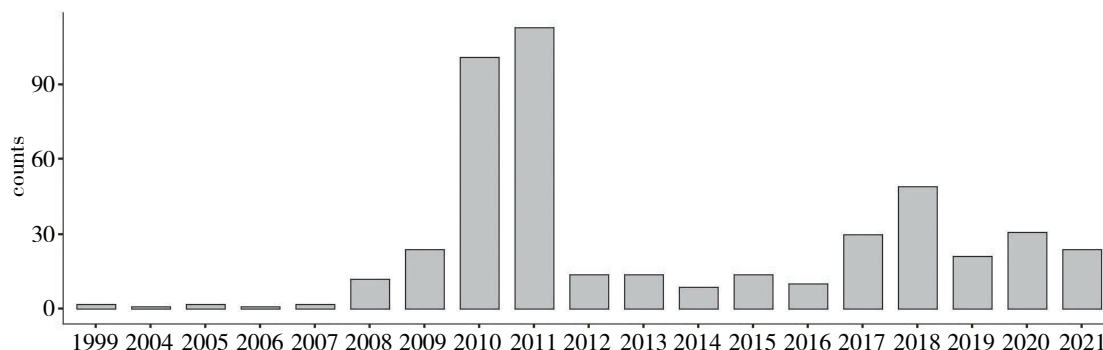


图 5 时滞 1 年内撤稿文章的发表年份频次柱状图

Figure 5 Histogram of the frequency of the year of publication of retracted articles within 1 year time lag

研究为了体现学科交叉情况,制作了领域共现图谱,如图6所示,并进一步计算了图网络的基础指标,网络的度中心性为0.91,紧密中心性为0.95,中介中心性为0.62,特征向量中心性为1。度中心性代表一个节点与其他节点发生直接联系的程度,在图6中核心节点为心理学(Psychology)。

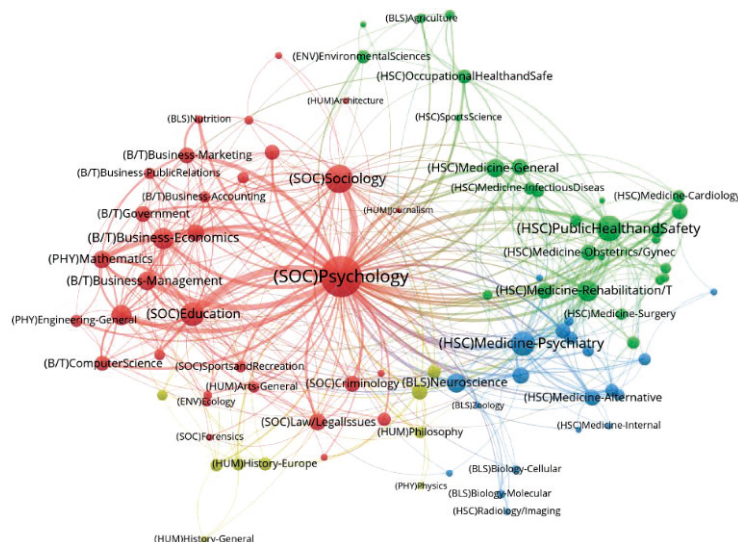


Figure 6 Study area co-occurrence map

在 Retraction Watch 网站上罗列出的 104 条撤稿原因中, 心理学撤稿原因涉及 87 条。在对撤稿主要原因的分析中, 由于部分文章经过勘误或关注后撤稿, 存在多个撤稿原因条目, 因此参照李艾丹等 (2019) 的措施, 取并集处理。在前 10 撤稿原因中, Notice-Limited or No Information 这一撤稿原因最多, 但在讨论具体原因时并不纳入排名 (王景周, 2021)。

在心理学撤稿原因生成共现图谱后,对其进行聚类,结果主要有6大类如图7所示。第一类:版权、审查相关;第二类:研究错误;第三类:学术不端类;第四类:文献状态不明类;第五类:担忧类;第六类:其他类型。

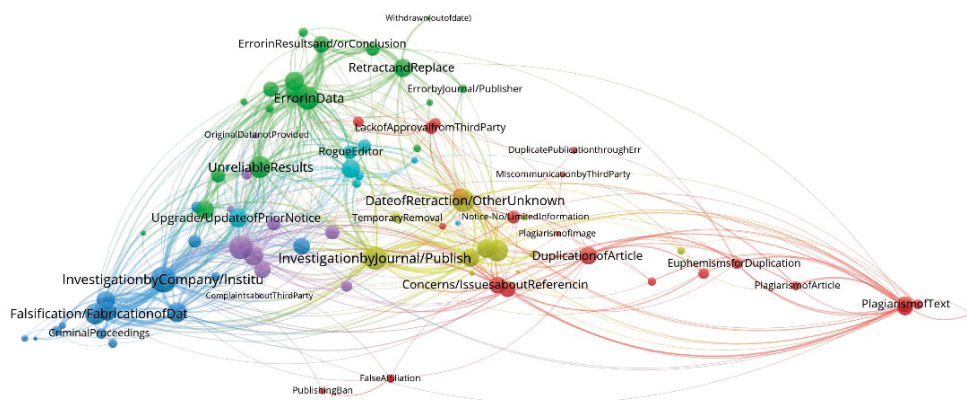


图7 撤稿原因共现图

Figure 7 Co-occurrence chart of reasons for withdrawal

对撤稿原因进行大类归纳,制作撤稿原因共现图谱,如图8所示,该图谱可以主要聚类为3大类,分别是未知或信息不全、研究本身问题和调查原因三类。

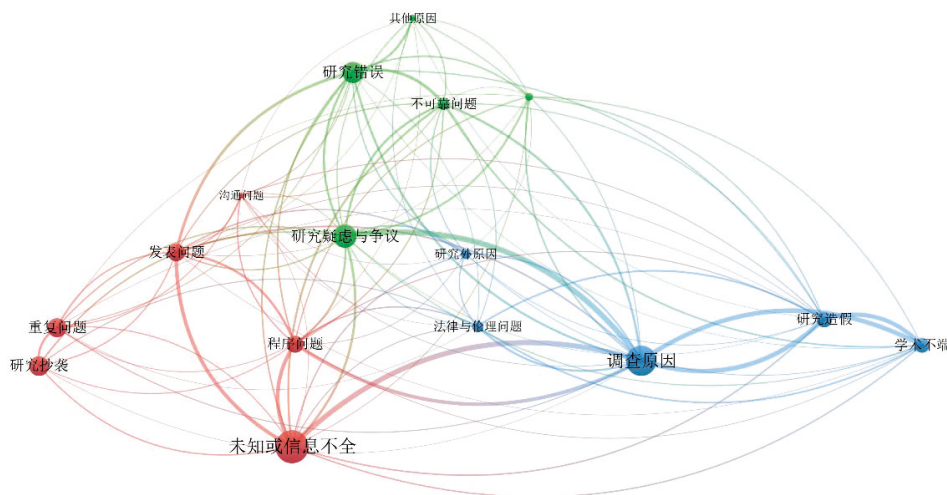


图8 撤稿原因大类共现图

Figure 8 Co-occurrence chart of major categories of reasons for withdrawal

进一步对撤稿原因大类进行频次统计,结果发现,未知或信息不全是撤稿的最高频次的原因,其次是调查原因和研究错误。与未知或信息不全共现次数最高的为研究抄袭、重复问题、发表问题、程序问题和调查原因;与调查原因共现次数最多的是未知或信息不全、研究疑虑与争议、研究造假、学术不端、研究外原因和法律与伦理问题;与研究错误共现次数最高的为不可靠问题、其他问题、研究疑虑与争议。

研究进一步根据撤稿原因所对应文章的结构进行分析，制作了共现知识图谱，如图 9 所示，其他类通常与数据和结果共现，方法和材料通常共现，全文、文本、引用、作者和图像共现更多。

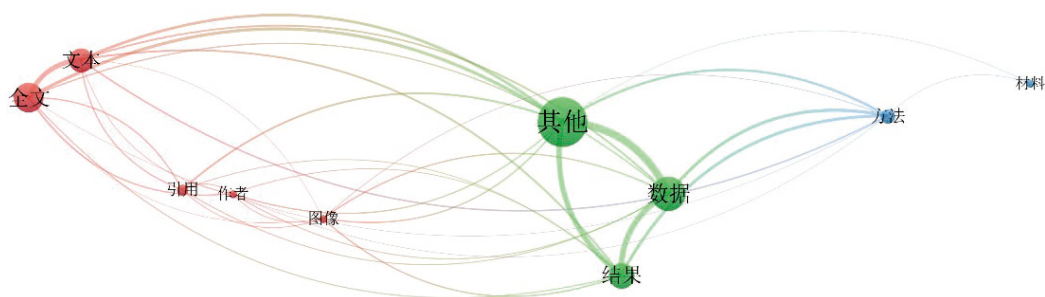


图 9 撤稿原因对应文章结构共现图谱

Figure 9 Co-occurrence mapping of article structure corresponding to reasons for retraction

进一步对撤稿原因对应文章结构频次进行了统计，结果显示，其他占据撤稿主要频次，即学术研究本身的正确性并非影响撤稿的主要因素，而是学术研究正确性以外的其他因素，在学术相关因素中，数据为最高频次原因，其次为全文、结果、文本和方法部分。

最后，根据学术不端的严重程度和主观恶意程度，对研究进行了频次统计，结果发现除去其他类别后，恶意的学术不端频次最高，其次是难以界定，无意错误最少。撤稿文献所涉及的国家/地区中，中国、美国、俄罗斯和英国均占到 10% 以上，四个国家/地区总占比达到 76.6%。为了更好地呈现撤稿情况的国际地域分布，进一步根据撤稿量绘制国际热力图，分布见图 10。

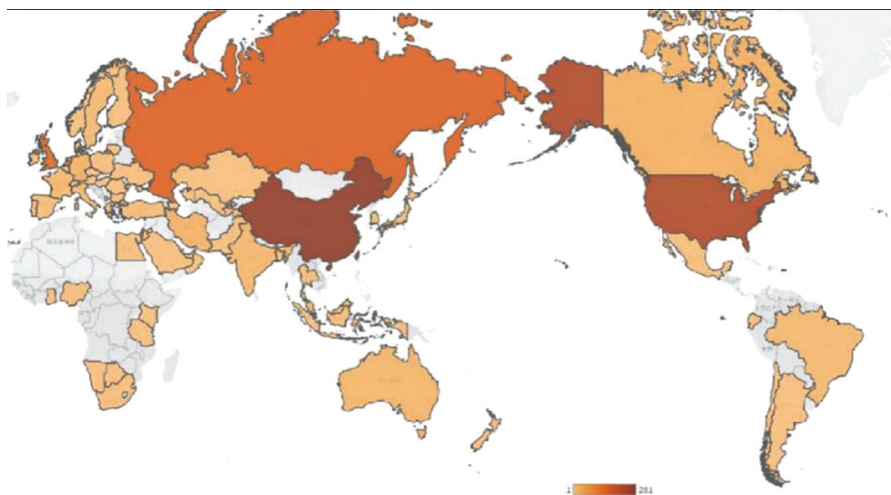


图 10 撤稿国家/地区分布图

Figure 10 Distribution of withdrawal countries/regions

此外，在分析涉及国家撤稿量的基础上，进一步对第一作者的第一单位所属国家进行分析，结果发现中国、美国、俄罗斯、英国和荷兰排名不变，加拿大、德国、澳大利亚排名下降，伊朗、印度排名上升，

先前未出现的中国台湾地区上升到前 10。

为了显示国家之间的合作程度，绘制了国家间的共现图谱，如图 11 所示，结果显示，中、美、英、加、澳、韩、日等国家 / 地区合作较为紧密，论文撤稿量较大的俄罗斯则相对跨国合作研究较少。

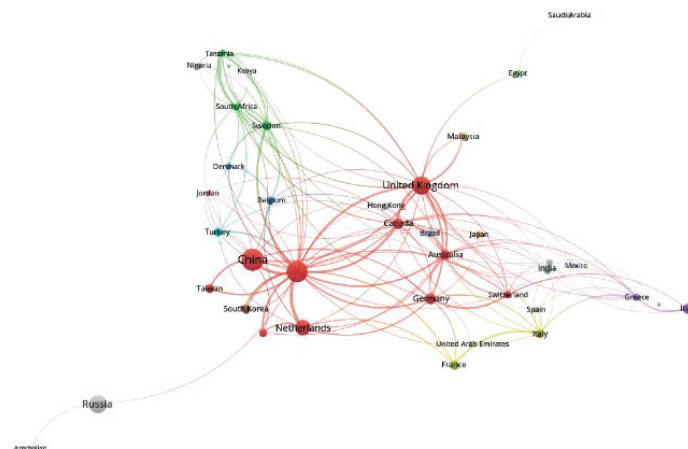


图 11 撤稿国家 / 地区共现图

Figure 11 Co-occurrence of withdrawal countries/regions

不同国家的科研水平、资助条件、学术文化均具有较大差异，因此，进一步分析了国家间的撤稿原因差异（保留了大于 10 频次的撤稿原因）。可以看出中国是未声明撤稿原因、报告缺乏客观、虚假同行评议、期刊审查最多的国家，其中由于未声明原因的撤稿文章中会议论文占据大多数，且会议主要在中国举办，因此更容易发生本国人虚假同行评议的现象。美国是数据、分析、方法、结论错误，数据、结论不可靠最多的国家，未明确报告撤稿声明的仅有 9 条，可以看出美国的撤稿机制比较完善，声明内容比较详实。俄罗斯是重复发表、文本错误、剽窃、替换文章最多的国家，在非英文为母语的国家，同一篇研究以母语和英语分别投递被发现的风险较小，并且剽窃行为具有较小的压力，由于俄罗斯是论文发表大国，因此重复发表、剽窃为主要撤稿原因。荷兰数据造假、官方 / 第三方审查为主要撤稿原因，由于荷兰学术审查制度较为完善（美国类似），可以看出第三方审查发挥的一定作用。

进一步通过对 Retraction Watch 中原始撤稿原因再次归类后，得到第一作者归属国家或地区的撤稿原因频次分布。其中，中国撤稿原因为未知或信息不全最多，程序问题、调查原因、研究错误和研究疑虑与争议也占有一定比例；美国最高频次的撤稿原因为研究错误，其次为调查原因、研究疑虑与争议、学术不端（撤稿声明中的统称，不包括抄袭、造假等原因）和不可靠问题；英国最高频次为研究疑虑与争议和调查原因；俄罗斯最高频次的撤稿原因为抄袭；荷兰最高为学术不端、调查原因和研究造假；伊朗、印度、德国、加拿大和中国台湾地区频次相对较少。

根据心理学研究框架，将撤稿原因进一步分类，中国数据原因撤稿最多，其次是结果部分；美国文本原因撤稿频次较高；俄罗斯非学术相关的撤稿原因占比最大，英国文本、数据、结果部分均有一定比例；其他国家频次较低不进行讨论。

根据撤稿是否属于作者主观恶意进行分类统计，中国、英国难以界定的撤稿原因比例最大，美国无

意错误、学术不端、难以界定撤稿原因均很多,俄罗斯和荷兰学术不端撤稿比例最高。之后分析了撤稿原因的时间变化趋势,结果发现2000年以前,撤稿原因主要以数据、结果的关注与担忧、公司/机构的审查为主,2001—2005年间,没有突出的撤稿原因,2006—2010年间,文章主要以伪造数据、期刊/出版商/第三方审查、未声明具体原因为主,2011—2015年间主要以违反政策、数据/方法/文本错误、结果不可靠、剽窃为主,2016—2020年间主要以分析、结论/结果错误、剽窃、虚假同行评议、撤回再次发表为主,2021年无突出撤稿原因。

进一步根据撤稿原因大类进行年份频次统计,结果发现,1943—2000年间主要撤稿原因为研究疑虑与争议和调查原因,2001—2005年间主要为调查原因和学术不端,2006—2010年间主要为未知或信息不全、调查原因、学术不端、研究造假,2011—2015年间主要为未知或信息不全、研究错误、研究抄袭,2016—2020年间主要为研究抄袭和研究错误,2021年各类原因频次较少,不具有代表性。

根据撤稿原因与心理学研究框架的对应进一步统计年份变化频次,可以发现,1943—2000年间主要撤稿对应部分为数据和结果,2001—2005年间主要为数据,2006—2010年间主要为数据、全文和结果,2011—2015年间主要为全文、数据、文本、结果,2016—2020年间主要为全文、数据、文本、结果。

进一步分析了不同文章类型的撤稿原因,其中综述文章主要撤稿原因为撤回旧文章并以新文章替换,有关文本方面的问题(如文本错误、抄袭)也占有一定比例。元分析文章最主要撤稿原因为分析错误,研究文章主要撤稿原因为机构审查和数据问题的担忧,临床研究主要撤稿原因为研究方法的错误,预印本撤稿原因主要为对数据和结果的担忧,待发表文章撤稿原因为结果、数据、作者贡献相关的问题和审查原因。

接着对不同类型文章的撤稿原因大类频次进行统计,可以看出,研究类文章最常见的撤稿原因为调查原因、研究错误、研究疑虑与争议、研究抄袭、学术不端、重复问题,会议类文章最常见的撤稿原因为未知或信息不全、调查原因、程序问题,综述类文章撤稿最常见原因为发表问题、未知或信息不全、研究抄袭,临床研究最常见原因为研究错误,元分析最常见的撤稿原因为研究错误,待发表研究最常见原因为研究疑虑与争议。

之后,针对不同类型文章及撤稿原因所对应文章结构部分的频次进行统计,结果发现,研究类文章对应数据、全文、结果、文本部分最多,会议论文主要集中在其他非文章结构本身问题上,综述文章主要集中在全文上,临床研究主要集中在结果和数据部分,元分析主要集中在数据部分,其他类型的文章频次较少,且主要频次集中在其他非文章结构本身的问题上,没有代表性。

在对期刊文献阅读是否有收费阻碍的分析上,绝大多数(95%)的心理学撤稿原始文献均为开放获取,该因素也印证了开源对撤稿监督的促进作用。不过,同样由于近年来OA期刊的快速增长,涌现部分“掠夺性期刊”(Predatory Publishers, Shrestha, 2020),存在大量的虚假同行评议现象,当受到研究者阅读检验后引发质疑甚至撤稿。

Retraction Watch数据库中纳入的学者中艾森克的撤稿声明最多,作为人格心理学领域的著名心理学家,其文章被大量撤稿受到了广泛的关注。77篇声明均发生在2020年2月10日以后,且该声明数达到69篇,涉及期刊*Perceptual and Motor Skills*和*Psychological Reports*分别为36和33篇。撤稿声明数第二多的学者是施塔珀尔,其学术不端事件在2011年引起心理学界广泛关注。撤稿声明量第三的学者

是万辛克（Wansink），Nature 曾针对性发表过报道（Fidler，2020）。三位学者总撤稿声明达到总数的15.8%，针对三位学者的撤稿曝光均引发心理学领域广泛的关注与讨论。共出现2357位作者，817位第一作者撤稿，730位第一作者仅出现1次。并且非第一作者的1606人中，1534为未当第一作者但被撤稿。

学者共现图谱也显示出艾森克、施塔珀尔和万辛克三位学者撤稿涉及的规模之大，如图12所示。

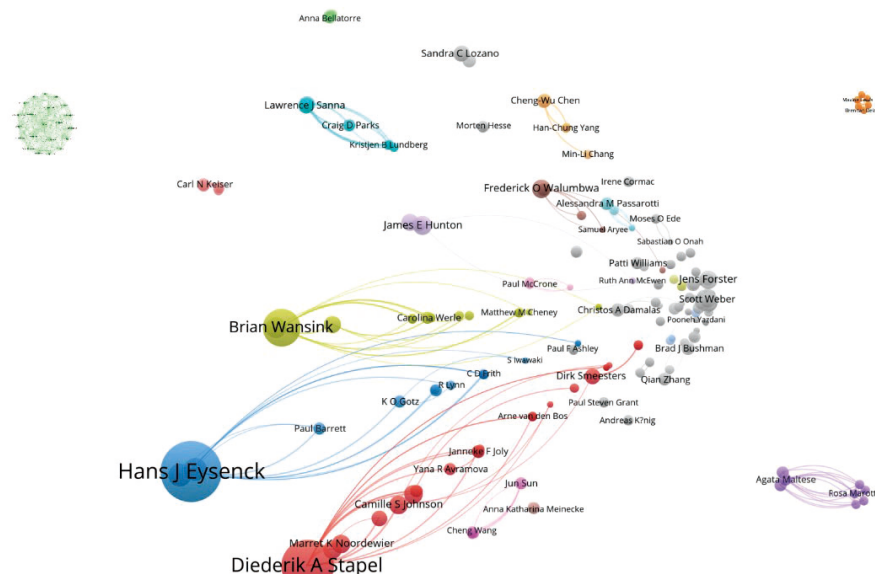


图12 撤稿作者共现图

Figure 12 Co-occurrence of retracted authors

对作者人数的频次进行统计。撤稿声明中，以往有研究发现单个作者比例为6.6%，由于未责任分散，远小于多位作者撤稿的5.0%（Foo，2011；Campos-Varela and Ruano-Raviña，2019），不过心理学领域的单个作者撤稿比例较高，占比24%，两位作者达到32%。在大于2位作者的研究中，沟通不畅导致一稿多投、署名争议等问题，同时也有责任分散的可能。

最后，针对本研究纳入数据的各项维度进行皮尔逊相关检验，结果发现，标题长度与作者数有显著的正相关，标题长度与时滞日具有显著的负相关，国家数量与作者数量具有显著的正相关。

4 讨论

4.1 结果解释

通过以上分析，主要得到以下结果：（1）心理学界自净能力较强（时滞较短、最大时滞跨度较大）；（2）撤稿具有以学者（艾森克、施塔珀尔和万辛克）、年份（2011年和2020年）、期刊（*Perceptual and Motor Skills*、*Psychological Reports*、*European Journal of Social Psychology*等）为核心引起的广泛关注为特点。顶级学者的撤稿事件会引发领域内、甚至领域外的大量讨论，督促学者、期刊等主体加强对研究的审查，增强本领域的学术规范意识；（3）心理学撤稿文献主要交叉学科为教育学和社会学。心理学与

药、精神病学、社会学、教育学、商业学等多个领域高度交叉,针对相似科学问题的不同领域采用的研究方法,从基因、表征、蛋白、功能、结构网络、影像学、行为(个人行为和社会行为)、质性过程的多层级研究对照,会得到相似或相左的印证结果,因此学科交叉有利于领域内的自我净化。并且在学科高度融合的当下,领域的界限逐渐模糊,在促进不同学者相互了解、合作的同时,也增大了跨学科研究者失误的风险,这要求学者在掌握好自身领域的专业知识的同时,要主动加强对合作者、相关领域研究内容、方法的理解,承担跨学科合作中的共有责任;(4)中国、美国、英国、俄罗斯为心理学研究撤稿大国。研究的撤稿量与领域内文献的发文量有很大的关系,撤稿量以发文量为基础,是一个领域国内学术圈层研究思想活跃交流、严格审查、自我净化的体现;(5)心理学研究主要撤稿原因为对数据质量的担忧。研究文章(Research Article)占据了撤稿文献的最大比重,其中由于数据原因撤稿占据了313条,结果部分占据147条,文本问题占据119条,这说明界定层、表征层和分析层三个部分都存在较大问题,且研究越是下游的环节出现的撤稿声明越多。从数据有效性的缺失为起点,无法有效地在分析层通过有效的数据分析方案,得到实践任务上的联系,从而无法建立抽象能力之间的各类关系。这造成了在分析层得到有效结果的缺失,并进一步加大了通过结果进行解释、利用文本引出研究核心观点的困难;(6)撤稿声明主要来源为无收费获取障碍的期刊文章。施塔珀尔事件后,center for open science(COS)搭建了open science framework(OSF)平台以促进心理学研究的大规模重复计划,开放科学、研究的可重复性概念得到了重视(胡传鹏等,2016),开放获取文献是开放科学运动中的一个重要环节,为读者对已经发表的研究进行阅读、审视提供了支持,促进学界的撤稿自净行为。但无付费获取障碍期刊的高比例也体现出部分低水平期刊的扩张现象,降低了研究的出版审查标准,增大了虚假的同行评议风险。

撤稿原因的分析发现,学术错误比学术不端更容易被撤稿,这与过去的研究结论相符(Steen et al., 2013; Steen, 2011)。学术错误大多是学者非主观恶意行为,而学术不端则是学者以发表文章为目的进行的恶意行为,更具有隐蔽性。并且学术不端范围内的撤稿原因,以对数据的操纵、伪造数据为主,较难从源头发现研究的问题,针对此类研究需要加强元分析、系统评价等研究的整合分析,同时加强对心理学研究可重复性问题的重视。

除此以外,研究还发现了2010年的特殊性,COPE首次在前一年即2009年颁布标准统一的撤稿准则,Staple学者在后一年即2011年大规模造假,开放科学运动和研究的可重复性被广泛呼吁。丁恩(Dinh)等于2019年也在研究中指出,一般文章对撤稿文章的引用在2010年有下降的趋势,但在2011年逐渐回升。同样,在邓支青和吴任力(2021)的研究中,2010年Retraction Watch Database的收录撤稿量、会议文献撤稿量均为高峰顶点。

4.2 撤稿问题建议

法内利(Fanelli)等于2015年分析了引发学者学术不端的相关因素,结果表明不完善的学术诚信政策、自我监督及相互批评文化的缺乏这两个因素可以显著预测学者的学术不端行为,而学者的性别、发表论文的压力则不能有效预测。另外,先前的研究发现,论文撤稿由学者自我报告错误的占比达到21.9%,研究者主动承认错误后,先前被引量由于撤稿导致的下降趋势逐渐改善,但非自我报告的撤回文章被引损失量即使在5年后依然会有影响(Lu et al., 2013)。撤稿不仅仅是第一作者、通讯作者、其他合作者

的责任,编辑、审稿人、出版商也需要履行自己的职责。并且撤回文章并非对研究者惩戒的全部,需要加大对未来学者发表文献的监督,因为有撤稿研究领域的学者发现,部分由于数据问题被撤稿者,甚至在之后的发表过程中继续使用虚假数据(Foo, 2011)。

以上都表明减少撤稿事件的发生,需要多方主体从以下多个层面进行改进。首先,政策层面上,国家近年来已经出台很多科研诚信和学术规范有关文件对相关概念、范围进行界定(张艳、焦艳, 2021),且中国科研诚信网也于2021年7月30日成立。但目前尚无英、美、日等国一样完善的学术科研诚信监督体系,我国仍迫切需要一套融合科研诚信教育、不端行为预防、研究全流程监督、不端行为举报处理及再教育、学术成果评价、公众媒体报道等方面的完善制度。

在研究主体方面,从本科生、硕士生、博士生入学起,加强科研诚信教育。有研究表明,高质量期刊的审核也更为严格,这导致其撤稿量也相应地高于低质量期刊(Cokol, 2008),因此,除了诚信教育,更应该加强学术能力的提升,较强的学术科研能力和科研思维是产出高质量文章的前提条件,从学者研究层面减少撤稿的可能。期刊及出版商推广学术不端检测工具的应用,其审稿人也需要加强对研究者投递稿件的高质量要求,协助研究者的文章内容提升并加以监督。另外,第三方监管也是重要的主体,对学术不端产生重要的影响,第三方通常可以是公众媒体、同行评议,这是撤稿研究发现、对学术界产生重大影响前制止的重要环节。

在技术层面,审查者可以利用大数据筛查的手段,对项目的资金使用情况异常值检测,结合区块链技术协助科研诚信管理平台的构建,学术不端的识别层面,随着目前筛查学术不端软件的进步(iThenticate and Turnitin, Steen et al., 2013)和剽窃的检测技术的发展(如基于引用的剽窃检测的引文分析(citation-based plagiarism detection, CbPD)、语义分析等),可以解决意译剽窃、翻译剽窃、结构剽窃、思想剽窃等多个维度的学术不端。

在学术交流层面,要促进学术成果的传播与交流。论文的合作者、同领域或交叉学科领域内潜在的合作者之间加强相互交流,既可以促进高质量研究的传播,又可以加快不可靠研究的发现,公开讨论目前也有Pubpeer、Twitter等平台为学者提供了交流思想的媒介。同时,也鼓励撤稿文章的学术不端者及其合作者,积极自我反省,主动报告不端行为,进一步主动撤稿。

在撤稿信息传播层面,也许要期刊在决定撤稿后,对撤稿声明进行及时、明确的信息传递,减轻对学术界的不利影响。通常撤稿具有较长的滞后性,迪卡利尔等人(Decullier et al., 2014, 2013, 2018)对此展开了多项研究。他们每三个月对Pubmed上的撤稿文献进行一次检索,发现直到35个月左右,撤稿数量才趋近于稳定。并且,撤稿声明也存在表述不清的现象,45%的撤稿文章并没有记录清楚基本信息(Steen, 2011),这一比例在他们的研究中也达到了22%,甚至在更早以前,91%的撤稿文章没有任何标记(Pfeifer and Snodgrass, 1992)。因此,需要采取一些措施减少撤稿声明发布的滞后性和透明度,在互联网资源中标明撤稿水印,推广作者投递稿件时使用ORCID以利于学术不端学者发表文献信息的追踪,推广文献独特标识引用符号以利于被引文章的追踪。

另外,有学者发现,一些研究者将会会议论文延伸到期刊论文,从而主动撤稿会议论文以促进研究在另一期刊上的发表(邓支青、吴任力, 2021),此类非恶意的撤稿也建议建立平台对其进行标识归纳,协助读者对重新投递发表文章的追踪,而不是止步在撤稿文章页面。以往的撤稿经历并不能完全否定所

有学者及其他已发表工作,更不能否定未来产生的科研成果。在深刻批评教育撤稿文章作者之后,也要建立一套给予改正错误,继续推进科研进步机会的制度,促进学术界的规范与科研的进步。

5 补充材料

研究正文涉及 33 个表格,见在线文档:

<https://k1ax0er2lc.feishu.cn/docx/P5P9dwFKxouoSHxTgFFc54yKnld>

参考文献

- [1] Almeida R M V R, Catelani F, Fontes-Pereira A J, et al. Retractions in general and internal medicine in a high-profile scientific indexing database [J]. Sao Paulo Medical Journal, 2015, 134 (1): 74-78.
- [2] Amos K A. The ethics of scholarly publishing: Exploring differences in plagiarism and duplicate publication across nations [J]. Journal of the Medical Library Association: JMLA, 2014, 102 (2): 87-91.
- [3] Azoulay P, Bonatti A, Krieger J L. The Career Effects of Scandal: Evidence from Scientific Retractions [J]. The Review of Economics and Statistics, 2015 (9).
- [4] Bar-Ilan J, Halevi G. Post retraction citations in context: A case study [J]. Scientometrics, 2017, 113 (1): 547-565.
- [5] Barbour V, Kleinert S, Wager E, et al. Guidelines for retracting articles [EB/OL]. [2023-02-10]. <https://publicationethics.org/node/19896>.
- [6] Bilbrey E, O'Dell N, Creamer J. A Novel Rubric for Rating the Quality of Retraction Notices [J]. Publications, 2014, 2 (1): 14-26.
- [7] Brainard J, You J. What a massive database of retracted papers reveals about science publishing's "death penalty" [EB/OL]. [2023-02-10]. www.science.org.
- [8] Budd J M, Sievert M, Schultz T R. Phenomena of Retraction [J]. JAMA, 1998, 280 (3): 296.
- [9] Wager E, Barbour V, Kleinert S, et al. COPE Guidelines for retracting articles [EB/OL]. [2023-02-10]. Committee on Publication Ethics, <https://doi.org/10.24318/cope>.
- [10] Campos-Varela I, Ruano-Raviña A. Misconduct as the main cause for retraction. A descriptive study of retracted publications and their authors [J]. Gaceta Sanitaria, 2019, 33 (4): 356-360.
- [11] Cokol M, Ozbay F, Rodriguez-Esteban R. Retraction rates are on the rise [J]. EMBO Reports, 2008, 9 (1): 2.
- [12] Decullier E, Huot L, Maisonneuve H. What time-lag for a retraction search on PubMed? [J]. BMC Research Notes, 2014, 7 (1).
- [13] Decullier E, Huot L, Samson G, et al. Visibility of retractions: A cross-sectional one-year study [J]. BMC Research Notes, 2013, 6 (1).
- [14] Decullier E, Maisonneuve H. Correcting the literature: Improvement trends seen in contents of retraction notices [J]. BMC Research Notes, 2018, 11 (1).
- [15] Dinh L, Sarol J, Cheng Y, et al. Systematic examination of pre-and post-retraction citations [J]. Proceedings of the Association for Information Science and Technology, 2019, 56 (1): 390-394.
- [16] Fidler F. Fraud, bias, negligence and hype in the lab —a rogues' gallery [J]. Nature, 2020, 583

- (7817): 515–516.
- [17] Fanelli D. Why Growing Retractions Are (Mostly) a Good Sign [J]. *PLoS Medicine*, 2013, 10 (12): e1001563.
- [18] Fang F C, Steen R G, Casadevall A. Misconduct accounts for the majority of retracted scientific publications [J]. *Proceedings of the National Academy of Sciences*, 2012, 109 (42): 17028–17033.
- [19] Foo J Y A. A Retrospective Analysis of the Trend of Retracted Publications in the Field of Biomedical and Life Sciences [J]. *Science and Engineering Ethics*, 2010, 17 (3): 459–468.
- [20] Furman J L, Jensen K, Murray F. Governing knowledge in the scientific community: Exploring the role of retractions in biomedicine [J]. *Research Policy*, 2012, 41 (2): 276–290.
- [21] Hu C, Wang F, Guo J, et al. The replication crisis in psychological research [J]. *Advances in Psychological Science*, 2016, 24 (9): 1504.
- [22] Hwang W S, Roh S I, Lee B C, et al. Patient-specific embryonic stem cells derived from human SCNT blastocysts [J]. *Science (New York, N. Y.)*, 2005, 308 (5729): 1777–1783.
- [23] Kochan C A, Budd J M. The persistence of fraud in the literature: the Darsee case [J]. *Journal of the American Society for Information Science*, 1992, 43 (7): 488–493.
- [24] Liu X, Chen X. Authors' noninstitutional emails and their correlation with retraction [J]. *Journal of the Association for Information Science and Technology*, 2020, 72 (4): 473–477.
- [25] Lu S F, Jin G Z, Uzzi B, et al. The Retraction Penalty: Evidence from the Web of Science [J]. *Scientific Reports*, 2013, 3 (1).
- [26] Moylan E C, Kowalczyk M K. Why articles are retracted: A retrospective cross-sectional study of retraction notices at BioMed Central [J]. *BMJ Open*, 2016, 6 (11): e012047.
- [27] Nath S B, Marcus S C, Druss B G. Retractions in the research literature: misconduct or mistakes? [J]. *Medical Journal of Australia*, 2006, 185 (3): 152–154.
- [28] Oransky A I. The first-ever English language retraction (1756) ?Retraction Watch [EB/OL]. [2023-02-10]. <http://retractionwatch.com/2012/02/27/the-first-ever-english-language-retraction-1756/>.
- [29] Oransky I A M. (2010, August 3) Retraction Watch [EB/OL]. [2023-02-10]. <http://retractionwatch.wordpress.com>.
- [30] Pfeifer M P. The Continued Use of Retracted, Invalid Scientific Literature [J]. *JAMA: The Journal of the American Medical Association*, 1990, 263 (10): 1420.
- [31] Pudovkin A I, Garfield E. Rank-normalized impact factor: A way to compare journal performance across subject categories [J]. *Proceedings of the American Society for Information Science and Technology*, 2005, 41 (1): 507–515.
- [32] Schubert A, Braun T. Relative indicators and relational charts for comparative assessment of publication output and citation impact [J]. *Scientometrics*, 1986, 9 (5/6): 281–291.
- [33] Sen B K. NORMALISED IMPACT FACTOR [J]. *Journal of Documentation*, 1992, 48 (3): 318–325.
- [34] Shrestha J. List of predatory journals and publishers [EB/OL]. [2023-02-10]. <http://eprints.rclis.org/40643>.
- [35] Steen R G. Retractions in the scientific literature: is the incidence of research fraud increasing? [J]. *Journal of Medical Ethics*, 2010, 37 (4): 249–253.
- [36] Steen R G, Casadevall A, Fang F C. Correction: Why Has the Number of Scientific Retractions Increased?

- [J]. PLoS ONE, 2013, 8 (7) .
- [37] The Scientist Magazine [EB/OL] . [2023-02-10] . [Scientist.com](https://www.the-scientist.com/). <https://www.the-scientist.com/>.
- [38] The jamovi project, jamovi-Stats [EB/OL] . [2023-02-10] . [Now.Www.jamovi.org](https://www.jamovi.org). <https://www.jamovi.org>.
- [39] Thompson Reuters. Science Watch: Top 20 countries in all fields, 2001–August 31, 2011 [EB/OL] . [2023-02-10] . <http://archive.sciencewatch.com/dr/cou/2011/11decALL/>.
- [40] The Retraction. The Retraction [EB/OL] . [2023-02-10] . [Retractiondatabase.org](http://retractiondatabase.org). <http://retractiondatabase.org/>.
- [41] Van Noorden R. Science publishing: The trouble with retractions [J] . Nature, 2011, 478 (7367) : 26–28.
- [42] van Eck N J, Waltman L. VOS: A New Method for Visualizing Similarities Between Objects [M] // Decker R, Lenz H J. Advances in Data Analysis. Studies in Classification, Data Analysis, and Knowledge Organization. Springer, Berlin, Heidelberg, 2007.
- [43] Wililam Broad, Wade N. Betrayers of the truth [M] . Simon And Schuster, 1982.
- [44] Wager E, Barbour V, Yentis S, et al. Retractions: guidance from the Committee on Publication Ethics (COPE) [J] . Obesity Reviews, 2010, 11 (1) : 64–66.
- [45] Wager E, Williams P. Why and how do journals retract articles?An analysis of Medline retractions 1988–2008 [J] . Journal of Medical Ethics, 2011, 37 (9) : 567–570.
- [46] Wang J, Ku J C, Alotaibi N M, et al. Retraction of Neurosurgical Publications: A Systematic Review [J] . World Neurosurgery, 2017 (103) : 809–814.
- [47] Yan J, Mac Donald A, Baisi L-P, et al. Retractions in orthopaedic research [J] . Bone & Joint Research, 2016, 5 (6) : 263–268.
- [48] 胡传鹏, 王非, 过继成思, 等. 心理学研究中的可重复性问题: 从危机到契机 [J] . 心理科学进展, 2016, 24 (9) : 1504–1518.
- [49] 李艾丹, 张希, 梁冰. 中国学者发表的国际科技论文被撤销现象分析 [J] . 中国科技期刊研究, 2019, 30 (10) : 1034–1038.
- [50] 王景周. 因“论文工厂”撤回论文的特征分布及治理路径 [J] . 中国科技期刊研究, 2021, 32 (12) : 1507–1518.
- [51] 周志新. 中文科技期刊被撤销论文特征分析及启示 [J] . 中国科技期刊研究, 2017, 28 (11) : 1065–1070.

Analysis of Reasons for Withdrawal of Psychology Paper

Kang Cunyi^{1,2} Wang Zijin^{1,2} Wang Wenxuan^{2,3}

1. Beijing Normal University, Zhuhai;

2. Psychological Corridor Project, Institute of Balanceomics and Interdisciplinary Standardized Research
Interactive System, Beijing;

3. Liquid Studio, Accenture Greater China, Shanghai

Abstract: Objective: Scientific research results should be based on traceable methods. Once the methods are faulty, the scientific research results are no longer reliable. Current psychological research suffers from a series of subjective and objective methodological irregularities, which undermine the reliability of empirical research and further trigger corrections and retractions. This paper provides suggestions for avoiding research irregularities in psychological research with a view to discussing the impact of irregular academic research on the academic community. Methods: The study was conducted with 1034 statements of concerns, corrections and retractions in the various subjects of the Retraction Watch Database for psychology. The study discussed the characteristics of retracted articles in the field of psychology based on the past framework of psychology research. Results: The most common reasons for retraction in the field of psychology were concerns and problems related to data, plagiarism, and duplicate publications. The most retracted journals are *Perceptual and Motor Skills* and *Psychological Reports*; the most retracted conference is ICEE held in Shanghai in 2011; the most retracted publisher is IEEE, followed by SAGE and Wiley respectively; the most retracted countries are China, USA, Russia and UK; the maximum retraction time lag is 78 years, the minimum is 0 days (i.e. Retracted in the same day), the average time lag is 5.6 years. Conclusion: Strengthening the standardization of psychology research and avoiding retractions due to academic reasons from the source as much as possible are conducive to the stable development of the field of psychology.

Key words: Psychological research framework; Retraction; Scientometric