Declining citation accuracy in polar research

T. McIntyre1 and N.S. Haussmann2

1Department of Life and Consumer Sciences, College of Agriculture and Environmental Sciences, University of South Africa-Science Campus, Florida, South Africa and 2Department of Geography, Geoinformatics and Meteorology, University of Pretoria, Hatfield, South Africa

Abstract

Accurate citation practices are important to ensure a robust knowledge base and overall trustworthy academic enterprise. The prevalence of poor citation practices has been assessed in multiple fields, resulting in estimates of inaccurate citations ranging typically between 15% and 25%. Here, we assessed the accuracy of citations in research articles extracted from 11 journals with a polar sciences focus. Thirty percent of citations from recent articles (published between 2018 and 2019) and 26% of citations between 1980 and 2019 were found to be inaccurate. We found no evidence for differences in citation accuracy between the journals assessed, or effects on citation accuracy associated with the number of authors, number of references, position of references or if a citation was a self-citation or not. Importantly, we present evidence for a decline in citation accuracy between 1980 and 2019 in polar sciences. Citation practices are unlikely to improve unless journals provide incentives for scholars to be more meticulous, and we recommend active monitoring of citation accuracy and citation appropriateness by reviewers and editorial staff.

Introduction

Science advances in an incremental fashion, whereby researchers typically rely on published research to inform their own work. Referring to previously peer-reviewed, published research is important not only to acknowledge the existing knowledge base but to provide support for assumptions, methods used, conclusions reached and arguments put forward. When citations are not accurately used and do not support the statements they are purported to, it not only presents the cited work in an unfair manner that misleads the reader but may lead to the spread of unsubstantiated information. It is therefore imperative that existing knowledge is portrayed accurately when cited to facilitate a robust knowledge base from which science can advance.

Citation inaccuracies can take many forms, ranging from easily detectible bibliographic errors to more significant errors whereby important information is misinterpreted or misrepresented (Martella et al., 2021). When errors misrepresented published data, quotes, claims and/or interpretations, they are termed “quotation errors” to distinguish these from the broader group of inaccurate citations that include bibliographic and typographical errors (Martella et al., 2021; Smith & Cumberledge, 2020). The accuracy of citation practices has been assessed for a number of fields, notably in many medical fields (reviewed in Jergas & Baethge, 2015), ecology (e.g. Drake et al., 2013; Todd, Guest, Lu, & Ming Chou, 2010; Todd, Yeo, Li, & Ladle, 2007), physical geography (e.g. Haussmann, McIntyre, Bumby, & Loubser, 2013), learning sciences (e.g. Martella et al., 2021) and, more recently, general high-impact science journals (e.g. Smith & Cumberledge, 2020). These findings are typically worrisome, with many studies suggesting that more than 20% of the citations in their respective fields are inaccurate (e.g. Jergas & Baethge, 2015; Martella et al., 2021; Todd et al., 2010).

Importantly, studies reporting temporal trends in citation practices are seemingly constrained to the medical literature. Here, Buijze et al. (2012) reported a decline in citation accuracies over a 10-year period for papers specific to orthopaedics, while Jergas and Baethge (2015) undertook a meta-analysis of citation practices papers in the medical field and reported a non-significant increase in citation error rates over time. Previous studies analysing citation practices in other fields typically provide snapshots in time, presenting results from assessments of citations made during a single year or small time frames of no more than two to three years (e.g. Drake et al., 2013; Haussmann et al., 2013; Smith & Cumberledge, 2020).

Various factors have previously been assessed as potential predictors of citation accuracy, often with contrasting outcomes between field-specific studies on their likely influence. For example, occasional reports of differences in citation accuracy between journals (e.g. Davids, Weigl, Edmonds, & Blackhurst, 2010; Reddy, Srinivas, Sabanayagam & Balasubramanian, 2008; Schulmeister, 1998) contrast with others that report no such differences within other fields (e.g. Buijze, Weening, Poolman, Bhandari, & Ring, 2012; Hansen & McIntyre, 1994; Smith & Cumberledge, 2020). Similarly, journal impact factors have been reported to be related to citation accuracy (e.g. Davids et al., 2010; Haussmann et al., 2013), while others have reported no evidence for such a relationship (e.g. Buijze et al., 2012; Todd et al., 2010). Other factors showing
similar non-consistent relationships with citation accuracy include the number of authors (e.g. Buijze et al., 2012; Reddy et al., 2008; Todd et al., 2010), number of references in the article (e.g. Buijze et al., 2012; Eichorn & Yankauer, 1987; Reddy et al., 2008; Schulmeister, 1998; Todd et al., 2010), as well as if a citation formed part of a string of citations or not (e.g. Buijze et al., 2012; Smith & Cumberledge, 2020).

Polar science is by its nature a broad term only limited by the geographical scope of applying to the polar regions, but encompassing multiple disciplines ranging from physical – and life sciences to the social – and political sciences. Despite the range of research fields encompassed, many polar scientists identify as being part of such an interdisciplinary polar sciences community. This is perhaps best exemplified through efforts such as the very successful International Polar Year (IPY) that brought together scientists and educators to contribute to a large number of multinational projects investigating issues dealing with earth, land, people, oceans, ice, the atmosphere, space, education and outreach under a single umbrella (Carlson, 2019), and resulting in the establishment and continued successes of organisations such as the Association of Polar Early Career Scientists (APECS) (Hindshaw et al., 2018).

Here, we assessed the citation practices by the polar sciences community. Specifically, we investigated the appropriateness of citations used in polar sciences, and which factors help explain the appropriateness of citations.

Methods

We broadly followed the approach of Todd et al. (2007) and selected a series of articles from journals with a polar research focus, from which we randomly selected citations to verify. We generated two independent datasets: one dataset to assess current citation practices and potential correlates; and a second dataset to assess temporal trends in citation practices.

To be included in our study, a journal had to (1) have a polar focus (i.e. explicitly publishing research related either to the Arctic, Antarctic or both), (2) have a current ISI (Institute for Scientific Information, Clarivate Analytics) rating and (3) be accessible via either the University of South Africa and/or University of Pretoria libraries. We further excluded journals in specialised and technical fields that we could not reliably assess, given our own academic expertise and experience (e.g. International Journal of Offshore and Polar Engineering). Accordingly, we identified 11 journals with a polar focus. Journals included in our analyses were Antarctic Science; Arctic; Arctic, Antarctic and Alpine Research; Arctic Anthropology; Cryosphere; Permafrost and Periglacial Processes; Polar Biology; Polar Record; Polar Research; Polar Science; and Polish Polar Research. From these journals, we selected the first and the last research article from each of the three most recent issues prior to 2020 that were accessible via either the University of South Africa and/or University of Pretoria libraries. Therefore, a total of 66 (11 × 2 × 3) articles were selected as primary articles to assess current citation practices. From each of these primary articles, a reference (from here on referred to as the secondary article) was randomly selected from the reference list using a random number generator to select the reference. The statement that the reference was supporting was searched for within the primary article’s text and retained if it was supported by a single citation and the reference article obtainable via either of the university libraries. If either of these criteria were not met, another reference was selected randomly from the reference list until we found a reference that met these requirements. When a selected reference was used multiple times in the text of the primary article to support various statements, we randomly selected from statements where the secondary article was the sole citation.

The secondary articles were read by both authors of this article (TM and NSH), who independently classified its appropriateness according to four categories. Citations were classified as offering clear support (category 1) if the cited article provided unambiguous support of the statement, either via statements in the text of the cited article or the results presented. Citations that did not corroborate the statement in the primary article via either statements in the text of the cited article or the results presented in the cited article were classified as offering no support (category 2). In these cases, the cited article could even contradict the assertion in the primary article. Ambiguous citations (category 3) were identified as such when either of the following were considered to best describe the citation:

- the cited article had been interpreted in one way but could also be interpreted in other ways, including the opposite point;
- the primary article was supported only by a section of the cited article – however that section was deemed contrary to the overall direction of the cited article;
- the primary article statement includes two or more components, but the cited article provides support for only one of them;
- it was unclear what statement was supposed to be supported (i.e. if there was no explicit statement in the phrase preceding a citation);
- or if, instead of the cited paper supporting the author’s statement directly, the cited article provides an example of the author’s statement without the reference indicating it as such (i.e. by preceding with “e.g.”). This did not always lead to ambiguity and was only classified as such if it was unclear if the citation was intended as an example of such a study/result or if it directly supported the primary article statement.

Lastly, citations were classified as empty citations (category 4) if the cited article simply cites other articles that support the assertion made in the primary article and does not provide support for the statement through its own findings. Citing a review article was not considered to be an empty citation if the support for the assertion was through a new insight or opinion offered by the author(s) of the review.

When classifications did not correspond, we selected the classification that was most in favour of the primary authors. In other words, if one assessor scored the citation as offering clear support, it was automatically recorded as an accurate citation (category 1), irrespective of the other assessor’s score. Both assessors needed to score a citation as unsupported (category 2) for it to be recorded as such, with the score defaulting to ambiguous (category 3) if there was disagreement and neither of the assessors scored it as offering clear support (i.e. if the citation was scored either as 2 and 4, or 3 and 4 by the two assessors). Similarly, both assessors needed to score a citation as empty (category 4) for it to be recorded as such.

To test for any possible predictors of citation accuracy, we further recorded for each of the 66 primary articles: the journal where published; the number of authors; the number of references; the position of the assessed citation (e.g. Introduction, Methods, Results, Discussion or other); as well as if it was a self-citation or not (i.e. at least one of the authors of the primary article was a co-author on the secondary article). We used binomial generalised linear models (GLMs) to assess any relationships between these variables and whether a citation was deemed appropriate
Table 1. Citation accuracy classifications (reported as percentages) for 222 published citations in polar sciences journals. Classifications are separated between current (“2018/19 only”) citations and “1980–2019” citations.

| Citation category           | 2018/19 only (11 journals, n = 66) | 1980–2019 (4 journals, n = 156) |
|----------------------------|-----------------------------------|----------------------------------|
| Clear support              | 69.7%                             | 73.7%                            |
| No support                 | 4.5%                              | 3.2%                             |
| Ambiguous support          | 16.7%                             | 20.5%                            |
| Empty citation             | 9.1%                              | 2.6%                             |

or not (categories of not appropriate citations having been grouped for this purpose such that citations were either deemed appropriate or not appropriate).

To assess temporal trends in citation practices, we additionally sampled articles from four journals where we were able to access articles published over a period approaching four decades (early 1980s until 2019). The four selected journals for this part of the analyses were *Arctic, Antarctic and Alpine Research* (1980–2019); *Polar Biology* (1982–2019); *Polar Record* (1980–2019) and *Polar Research* (1982–2019). Here, we selected one article per year per journal, identified as being the middle article from the first issue of each year. In cases where there were an equal number of articles in a volume, we again used a random number generator to pick one article between the two middle articles. Therefore, a total of 156 articles (40 + 38 + 40 + 38) were selected here to assess temporal trends in citation practices. Here again we selected a reference (secondary article) randomly from the reference list and classified its appropriateness. To assess temporal trends in citation accuracy, we also used a binomial GLM to model the relationship between the likelihood of a citation being appropriate and the year in which the article was published.

All analyses were undertaken in the R environment (R Version 4.0.3; R Development Core Team, 2020), with statistical significance set at \( p < 0.05 \), and summaries reported as means ± standard deviation, unless otherwise noted.

Results

Primary papers were authored by a median number of three authors (mean = 3.5 ± 2.5) and cited a mean number of 51.8 (± 31.5) references. Assessed citations were positioned most often in the Discussion (\( n = 69 \)), followed by the Introduction (\( n = 49 \)), Methods (\( n = 48 \)) and other (\( n = 48 \)), Results (\( n = 5 \)) and Conclusion (\( n = 3 \)). Twenty-one out of the 222 citations were recorded as self-citations where at least one of the authors was also the author or co-author of the cited work.

Overall, there was 81% agreement (180 out of 222) between the citation accuracy classifications of the two assessors. We found clear support for 46 out of 66 current (i.e. 2018 and 2019) citations (i.e. 70%) across the 11 journals sampled, and clear support for 115 out of 156 citations (74%) across the 4 journals considered between 1980 and 2019 (Table 1). Citations that did not offer clear support were most often classified as providing ambiguous support (\( n = 43 \)), followed by empty citations (\( n = 10 \)) and eight citations were classified as providing no support (Table 1).

Model outputs provided no evidence for significant relationships between the citation accuracy (i.e. if clearly supported or not) and journal (\( X^2 = 10.9_{10,65}, \ p = 0.37 \)), number of authors of primary article (\( X^2 = 0.09_{1,65}, \ p = 0.76 \)), number of works cited by primary article (\( X^2 = 0.02_{1,65}, \ p = 0.89 \)), nor the position of the citation (\( X^2 = 7.27_{5,65}, \ p = 0.2 \)). Similarly, we found no evidence for a relationship between citation accuracy and whether citations were self-citations or not (\( X^2 = 0.15_{1,65}, \ p = 0.7 \)).

Year of publication was related to citation accuracy (\( X^2 = 4.28_{1,55}, \ p = 0.04 \)), and the probability of a citation being accurate decreased between the early 1980s and 2019 (Fig. 1).

Discussion

Our results suggest that between 26% (1980–2019) and 30% (2018/2019 only) of citations in polar sciences journal articles are not accurate. Such an overall rate of inaccuracy is similar to overall rates of approximately 25% reported for ecology (Todd et al., 2007), marine ecology (Todd et al., 2010), medical research (Jergas & Baethge, 2015; Mogull, 2017) and top-ranked general science journals (Smith & Cumberledge, 2020). Quotation errors specifically (i.e. no support and ambiguous support) were evident in 24% (1980–2019) and 21% (2018/2019) of the citations. These error rates fall within the ranges reported for other fields (e.g. learning sciences 26% (Martella et al., 2021), general science 25% (Smith & Cumberledge, 2020)) and seem to suggest such error rates to be relatively consistent in recent years between various academic fields.

Similar to the results reported elsewhere (e.g. Haussmann et al. 2013; Todd et al. 2010), we believe our results are somewhat conservative for two reasons. Firstly, because we gave the benefit of the doubt to the authors when ratings did not agree (19% of the cases), which probably resulted in fewer quotation errors being identified. Secondly, while the potential influence of restricting our analyses to single citations and excluding string citations was not tested, we initially expected that authors would likely be more conscientious when using a single reference to support a statement, compared to using a string of citations. This argument was raised in Todd et al. (2010) and Haussmann et al. (2013) and supported by the results of Buijze et al. (2012) who reported most inaccurate citations in orthopaedic journals to be present in string citations. However, Smith and Cumberledge (2020) recently reported the opposite – an increased likelihood of a citation being accurate if used as part of a string of citations, as opposed to singularly. It should be noted here that they required only a string of citations in its totality to support a given statement, and not for individual citations in a string to independently support the proposition, which would have increased the likelihood of string citations to provide full support to statements (Smith & Cumberledge, 2020).

Our results did not provide any support for differences in citation accuracy associated with the journal, number of authors, number of references, or if the citation was a self-citation or not. This is also broadly in agreement with other studies that did not find differences in citation accuracies associated with such variables (e.g. Jergas & Baethge, 2015; Smith & Cumberledge, 2020; Todd et al., 2010) – but see Buijze et al. (2012) who reported a slight correlation between the number of authors and the likelihood that a citation was inaccurate in the orthopaedic literature, and Haussmann et al. (2013) who reported a slight difference in citation accuracy linked to the impact factor of journals in the field of physical geography. The equal likelihood of self-citations and other citations being accurate is surprising to some extent, given that authors are presumably less likely to miscite their own publications. Nonetheless, out of the 21 self-citations we identified, 2 did not provide any support for the statements they referred...
to. The presence, albeit low prevalence, of such findings are similar to those reported for journals in physical geography (Haussmann et al., 2013) and are of concern due to the possibility of authors deliberately misciting their own work.

Perhaps the most concerning finding in our results is the decline in accuracy of citations between the early 1980s and now (Fig. 1). We are only aware of a single other publication that reported a temporal trend in citation practices – Buijze et al. (2012) also reporting a decrease in citation accuracy in orthopaedic papers between 2000 and 2009. The potential reasons for such declines in citation accuracy are numerous. Possibly one of the simplest explanations for this trend is that authors are increasingly hasty and under pressure to complete manuscripts given increased pressures associated with competitiveness in academia, leading to increasingly sloppy practices when citing publications (Todd et al., 2007). Furthermore, we speculate that a potential increase in the proportion of research outputs from universities authored by students (as opposed to principal investigators) (Al-Busaid & Al-Shaqsi, 2015; Andersen, Østergaard, Fosbol, & Fosbol, 2015; Kan et al., 2021) may contribute to increased inaccurate citation practices. Students are mostly inexperienced in academic writing and the publication process, and mentors may find it difficult to verify each student’s references when advising multiple students at any given time. While our results provided no support for a link between citation accuracy and the number of references, it is worth noting that there was an increasing trend between 1980 and 2019 in the number of references cited by papers in our sample (results not shown). The increasing pressure on academics to generate publications and to garner citations of their own work may of course also fuel deliberate malpractice when it comes to citations (see examples in Lockwood, 2020).

Regardless of the reasons for inaccurate citation practices, the effects are deleterious to science as a whole and undermine the overall trustworthiness of the scientific process – something which perhaps requires more protection now than ever before given widespread commercial and political pressures seeking to increase scepticism of science (Druckman, 2017). Various approaches to improve citation practices have been proposed by authors. These are typically separated by approaches applicable to editors and publishers vs approaches applicable to authors, and have been summarised by others (e.g. Jergas & Baethge, 2015). Perhaps the most thorough of these interventions is to include exhaustive technical editing that includes checking the appropriateness of citations as part of the editing process between acceptance of an article and its publication. This will likely lead to improved citation practices (e.g. Wager & Middleton, 2002), but practical implementation is likely out of reach of most journals in terms of available technical capacity. Random audits by journals (as proposed in Todd et al., 2007) and requesting ad hoc checks to be carried out by reviewers are likely more practical and can be facilitated by journals requiring authors to include page numbers with in-text citations (Smith & Cumberledge, 2020).

Some have argued that the responsibility ultimately lies with the authors to cite appropriately (Todd et al., 2010), and that citation practices may improve through appropriate training and the building of an ethical culture (Drake et al., 2013). While we broadly agree with these sentiments, there is no evidence for improvement in citation practices, despite multiple publications during the last two decades highlighting the problem. We therefore do not expect to see any broadscale improvement in citation practices without active, explicit encouragement and enforcement of adequate citation practices by all parties involved in the reviewing process. Ultimately, we encourage journals to actively monitor citation practices and explicitly encourage proper citations by making use of interventions such as ad hoc checks and occasional audits at a minimum, and where feasible, including reviews of citations as part of the technical review and editing processes.

Acknowledgements. The authors are grateful to two anonymous reviewers for their constructive comments that improved this paper. Access to the relevant

Fig. 1. Predicted probability of citations offering clear support as function of year of publication. Dots indicate individual citations informing the predictive GLM and are offset from one another on the figure.
literture used in our analyses was provided through the University of South Africa and the University of Pretoria.

Financial support. This research received no specific grant from any funding agency, commercial or not-for-profit sectors.

Conflict of interest. None.

References

Al-Busaid, I. S., & Al-Shaqsi, S. Z. (2015). Students’ contribution to the New Zealand medical journal: A 14-year review. New Zealand Medical Journal, 128(1412), 47–52.

Andersen, S. B., Østergaard, L., Fosbøl, P. L., & Fosbøl, E. L. (1994). Reference citations in radiology: An old problem revisited. Surgery, 6(2), 71–75. doi: 10.1016/0346-549. doi: 10.1371/journal.pone.0184727

Davids, J. R., Weigl, D. M., Edmonds, J. P., & Blackhurst, D. W. (2010). Accuracy of references in general science journals. PLoS ONE, 12(9), 1–17. doi: 10.1371/journal.pone.0184727

Druckman, J. N. (2017). The crisis of politicization within and beyond science. Nature Human Behaviour, 1(9), 615–617. doi: 10.1038/s41562-017-0183-5

Eichorn, P., & Yankauer, A. (1987). Do authors check their references? A survey of accuracy of references in three public health journals. American Journal of Public Health, 77(8), 1011–1012. doi: 10.2105/ AJPH.77.8.1011

Hansen, M. E., & McIntire, D. D. (1994). Reference citations in radiology: Accuracy and appropriateness of use in two major journals. American Journal of Roentgenology, 163(3), 719–723. doi: 10.2214/ajr.163.3.8079876

Haussmann, N. S., McIntyre, T., Bumby, A. J., & Loubser, M. J. (2013). Referencing practices in physical geography: How well do we cite what we write? Progress in Physical Geography, 37(4), 543–549. doi: 10.1177/0309133313482135

Hindshaw, R. S., Marisah, H., Vick-Majors, T. J., Thornton, A. E., Pope, A., Zaika, Y., . . . Fugmann, G. (2018). A decade of shaping the futures of polar early career researchers: A legacy of the International Polar Year. Polar Record, 54(3–6), 312–323. doi: 10.1017/S0032247418000591

Jergas, H., & Baethge, C. (2015). Quotation accuracy in medical journal articles-A systematic review and meta-analysis. PeerJ, 2015(10), 1–20. doi: 10.7717/peerj.1364

Kan, C. K., Qureshi, M. M., Paracha, M., Sachs, T. E., Sarfatty, S., & Hirsch, A. E. (2021). Effect of medical student contributions on academic productivity: Analysis of student authorship over time. Advances in Medical Education and Practice, 12, 481–489. doi: 10.2147/AMEP.S301041

Lockwood, M. (2020). Editorial: Citation malpractice. Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences, 476(2242), 0–2. doi: 10.1098/ rspa.2020.0746

Martella, A. M., Yaculla, J. K., Martella, R. C., Marchand-Martella, N. E., Ozen, Z., Karatas, T., . . . Karpicke, J. D. (2021). Quotation accuracy matters: an examination of how an influential meta-analysis on active learning has been cited. In Review of Educational Research, 91. doi: 10.3102/0034654321991228

Mogull, S. A. (2017). Accuracy of cited “facts” in medical research articles: A review of study methodology and recalculation of quotation error rate. PLoS ONE, 12(9), 1–17. doi: 10.1371/journal.pone.0184727

Reddy, M. S., Srinivas, S., Sabanayagam, N., & Balasubramanian, S. P. (2008). Accuracy of references in general surgical journals – An old problem revisited. Surgeon, 6(2), 71–75. doi: 10.1016/S1479-666X(08)80067-4

Schulmeister, L. (1998). Quotation and reference accuracy of three Nursing journals. Journal of Nursing Scholarship, 30(2), 143–146. doi: 10.1111/j. 1547-5069.1998.tb01269.x

Smith, N., & Cumberledge, A. (2020). Quotation errors in general science journals. Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences, 476(2242). doi: 10.1098/rspa.2020.0538

Todd, P. A., Guest, J. R., Lu, J., & Ming Chou, L. (2010). One in four citations in marine biology papers is inappropriate. Marine Ecology Progress Series, 408, 299–303. doi: 10.3354/meps08587

Todd, P. A., Yeo, D. C. J., Li, D., & Ladle, R. J. (2007). Citing practices in ecology: Can we believe our own words? Oikos, 116(9), 1599–1601.

Wager, E., & Middleton, P. (2002). Effects of technical editing in biomedical journals: A systematic review. Journal of the American Medical Association, 287(21), 2821–2824. doi: 10.1001/jama.287.21.2821