

引文内容分析视角下中文造假论文被引分析*

——以国家自然科学基金委员会查处的造假论文为例

杨 珠

《泉州师范学院学报》编辑部, 362000, 福建泉州

摘要 以国家自然科学基金委员会查处的中文造假论文为例, 基于引文内容分析其被其他文献引用的情况。调查表明: 11 篇中文造假论文共被有效引用 34 次; 在定性为造假论文前被引用 26 次(占 76.47%), 在定性为造假论文后被引用 8 次(占 23.53%, 其中 1 次为负面引用); 从引用情感来看, 15 次(占 44.12%) 为正面引用, 18 次(占 52.94%) 为中性引用, 1 次(占 2.94%) 为负面引用。建议加强发现造假论文的机构(个人)、作者单位与学术期刊等的沟通交流, 督促造假论文尽快被撤销并在撤销声明中阐述具体原因; 基于科学研究不断发展的特点以及学术不端行为长期存在的现状, 科研工作者应具备质疑精神, 对前人的研究结果进行借鉴和识别; 关注造假论文的正面引用, 应警惕连环造假行为。

关键词 造假论文; 引用分析; 中文论文; 学术不端; 自然科学

Citation analysis of Chinese falsified papers from the perspective of citation content analysis: taking the falsified paper investigated by the National Natural Science Foundation of China as an example//YANG Zhu

Abstract Take the falsified papers in Chinese investigated by the National Natural Science Foundation of China as an example, we analyzed their citations by other literature based on the content of the citations. The survey shows the following results. Altogether, 11 falsified papers in Chinese were cited 34 times, among which 26 times (76.47%) before being classified as falsified papers and 8 times (23.53%, including one negative citation) after. In terms of citation sentiment, 15 times (44.12%) were positive citations, 18 times (52.94%) were neutral and one time (2.94%) was a negative. It is recommended that communication should be strengthened between the institution (individual) that discovered the falsified paper, the author's affiliation and the academic journal, etc., and the falsified paper should be withdrawn as soon as possible with specific reasons. Based on the evolving nature of scientific research and longstanding academic misconduct, researchers should have questioning spirit and be able to learn from and discriminate previous results, and pay attention to the positive citations of falsified paper to avoid sequential fraudulent.

Keywords falsified papers; citation analysis; Chinese papers; academic misconduct; natural sciences

Author's address Editorial Department of Journal, Quanzhou Normal University, 362000, Quanzhou, Fujian, China

DOI:10.16811/j.cnki.1001-4314.2022.03.010

近年, 国内外媒体揭发我国各类学术不端论文的新闻频发, 学术不端论文受到大众和学术界的关注和重视。造假论文与一文多发、抄袭等其他学术不端论文相比, 其性质更恶劣, 对社会和学术界的影响更严重^[1]。当前, 学者关于中文造假论文的研究主要集中在如何防范造假论文的发表, 如识别造假数据^[2]、造假图片^[3]、造假作者信息^[4]等。相对于中文期刊数据库, 国外一些期刊数据库对已发表论文的撤销较成熟, 相关数据的获取较容易, 因此, 对已发表的造假论文被引用的研究主要集中在外文论文。例如, Schneider 等^[5]以 1 篇被撤销的造假论文为例, 从其被引用情况探讨数字文献时代撤销论文的传播问题。张春博等^[6]以 1 篇自然科学高被引撤销论文和 1 篇社会科学高被引撤销论文为例, 通过对其全文引文进行分析, 探讨科学自我纠错机制的作用发挥情况。袁子晗等^[7]以哈佛大学心脏干细胞撤稿事件为例, 对 16 篇撤销的造假论文施引文献引用态度进行标记与统计, 探讨其变化轨迹。目前, 对已发表中文造假论文被引的研究较少, 仅丛敏等^[8]对 1 篇已被期刊数据库撤销的中文造假论文进行追踪分析, 探讨了该造假论文的引用特征。

目前, 造假论文研究主要集中在论文发表前的预防, 对发表后的造假论文研究较少。而对发表后的造假论文也主要集中在已撤销的造假论文, 这是因为从数据获取来说较为方便。此外, 已发表的造假论文由于在期刊数据库中无法直接查询获得, 数据获取难度较大, 研究存在一定的难度。在此背景下, 本研究针对已发表却未被撤销的中文造假论文, 从引文内容分析的视角出发, 探讨中文造假论文被引用情况, 揭示造假论文在学术界造成的影响及其原因, 分析科学自我纠错机制在造假论文中的作用, 以期对相关论文撤销甚至学术不端行为的研究以及期刊、管理部门对学术不端事件的处理提供一定的参考。

1 研究方法对象

在国家自然科学基金委员会官网中查询其监督委

* 中国高校科技期刊研究会学术诚信与版权专项基金(CUJS-CX-2021-008)

员会的“处理决定”,调查2015—2016年间国家自然科学基金委员会查处的不端行为案件处理决定^[9],提取涉及中文造假论文的基本信息。这些造假论文发表在2008—2010年,2015年被国家自然科学基金委员会定性为造假。笔者于2021年12月25日在中国知网调查其被引用情况。该调查时间点既可调查论文造假定性前的引用情况,也可调查论文造假定性后的引用情况,且二者的时间段较为均衡。调查时,造假论文对施引文献的影响采用人工通篇阅读施引文献的全文文本进行识别标注和分析,最终生成本文所需的施引情况。

2 结果与分析

2.1 论文基本情况

调查2015—2016年间国家自然科学基金委员会查处的不端行为案件处理决定,共得到中文造假论文13篇,其中1篇未被中国知网收录,1篇被撤销。因此,共得到中文造假有效论文11篇,具体信息见表1。这11篇论文分别发表在2008年(5篇)、2009年(2篇)、2010年(4篇),涉及期刊8种。期刊影响因子最高的为1.117,最低的为0.205,中位数为0.860。其中4篇造假论文发表在中国科学引文数据库CSCD收录期刊中。这些论文在期刊数据库中均未标注为造假论文,因此也未被中国知网期刊数据库加以撤销。也就是说,这些造假论文以正常论文的形式存在于期刊数据库中,供读者正常下载、阅读和引用。

表1 造假论文的基本信息

论文编号	发表年份	期刊影响因子	下载量	被引次数
1	2010	0.860*	125	5
2	2010	0.869	155	5
3	2009	0.713	77	1
4	2010	0.507	132	3
5	2008	0.205	63	0
6	2008	0.656	149	2
7	2009	1.117*	96	4 [#]
8	2008	1.117*	359	11 [#]
9	2008	0.205	99	1
10	2008	0.882	135	0
11	2010	1.117*	170	4

注:*表示CSCD收录期刊;#为中国知网引用数据,实际有效被引次数均减少1次。

2.2 论文被引用次数

被引用次数体现了文献的学术影响情况。截至2021年12月25日,11篇造假论文共被引用36次(其中2次为施引文献的文末参考文献有该造假论文,但正文无论论文的引用标示符,故有效被引用次数应为34次,后文所说的被引用次数皆为有效被引用次数),

最高频次为10次,最低频次为0次,中位数为3次,平均值为3.1次。按施引文献的载体分类可分为:被期刊引用21次(占61.76%),被硕士学位论文引用12次(占35.29%),被博士学位论文引用1次(占2.94%)。按施引文献的作者分类:被造假论文的课题组或者本校相关研究人员引用11次(占32.35%),被其他人员引用23次(占67.65%)。按造假论文定性前后的施引时间分类可分为:未被鉴定为造假论文时被引用26次(占76.47%),被鉴定为造假论文后被引用8次(占23.53%,其中1次为负面引用)。

被引用的位置是指造假论文在施引文献中被引时出现的位置,如前言、方法、分析讨论和结论等。调查显示,34次被引有11次(占32.35%)出现在综述类文献(包括硕博学位论文的综述部分和期刊综述类论文),其余分布在研究性学术论文的情况分别是前言(10次,占29.41%)、方法(2次,占5.88%)、分析讨论(11次,占32.35%)和结论(0次)。在研究性学术论文中,造假论文被引用主要集中在前言和分析讨论中。前言主要是阐述相关前期研究进展的引用,而分析讨论主要是引用造假论文的结论或者内容对自己的研究加以论证支持。

造假论文在被发现前,是以正常论文形式在各大期刊数据库中存在,因此被引用属于正常情况。造假论文被发现需要一定的时间,因为造假论文多数为数据、图片等造假,与一文多发、抄袭相比,需要通过分析数据和图片、重复实验甚至是相关知悉内部信息的人员举报才可能被发现,总体而言它比较难被直接发现。而此次调查的造假论文是由国家自然科学基金委员会监督委员会收到举报而调查发现的,相关处理决定是:撤销相关已获国家自然科学基金项目资助,并追回已拨经费,取消造假人员的国家自然科学基金申请资格一定年限,并给予造假人员通报批评。相关流程未督促造假人员个人、机构或期刊对造假论文进行标识并撤销。国家自然科学基金委员会监督委员会公示这些造假论文的次年(2016年),《高等学校预防与处理学术不端行为办法》^[10]发布,指明高校是学术不端行为预防与处理的主体。2018年,《关于进一步加强科研诚信建设的若干意见》^[11]更详细地指出,作者机构应是科研诚信建设的第一责任主体;学术论文若存在违背科研诚信要求的情况,应要求相应违规人员采取撤回论文等措施消除不良影响。这些办法与意见发布在发现上述造假论文之后,但无一作者机构督促造假人员对这些造假论文进行追溯撤销。此外,由于信息的不对称,多数相关期刊无法获知论文的造假行为或一些期刊对撤销论文存在抵触行为,未对论文加以撤销。

尽管撤销论文不能完全制止其学术影响力的传播,但是相关调查研究显示,撤销论文可大幅度降低其学术影响力^[12]。多重关卡失守,最终造成论文即使已被权威部门定性为存在造假行为,但仍以正常论文的形式存在于各大期刊数据库中供读者下载、阅读和引用,从而使造假论文存在一定程度的被引用。

2.3 论文被引用情感

引用情感可被分为正面引用、中性引用和负面引用^[13]。采用人工通篇阅读施引文献,结合语句本身以及前后文对引用情感进行分类,如引用文献是为自己的研究提供论证支持或者参考采用相关的实验方法等,为正面引用;引用文献是为自己的研究而客观阐述研究背景的,为中性引用;引用文献是为指出该文献的不足或者对该文献持怀疑态度或其他否定引用的,为负面引用。将11篇造假论文的34次引用情感加以分类,得到:15次(占44.12%)为正面引用,18次(52.94%)为中性引用,1次(占2.94%)为负面引用。15次正面引用多为论证支持而引用;18次中性引用多为综述文章的引用;1次负面引用并非是实验验证加以否定,而是文章^[14]讨论学术论文出版后论文信息变更中存在的问题,如多数问题是论文未被主动撤销,以该造假论文为例加以说明。

在自然科学领域,前人已发表论文中的一些重要实验,后人可能需要重复实验进行验证;而一些不是很重要的实验结果且该结果又符合自己预期的,后人一般可直接引用该结果来支持自己的研究内容,无需重复实验进行核验,否则不仅会浪费资源,而且使科研工作者的科研工作量陡增。本次调查的数据显示,过半的引用为中性引用,体现了引用文献态度较为谨慎的科研工作者仍占多数。虽然正面引用的数量也不低,但调查中未发现有强烈认同造假论文的正面引用。负面引用是一篇关于学术论文出版后论文信息变更研究论文的引用,而非实验类验证论文的否定引用。总体来看,造假论文被引情况暂未体现科学研究过程中的自我纠错功能。这除了与当前科研的自我纠错能力有待提升有关外,也与文章并非是高影响力论文有关。高影响力论文由于读者受众较多,受关注度较高,被验证的机会较多,将更容易被其他研究者的相关实验验证而发现造假行为。例如,一篇2001年发表的自然科学类高影响力造假论文,截至2018年底被引用1103次,其中591次(占53.58%)为中性引用,89次(占8.07%)为负面引用^[6]。虽然该文献负面引用的比例不高,但是也体现了科学研究过程中一定程度的自我纠错功能。与该高影响力造假论文相比,本次调查的造假论文篇均被引次数仅为3.1次,其学术影响力总

体不高,被同行验证、发现造假的机会不多。该高影响力造假论文的调查结果与本文调查结果既有相似之处,亦有不同之处。相似之处为中性引用皆超过半数;不同之处为该文负面引用占8.07%,而本文调查结果并未发现实验验证的负面引用。

3 思考与建议

3.1 造假论文应被撤销并告知具体原因

由于发现造假论文的机构(个人)、作者单位与学术期刊的沟通机制较为缺乏,我国论文撤销制度有待完善以及撤销学术不端论文的意识有待增强等诸多原因,造成一些论文即使被发现存在造假行为却未能得到及时撤销,使造假论文仍然继续传播。自然科学领域的造假论文,如未被撤销仍以正常论文的形式存在而被读者阅读、参考,不仅可能造成学术方向的误导,浪费其他科研人员的实验资源和科研精力,而且可能造成科研共同体的不信任以及公众对科研学术的不信任^[15]。如果造假论文是医学类论文,严重的话甚至可能威胁到患者的身心健康。造假论文的危害不言而喻,一经发现应尽快加以撤销。要提高学术界的学术不端论文撤销意识,并加强发现造假论文的机构(个人)、作者单位与学术期刊等的沟通交流,尽快制止造假论文学术影响力的进一步扩散。在撤销造假论文时,应刊登论文撤销声明,指出论文存在造假行为,还应在撤销声明中详细注明撤销的具体原因,如哪些方法、数据、图片等存在造假行为,让读者了解为何撤销论文;而非直接声明“该文存在造假行为”却不刊登具体造假原因,使读者无法了解论文因何种造假而撤销。此外,造假论文可能并非全文造假,也许是部分内容造假,不能对全文进行全盘否定。对一些读者来说,部分未造假内容亦有可借鉴之处,甚至是部分造假内容亦可提供研究规避参考价值。

3.2 科学研究应有质疑精神

科学研究是一个不断发展变化的过程。在科学技术发展的过程中,某些领域可能不断有新的规律和理论产生,更新甚至淘汰原有的规律和理论,以发生质的改变。科学研究不断发展的特点决定了科研工作者的科学研究站在前人科研成果基础上的同时,还应具备质疑精神。质疑精神是促进科学进步的基本动力之一。如果科研工作者在科学研究中没有质疑精神,科学技术将保持原状,无法得到任何发展。此外,受当前学术评价体系、浮躁的科研风气以及各类利益驱使,部分科研工作者违背学术规范并制造各类造假论文。尽管中央及各部委发布一系列细化的各类科研诚信政策,表明我国对建设科研诚信和整治学术不端行为的

态度与决心,但是再严厉的防范和惩处措施也不能完全杜绝学术不端行为的发生,只能将学术不端行为降低到一定程度。在这种背景下,科研工作者应具备批判的质疑精神,对前人的研究结果进行借鉴和识别,即使是一些当下学术界权威人士的科研成果,如一些实验结果一直无法重复,也应勇于提出意见相左的科研结论,而不能盲目跟随权威。如韩国曾经的民族英雄“克隆之父”黄禹锡的造假事件、美国贝尔实验室“科学神童”舍恩的造假事件、德国著名的赫尔曼-布拉赫的造假事件等^[16],造假人员皆为当时行业的权威人士。尊重客观的实验结果,发现问题并提出质疑,是捍卫科学精神的重要方式。

3.3 应警惕连环造假行为

建立在造假文献基础上的研究,或者强烈认同造假文献研究结果的研究,应警惕其学术不规范或学术不端的可能性。本次调查中,中性引用占 52.94%,多数是作为研究背景而被引用,表明过半的科研工作者对他人研究成果的引用较为谨慎;正面引用仍有 44.12%,尽管施引文献皆未出现强烈支持造假论文结果的正面引用,但对造假论文的认同也应加以适当关注,警惕造假行为的连环进行,尤其是造假论文作者本人及其硕博、同课题组成员对造假论文正面引用的文献,更应是学术造假关注的重点对象。与此类似的是,1 篇研究妊娠高血压综合征的内容质量较差的论文,选取的 66 例样本中竟有 40 例男性患者,该显而易见的造假论文却在发表后 7 年多内被其他文献引用 62 次^[8]。这些施引文献是否存在连环造假行为也应得到关注。相关科研管理机构、作者单位在调查发现造假论文后,不应止步于定性、撤销造假论文和惩处造假行为人,还应进一步追踪该作者相关论文是否涉嫌造假等学术不端行为。如有可能,追踪强烈认同或支持正面引用造假论文观点的其他文献是否涉及学术不端等,这也是发现学术不端行为的可行办法之一。

4 结束语

以国家自然科学基金委员会查处的中文造假论文为例,从引文内容分析其被其他文献引用的情况,发现造假论文由于未被及时撤销而不断被其他文献引用,虽然过半的引用为中性引用,但并未出现实验验证的负面引用,未达到自然科学领域的学术自我纠错功能。由于本文可获取的样本数量较少,并不能代表自然科学领域造假论文的整体现状,仅希望以此为鉴,科研管理机构、科研工作者、作者单位和学术期刊等相关责任者能在自己职责范围内共同努力,推动我国风清气正的科研环境的形成。

5 参考文献

- [1] 霍勒斯·弗里兰·贾德森. 大背叛: 科学中的欺诈[M]. 张铁梅, 徐国强, 译. 北京: 生活·读书·新知三联书店, 2011
- [2] 余菁, 邬加佳, 孙慧兰, 等. 科技论文数据造假的核查策略和统计学方法验证[J]. 中国科技期刊研究, 2021, 32(6): 770
- [3] 徐彤阳, 任浩然. 数字图书馆视域下学术论文图像篡改造假检测研究[J]. 现代情报, 2018, 38(7): 81
- [4] 李新根. 学术论文作者信息造假现象分析及防范对策[J]. 编辑学报, 2020, 32(4): 431
- [5] SCHNEIDER J, YE D, HILL A M, et al. Continued post-retraction citation of a fraudulent clinical trial report, 11 years after it was retracted for falsifying data[J]. *Scientometrics*, 2020, 125: 2877
- [6] 张春博, 丁堃, 王贤文, 等. 全文引文分析视角下的造假论文学术影响研究[J]. 科学学研究, 2021, 39(4): 577
- [7] 袁子晗, 张红伟. 学术不端撤销论文施引文献引用态度分析: 以哈佛大学心脏干细胞撤稿事件为例[J]. 中国科技期刊研究, 2021, 32(4): 465
- [8] 丛敏, 王景周. 期刊数据库被删除论文的引用特征: 1 篇研究妊娠高血压综合征的论文追踪分析[J]. 中国科技期刊研究, 2021, 32(3): 418
- [9] 国家自然科学基金委员会. 2015—2016 年查处的科研不端行为案件处理决定[A/OL]. [2021-12-25]. https://www.nsf.gov.cn/Portals/0/fj/fj20161212_04.pdf
- [10] 教育部. 高等学校预防与处理学术不端行为办法[A/OL]. [2021-12-28]. http://www.moe.gov.cn/srcsite/A02/s5911/moe_621/201607/t20160718_272156.html
- [11] 中共中央办公厅, 国务院办公厅. 关于进一步加强科研诚信建设的若干意见[A/OL]. [2021-12-28]. http://www.gov.cn/xinwen/2018-05/30/content_5294938.htm
- [12] DINH L, SAROL J, CHENG Y 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
- [13] 刘盛博, 丁堃, 张春博. 基于引用内容性质的引文评价研究[J]. 情报理论与实践, 2015, 38(3): 77
- [14] 王景周. 学术论文出版后信息变更的可回溯管理[J]. 中国科技期刊研究, 2019, 30(8): 846
- [15] 杨珠. 中文自然科学基金类造假论文的后期处理: 以国家自然科学基金委员会查处的科研不端行为案件为例[J]. 编辑学报, 2018, 30(6): 601
- [16] 中国科学院. 科学与诚信: 发人深省的科研不端行为案例[M]. 北京: 科学出版社, 2013

(2022-01-26收稿;2022-02-08修回)