

科技论文中值得推敲的数据

霍振响 屈李纯[†]

《西北农林科技大学学报(自然科学版)》编辑部,712100,陕西杨凌

摘要 针对编辑工作实际,秉承推敲精神,提出先化整为零、后综观全文的编辑加工策略,对摘要、引言、方法、结果和结论中的数据特征、错误类型,以及全局观视阈下论文中常见的数据前后文不一致错误,进行了实例分析,阐明隐蔽数据错误的审读过程。

关键词 科技论文;数据错误;比对法

Data worthy of scrutiny in scientific papers // HUO Zhenxiang, QU Lichun

Abstract In accordance with the practice of editorial work, adhering to the spirit of scrutiny, we propose the strategies of first breaking up the text, then reviewing the whole text, and case analysis has carried out in order to explain some data errors based on the data characteristics, types of errors, and editing methods of the abstract, introduction, methods, results and conclusions, and finally elaborate the hidden data error by checking and approving the process.

Keywords scientific papers; data error; comparison method

Authors' address Editorial Department of Journal of Northwest A&F University (Nat Sci Ed), 712100, Yangling, Shaanxi, China

DOI:10.16811/j.cnki.1001-4314.2016.05.009

唐朝著名苦吟派诗人贾岛为诗句“鸟宿池边树,僧敲月下门”中用“推”还是用“敲”苦苦思索,并引出千古流传的“推敲”故事。可见贾岛诗句用词之求精求准。数据作为科技论文的核心,贯穿论文始终,从试验设计、试验结果到讨论,无一不是靠数据来支撑,若无相关数据佐证,则终乏说服力;因此,科技论文中的数据也理当要求精准。但是,由于作者写作水平所限、疏忽大意或是关注焦点不同,其所投稿件中的数据

问题不一而足,这就要求编辑对于论文中的数据也应有“推敲”精神,在编辑加工过程中,时刻做到“心中有数”^[1],对于异常或者直觉有问题的数据学会多思多想、多查证、多与作者沟通,以确保论文刊出时数据准确无误,避免数据传播过程中出现严重的错误,导致经济损失^[2]。

对于论文中的常见数据错误类型、成因及审读方法,编辑同人进行了详细的总结和分析^[3-7];但这些研究多从数据错误归类或者方法论角度进行讨论,而编辑日常工作中接触的多是由摘要、引言、方法、结果、结论(讨论)及参考文献几部分组成的完整文稿,因此,若能就论文各组成要素中的数据特征、错误类型及审读方法进行探讨,对于数据的编辑加工则更具有指导意义和价值。因此,本文针对论文基本组成要素(由于参考文献中的数据多是出版年卷期、页码等固定数据,故不做讨论)中的数据特征、错误类型及审读方法进行研究,并对全局观视阈下论文中常见的数据前后文不一致错误进行分类阐述,以与编辑同人探讨。

1 论文有关组成部分中的数据

1.1 引言中的数据

1)数据特征。引言是围绕既定的立意选题,对相关研究现状、存在问题及研究切入点进行描述的内容;因此,引言中出现的数据多源自文献,是已经公布、发表或经调查证实的数据,用以说明相关研究成果。

2)错误类型。由于引言中的数据多源自文献,因

- [10] HU H Y,ZENG Z P,LIU J Z,et al. Key elements controlling oil accumulation within the tight sandstones [J]. Journal of Earth Science,2015,26(3):336
- [11] MATSUMOTO K,GOOSSENS S,ISHIHARA Y,et al. An improved lunar gravity field model from SELENE and historical tracking data: revealing the farside gravity features [J]. Journal of Geophysical Research,2010,115: E06007
- [12] MAZARICO E,LEMOINE F G,HAN S C,et al. GLGM-3: a degree-150 lunar gravity model from the historical tracking data of NASA Moon orbiters [J]. Journal of Geophysical Research,2010,115: E05001

- [13] 中国社会科学院语言研究所词典编辑室. 现代汉语词典 [M]. 6 版. 北京:商务印书馆,2015:826
- [14] XIN H,JIANG S Y,YANG J H,et al. Rare Earth element geochemistry of phosphatic rocks in Neoproterozoic Ediacaran Doushantuo Formation in Hushan Section from the Yangtze Gorges Area, South China [J]. Journal of Earth Science, 2016,27(2):206
- [15] WANG B G,JIN M G,LIANG X. Using EARTH model to estimate groundwater recharge at five representative zones in the Hebei Plain,China[J]. Journal of Earth Science,2016, 26(3):430

(2016-05-10 收稿;2016-07-26 修回)

此错误较少;但并不排除如作者阅读文献时曲解文意、断章取义造成引用错误,作者对原文献数据进行归纳总结时出现错误,甚至原文献数据本身错误^[4]等情况。

例如本刊接收的一篇文稿,引言中有“马铃薯产量水平只有 14.42 t/hm^2 , 还未达到世界平均水平 (16.9 t/hm^2) 的 75%”^[8]。从行文组织来看,这些数据没有什么问题;但稍加留心就会发现, $(14.42/16.9) \times 100\% = 85.3\%$, 而不是 75%。为了核实是作者引用错误还是原数据错误,编辑逐本溯源查阅了其来源文献,发现该文献中对应内容是“(中国马铃薯)单产水平只有 14.42 t/hm^2 , 还达不到世界平均水平 (16.9 t/hm^2) 的 75%”。可见原文献数据就是错误的。返修时将此问题指出,修改稿中作者将 75% 改为 85%。

3) 审读方法。引言中的数据多引自文献,若编辑发现错误,可直接将意见反馈给作者,让其核对并修改;若返回稿中未做改动,编辑则应查阅原文献核实,在排除编辑个人理解偏差的前提下,与作者沟通进行修改,防止文献数据传播过程中的二次错误出现。

1.2 方法中的数据

1) 数据特征。方法是论文的重要组成部分,是研究结果的源头,只有方法科学、设计合理,得到的结果才可能准确可靠。方法部分不仅包括试验材料、试剂仪器设备及其来源、试验地点、试验时间、区域概况,而且包括试验规模(如样本量)、设计思路、测定指标及统计方法等;因此会涉及很多数据,如来自年鉴、地方志、相关文献的概况描述类数据,试验设计数据,以及计算公式等,数据特征比较丰富。

2) 错误类型。方法中的数据错误类型较多,如经纬度颠倒、数据计量单位错误、区域概况引用数据出错、计算公式出错^[5]等。

例如本刊接收的一篇稿件中,关于研究区域概况有“乾安县全县总面积 $3\,616.6 \text{ km}^2$, 总人口 306 402 万”^[9]。初看上去没有什么问题,但是仔细斟酌就会发现,“总人口 306 402 万”,相当于一个县城人口有 30 亿多,显然是错误的。这样的数据如果不经认真推敲,

刊登出去,就会闹大笑话。

3) 审读方法。对于方法中如数据计量单位等错误,经验丰富的编辑因为积累了很多常见数据知识,因此一旦出现错误,一眼就能看出,对于确定的可以直接修改,无法确定的交由作者核实修改;而经验不足的编辑发现此类错误则比较吃力,因此在加工各类文章时,要有意主动记忆一些常见数据,丰富自己的常识库,提高业务能力。

1.3 结果中的数据

1) 数据特征。结果是论文研究的主要承载者,具有重要的承上启下的作用,其既是方法的终结,也是讨论或结论的来源;因此,结果中会出现各种各样的数据。这些数据在论文中多以图表的形式呈现,并多有对应的文字分析。文字分析中出现的数字主要有直接数据和间接数据 2 种:直接数据可以从图表中看出;间接数据则由图表中的数据数据进行二次计算或归纳得到,是对图表数据的延伸。

2) 错误类型。由于结果中数据众多,因此成为数据错误的“重灾区”。对于有对应文字分析的图表数据,直接数据错误编辑只要认真比对即可发现。间接数据的某些错误,编辑只需通过简单计算往往就可发现;但对图表进行归纳总结后用以描述相关趋势及差异等的数字,有些可以通过观察图柱或线条变化趋势判断对错,有些数据错误则比较隐蔽,需要编辑有一定的专业知识、逻辑常识和工作经验才可发现。在笔者工作过程中,就发现了如下比较隐蔽的数据错误。

① 逻辑错误。在本刊接收的一篇文稿中,结果部分有“由图 2 可知,……,Shannon-Weiner 多样性指数和均匀度随林冠开度的增大而增加,呈直线正相关关系”^[10]。

推敲过程:在二校后通读时发现文字叙述(呈直线正相关关系)与图 2(见图 1)的数据($r = -0.159\,5$ 和 $r = -0.285\,0$)不符,经查阅作者原稿发现,2 个 r 都是正值,这可能是排版人员失误而责任编辑一校时疏忽造成的。

② 违背生活常识类错误。本刊接收的一篇文稿中,结果部分有“通过调查 2011 年西安浐灞生态区周边的房价,得到消费价格见表 3”^[11]。

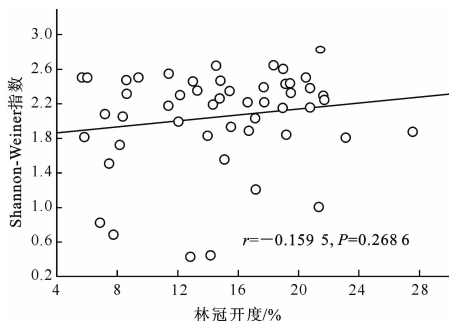


图 2 广东省银瓶山自然保护区林下物种多样性指标与林冠开度的关系

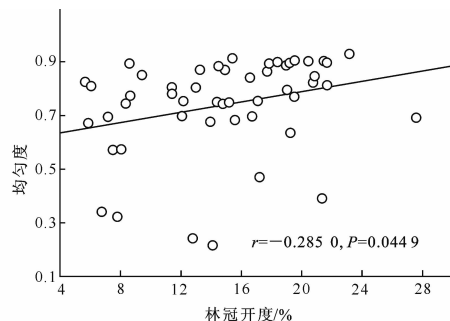


图 1 结果中的数据错误实例

表1 结果中的数据错误实例

表3 西安浐灞生态区房地产消费价格

到生态区的距离/m	距离 中值/m	房价/ (元·m ⁻²)	房屋间 差价/元
80~130	105	8 200	
130~180	155	7 200	1 000
180~230	205	6 300	900
230~280	255	5 500	800
280~330	305	5 000	500
330~		4 700	300

推敲过程:针对原表3(见表1)没有对应的文字分析,从表格来看,没有超大、超小、有效位数不一致等异常数据,因此在修改稿中并未发现问题;但笔者在一校过程中,对前2列的数据产生了疑问,与生态区距离仅80 m,如何保证生态区的正常发展?因此怀疑数据有误,在请作者校对时提出了此问题,作者校对结束后,改为表2。

3)审读方法。由于结果部分涉及数据众多,因此编辑首先要对图表中的数据逐一核查,并与对应的文字内容进行比对,保证前后一致;同时对一些间接数据,在比对核查的基础上,还应适当深究,发现其中的隐蔽性错误;对于自身欠缺的专业知识,要积极查阅相关资料,或向作者请教,与作者沟通后进行修改,以免刊印时出错。

表2 对原表3的修改结果

表3 西安浐灞生态区房地产消费价格

到生态区的距离/m	距离 中值/m	房价/ (元·m ⁻²)	房屋间 差价/元
800~<1 300	1 050	8 200	
1 300~<1 800	1 550	7 200	1 000
1 800~<2 300	2 050	6 300	900
2 300~<2 800	2 550	5 500	800
2 800~<3 300	3 050	5 000	500
≥3 300		4 700	300

1.4 摘要或讨论中的数据

1)数据特征。摘要和讨论中的数据有一定的共同点。其中,摘要独立于正文存在,是科技论文非常重要的组成部分,是读者快速了解研究内容的一个窗口,也是主要的检索对象。摘要中的数据,一般直接或间接来自正文结果;讨论是对结果的归纳总结,是研究问题得以解决的落脚点,其中的数据,有些是源自结果或对结果数据的延伸,还有些则是引自同类文献,以说明研究差异或创新性。

2)错误类型。摘要或讨论中源自结果的数据,主要错误形式就是与文中其他部分不一致。由于数据不一致是论文中比较常见的一种错误形式,因此另立一节展开叙述;对源自文献的数据,一方面要核对能否说

明相应观点,另一方面则应如引言中数据一样注意引用数据的合理性和准确性。

3)审读方法。对于摘要或讨论与文中不一致的数据,可以通过比对发现错误;对于源自文献的数据错误,则可采取引言中的数据审读方法发现并予以解决。

2 全局观视阈下的论文数据

综观科技论文核查全文数据^[12],是编辑加工的必经步骤。因为同一组数据可能散见于论文各相应部分中;因此在综观全文核查数据时,常见的错误就是同一组数据前后文不一致,如中英文摘要不一致、方法或结论与结果不一致、摘要与正文中不一致,等等。

同一组数据前后文不一致的错误虽然比较常见,但只要编辑在修改或校对文稿过程中,对反复出现的同一组数据进行全文比对,即可发现并解决。对于数据前后文不一致的错误,编辑同仁也曾有探讨^[5,7],本文简要列举几种。

1)中英文摘要不一致。本刊接收的一篇文稿中文摘要要有“最大理论预测值(6.63%)”^[13],而英文摘要中对应数据却是6.72%。经过全文核查比对,发现英文摘要中出现错误,应为6.63%。

2)试验方法、结论与结果不一致。在本刊接收的《地鳖虫提取物对大鼠运动能力和骨骼肌抗氧化酶活性的影响》^[14]一文中,有关试验设计的描述,原稿中其他处都是“按大鼠体质量设为0.5,1.0,1.5 g/kg”,仅1.2.1节是“0.5,1.0和2.0 g/kg”,显然是作者疏忽导致。可将1.2.1节直接改为0.5、1.0和1.5 g/kg。

在另一篇文稿中,对于目的基因的扩增,方法部分的描述为“采用20 g/L琼脂糖凝胶检测PCR扩增产物,并用琼脂糖凝胶试剂盒回收、纯化目的片段”^[15];但在对应结果中,却是“经10 g/L琼脂糖凝胶分析,PCR扩增的BPE 123和BPE 275片段的长度分别为462和762 bp,与预期结果一致”。同一数据前后文不一致,但对此错误编辑无法确定孰对孰错。经与作者讨论,确定均应为“20 g/L”。

3)其他不一致。本刊接收的一篇文稿,引言中有“ORF3编码基因包括ORF3a和ORF3b 2个基因,不同毒株的ORF3b都为744 bp,而ORF3a因毒株的不同而长度不等”^[16],而方法中相关内容为“……用于扩增ORF3a、ORF3b基因全长,长度分别为219和735 bp”。前句是总结性描述,后句是针对具体毒株TGEV;但从逻辑上来看,既然“不同毒株的ORF3b都为744 bp”,为何TGEV ORF3b长度却是735 bp呢?对此疑问请作者核对,作者反馈结果是引言中错误,应为735 bp。

对于同一组数据全文不一致的错误,最常用且最有效的审读方法就是比对法。只要编辑牢记,见到文中多处出现的同一组数据,无论散见于何处,即进行比对,错误基本就可消除。比对法,是编辑初学阶段比较容易上手和掌握的一种方法,不仅对于核查数据有好的效果,在其他内容编校中也能发挥很大的作用。

3 结束语

在科技论文数据的编辑加工中,采取“化整为零,综观全文”的加工策略,不仅可以提高编辑尤其是新入职编辑的工作效率,而且可以最大程度地解决论文中出现的数据错误。具体而言,对一篇文稿数据的编辑加工,可以采用如下步骤。

1)化整为零,以小节为单位进行加工。根据不同组成要素中的数据特征,以恰当的方法如计算法、比对法、常识法、逻辑推理法等,有针对性地进行推敲,以小节为单位进行编校加工也可以强化数据外其他内容的编校质量。

2)综观全文,消灭同一内容前后文不一致的错误。在化整为零逐节处理之后或处理过程中,还应利用比对法对同一数据进行前后文核查;因为论文中比较常见的数据错误类型就是前后文不一致,且这类错误不仅在数据中存在,在同一概念、药品、基因符号、试验处理、结论等内容中也都有可能存在。

“化整为零,综观全文”的加工策略不仅有利于全盘把握论文中的统计数据,通过审读推敲消除其中错误,对其他内容的编辑加工也有一定的实用性;同时,由于各行各业数据表述方式不同,尚无统一的标准和规范可以参考,编辑在处理论文中的数据时,始终要保持对数据的敏感度,通过单一数据核查、同一组数据比对,消除数据错误,确保刊发论文的质量。

4 参考文献

- [1] 陈浩元.科技书刊标准化18讲[M].北京:北京师范大学出版社,1998:285
- [2] 游俊,胡小洋,赵燕.科技论文中的数据问题及原因分析

- [J].江汉大学学报(自然科学版),2012,40(4):135
- [3] 张红林,邓薇,李寿星.科技论文中数据错误的识别与防范[J].武汉科技大学学报(社会科学版),2006,8(5):140
- [4] 朱久法,张彩虹.学术论文中常见的一些数据问题及对编辑工作的要求[J].中国科技期刊研究,2010,21(4):546
- [5] 武英耀.科技论文常见数据错误类型分析[J].编辑学报,2015,27(6):149
- [6] 游俊,赵燕,胡小洋.科技论文中数据问题的编辑审读技巧[J].编辑学报,2012,24(6):536
- [7] 钟羨芳.科技期刊编辑甄别数据错误的方法[J].科技与出版,2015(8):49
- [8] 徐胜涛,刘景辉,张磊,等.土壤改良剂对马铃薯水分生产效率、产量及品质的影响[J].西北农林科技大学学报(自然科学版),2012,40(11):48
- [9] 张清华,韩梅,杨利民.1949—2008年吉林省乾安县能值生态足迹的动态研究[J].西北农林科技大学学报(自然科学版),2012,40(5):174
- [10] 王永强,蔡燕茹,曾焕忱,等.不同林冠开度下亚热带林下植物的组成和多样性[J].西北农林科技大学学报(自然科学版),2016,44(5):68
- [11] 高鑫,解建仓,汪妮,等.基于物元分析与替代市场法的水资源价值量核算研究[J].西北农林科技大学学报(自然科学版),2012,40(5):229
- [12] 于永田.科技论文中的数据舛误与编辑加工[J].编辑学报,1993,5(3):143
- [13] 苏晨曦,陈文强,彭浩,等.微波辅助提取香菇多糖工艺的响应面优化[J].西北农林科技大学学报(自然科学版),2015,43(11):200
- [14] 高颖辉,周万红,马贺,等.地鳖虫提取物对大鼠运动能力和骨骼肌抗氧化酶活性的影响[J].西北农林科技大学学报(自然科学版),2015,43(11):24
- [15] 孙志华,张辉,刘来珍,等.布鲁氏菌IV型分泌系统效应分子BPE123和BPE275基因的原核表达与鉴定[J].西北农林科技大学学报(自然科学版),2016,44(5):2
- [16] 曹贝贝,兰培英,韩丽,等.猪TGEV HN-2012株ORF3a和ORF3b基因遗传变异分析及其真核表达研究[J].西北农林科技大学学报(自然科学版),2016,44(4):11

(2016-05-05 收稿;2016-06-02 修回)