Referencing practices in physical geography: how well do we cite what we write?

Natalie S. Haussmann

University of Pretoria, South Africa

Trevor McIntyre

University of Pretoria, South Africa

Adam J. Bumby

University of Pretoria, South Africa

Michael J. Loubser

University of Pretoria, South Africa

Corresponding author:

Natalie S Haussmann, Department of Geography, Geoinformatics and Meteorology, University of Pretoria, Private Bag X20, Hatfield, South Africa.

Email: natalie.haussmann@up.ac.za

Tel: +27 (21) 420 4049

Abstract

Accurate citation practices are important, both from an ethical and scientific point of view. Using an easily reproducible, previously published method, we assess citation accuracy in 120 articles published in the first half of 2011 and listed under 'Physical

Geography' in Thomson Reuters' ISI Web of Knowledge Science Citation Index. Our results indicate that at least 19% of citations in physical geography do not provide clear support for the statements they are meant to support. These results are in line with previously published findings for "field-orientated" sciences. We propose that both authors and editors help remedy this problem, by employing more rigorous writing and editing practices.

Keywords : bibliometrics, citation, misconduct, physical geography, referencing

I Introduction

As many an academic who has been miscited can attest, sloppy citation practices can be annoying. Miscitation can take various forms, ranging from giving credit where it is not due to misconstruing someone's findings. Although some scientists undoubtedly welcome any citations, regardless of their accuracy, inaccurate citations provide a disservice, both to those whose work is being misrepresented and to the readers of the article, who are being misled (Gavras, 2002). Citations are meant to highlight the existing knowledge base, enabling the researcher to build thereupon. Rigorous citation practices thus facilitate scientific progress within a field. Citing correctly is therefore important both from an ethical and a scientific point of view.

A number of recent publications have looked at various aspects of publication and citation within academia. For example, Sin (2011) looked at the impact of co-authorship on citation frequency, whereas Bar-Ilan (2008) critically analysed the h-index, a now widely used index of author citation frequency. More specifically within geography, Slyder et al. (2011) compared citation patterns and lifespans between forestry and geography, and Schuermans et al. (2010) analysed publication trends in human geography.

Citation accuracy, i.e. the extent to which citations support the assertions they are meant to support, has, however, predominantly been quantified within the health sciences (e.g. Fenton et al., 2000; Gosling et al., 2004, but see also Todd et al., 2010). Findings within the medical fields show that the percentage of accurate citations range from more than 90% to well below 70%, depending on the subdiscipline (see Table 2, Results). Various potential reasons have been suggested for citation inaccuracy, including hasty write-up (Gavras, 2002), inadequate full-text access to journals (Todd et al., 2007) and language-related issues (Fenton et al., 2000).

Using an easily reproducible methodology that can be applied to any academic field, citation accuracy has been quantified in ecology (Todd et al., 2007) and marine biology (Todd et al., 2010). As far as we know, no one has conducted a similar study for

physical geography. The question therefore remains: how well do physical geographers cite?

II Research Methodology and Data

To answer this question and ensure that our results were comparable with those of other fields (viz. ecology and marine biology), we followed the method of Todd et al. (2007). A search was conducted for journals with an impact factor greater than one, listed under 'Physical Geography' in the Thomson Reuters' ISI Web of Knowledge Science Citation Index. From these, all journals accessible via the University of Pretoria online search network were selected. In total 20 journals matched all search criteria. From these journals, the first, the middle and the last research article from each of the two most recent issues prior to July 2011 were selected. Therefore, a total of 120 (20 x 2 x 3) articles were selected as primary articles. From each of these primary articles, a reference (from hereon referred to as the secondary article) was randomly selected from the reference list and the statement that the reference was supporting was searched for within the primary article's text. Only references from journals held by the University of Pretoria were selected. Furthermore, only assertions supported by a single citation were included. The secondary articles were downloaded and read by all four authors of this article, who subsequently classified its appropriateness according to four categories

(see Table 1). The final decision on a citation's appropriateness was determined by the classification of the majority, with the verdict being in favour of the primary article's authors if there was any doubt (Todd et al., 2007).

Table 1. Definitions of citation categories (adapted from Todd et al. 2007).

Category	Definition
Clear support	The cited article provides unequivocal support of the assertion, via either statements in the text of the cited article or the data presented in the cited article.
No support	The cited article does not in any way substantiate the assertion via either statements in the text of the cited article or the data presented in the cited article. The cited article may even contradict the assertion in the primary article.
Ambiguous	a.) The material (either text or data) in the cited article has been interpreted one way, but could also be interpreted in other ways, including the opposite point. b.) The assertion of the primary article is supported by a portion of the cited article, but that portion runs contrary to the overall thrust of the cited article. c.) The assertion includes two or more components, but the cited article only supports one of them.
Empty citation	Also called "lazy author syndrome" (Gavras 2002). The cited article simply cites other articles that support the assertion made in the primary article. Citing a review article is acceptable if the support for the assertion is, for example, a new insight or opinion offered by the author (s) of the review.

Next, the number of authors and the number of references in the primary article were noted, as well as the Impact Factors (ISI, 2010) of both the primary and secondary articles' journals and whether or not the selected citation was a self-citation. Lastly, the h-index of the first author was recorded from ISI Web of Science Author Citation Report, as well as whether or not the first author was a native English speaker (based on surname and affiliation). Following Todd et al. (2007), chi-square tests and Pearson's correlations were used to assess whether associations existed between any of these variables and citation accuracy. We also tested the overall agreement between all four raters using an Exact Fleiss' Kappa statistic (Conger 1980; Fleiss 1971), and agreement was assessed according to the classes proposed by Viera and Garrett (2005). Accordingly, a Kappa statistic of < 0.2 was considered to indicate slight to no agreement; 0.21 – 0.4 fair agreement; 0.41 – 0.6 moderate agreement; 0.61 – 0.8 substantial agreement; and > 0.8 almost perfect agreement. All analyses were undertaken in the R programming environment (R Version 2.15.1, R Development Core Team 2012), with statistical significance set at $p < 0.05$.

III Results

The cited article clearly supported the citation in the primary article in 80.8% of the cases. No support was provided for the assertion in 10.8% of the cases. In the majority

of these instances, the information could simply not be found in the cited article (e.g. the majority of the species mentioned as supposedly characterizing a certain region's vegetation could not be found in the citation provided). In one case, however, the cited article had found the direct opposite to what the primary article claimed it had found (lower assemblage abundance as opposed to higher assemblage abundance in urban areas). Ambiguous and empty citations both constituted 4.2% of the cases. Ambiguous citations were mostly cases where the information had been interpreted one way, but could also have been interpreted in another way. For example, a citation on the effect of urbanization on generalist beetles failed to mention that the cited article discusses different categories of generalists, not all showing the supposed trend. We also found cases where it was not clear what exactly the assertion was that was meant to be supported, presumably as a result of a misplaced citation. For example, when starting a sentence with "several oceanic fronts", immediately followed by a citation, it is not clear what the reference is supposed to support. Empty citations were all cases where an article had been cited that mentions a finding or even discusses the finding, but the finding was not one of the results of the article itself. For example, a review on temperature inferences from tree rings should not be used to demonstrate a reduction in sensitivity of tree growth to changing temperature, and such a citation was therefore classified as empty.

No associations were found between citation accuracy and number of authors (df = 4, X^2 = 14.26, p = 0.28), number of references in the primary article (df = 5, X^2 = 22.78, p = 0.09), the impact factor of the secondary article (df = 3, X^2 = 8.06, p = 0.53), the h-index of the first author (df = 3, X^2 = 11.09, p = 0.27) or whether or not the first author was a native English-speaker (df = 1, X^2 = 2.08, p = 0.56). However, a weak, but significant correlation was found between the number of clear supports and the impact factor of the journal that the primary article was published in (n = 20, r = 0.46, p = 0.04, Figure 1). Twenty-two of the assessed citations were self-citations. Of these, 19 were

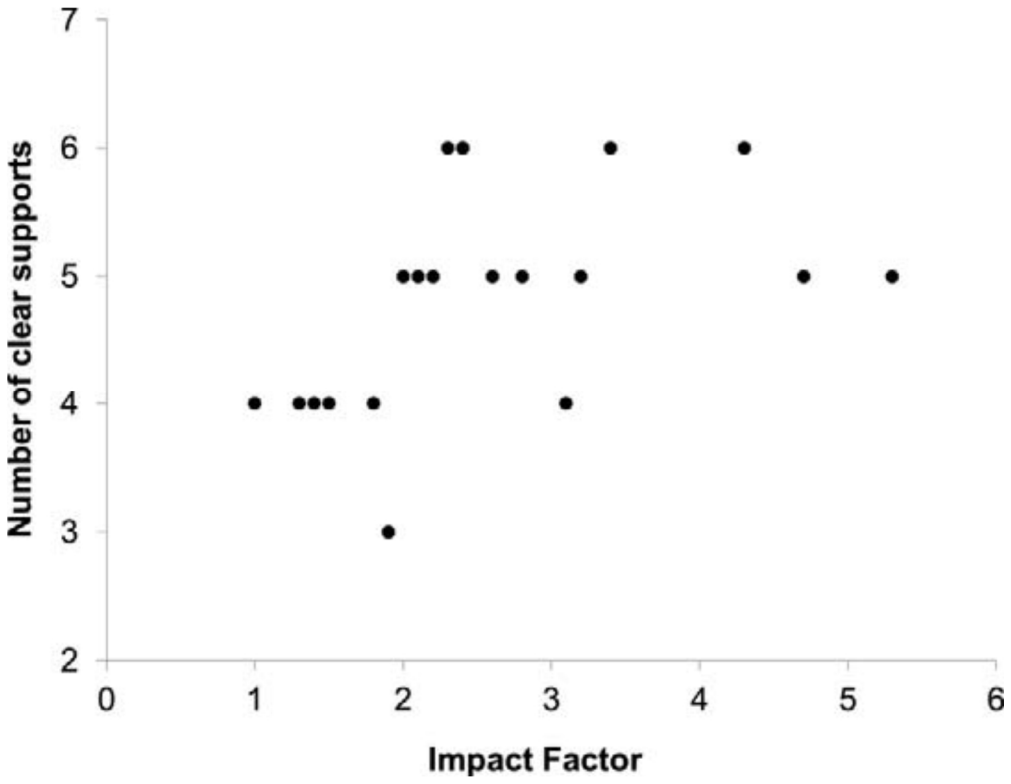

Figure 1. Relationship between journal impact factor and the number of clear support scores that were awarded to citations from the journal (n ¼ 20)

classified as having clear support (Category 1), and 3 as having no support (Category 2). Whether or not a citation was a self-citation did not have a significant relationship to the citation accuracy (df = 1, X^2 = 2.54, p = 0.47). There was fair agreement between raters and we obtained an Exact Fleiss' Kappa statistic of 0.27 (n = 120, Raters = 4).

IV Discussion

The percentage of clearly supported citations for articles in physical geography journals is similar to the values reported for anatomy publications, and slightly higher than what has been shown for ecology and marine biology publications (Table 2). As for the results reported by Todd et al. (2010), our results are likely to be overestimates of clear support for citations for two reasons. Firstly, we only considered single citations. We suspect that authors are more careful in selecting an appropriate citation when it is the only citation backing up a statement, than when it forms part of a whole string of citations. We therefore expect that the number of inaccurate citations would increase if we were to include assertions supported by multiple citations in our study. Secondly, equivocal results due to disagreement between raters were scored conservatively, giving the benefit of the doubt to the primary authors. The amount of disagreement was evidenced by the comparatively low Fleiss' Kappa value of 0.27, indicating only fair

Table 2. Percentage of articles providing clear citation support to assertions in articles arranged according to subject area (adapted from Todd et al. 2007). n = number of articles studied.

Subject area	**Clear support (%)**	**n**	**Source**
Nursing	93.3	180	Schulmeister (1988)
Radiology	90.5	95	Hansen & McIntire (1994)
Manual therapy	87.7	320	Gosling et al. (2004)
Burns and burn care	86.3	117	Al-Benna et al. (2009)
Otolaryngology/ head and neck surgery	83	153	Fenton et al. (2000)
Anatomy	80.9	272	Lukić et al. (2004)
Physical geography	**80.8**	**120**	**The present study**
Ecology	76.1	306	Todd et al. (2007)
Marine biology	75.8	198	Todd et al. (2010)
Ophthalmology	75	200	Buchan et al. (2005)
Surgery	70.8	137	Evans et al. (1990)
Emergency medicine	64.8	145	Goldberg et al. (1993)

agreement. This meant that a substantial number of citations were scored conservatively in favour of the primary authors.

Todd et al. (2010) concluded that the similarity in citation accuracy between ecology and marine biology may be related to the overlap in general discipline type, with both fields forming part of "field-orientated or environment-related biology". Given that physical geographers only fare marginally betters in terms of citation practices, we suggest that the prevalence of miscitations is unlikely to be related to biology specifically, although it may be indicative of general citation practices amongst "field-orientated" scientists.

Similar to Todd et al. (2010), we assessed relationships between citation accuracy and a number of likely influential factors. These included the number of authors, number of references, the h-index of the first-author, whether the first authors were native English speakers or not, and whether a citation was a self-citation or not. Citation practices were considered likely to improve with increased numbers of authors, given that multiple authors were more likely to identify miscitations in draft manuscripts. An opposite trend was expected with regards to numbers of references in the primary article, with miscitations expected to increase with increased numbers of references, due to increased opportunity for honest mistakes. However, similar to Todd

et al. (2010), our results indicated that neither number of authors nor number of references showed significant relationships with citation accuracy. The lack of relationships here is perhaps indicative of a tendency amongst secondary authors not to be very involved in the actual writing of manuscripts, but rather to be involved in other aspects of the research such as experimental design and data collection. Furthermore, the h-index of the first author also did not show any clear relationship with citation accuracy, indicating that well-cited authors do not employ better citation practices than authors with fewer citations.

The first language of the lead author was expected to potentially influence citation practices due to an increased likelihood of misinterpretations associated with language difficulties. However, our results indicated that non-native English speakers evidently did not show any differences in citation practice when compared to native English speakers. We assessed whether first authors were likely to be native English speakers or not on the basis of their surnames and affiliations. This subjective evaluation process inevitably introduced some bias into the results reported here and may have influenced the conclusions reached. Therefore, while we found no evidence that first language had any influence on citation practices, we are cautious in our interpretation of this result.

Self-citations were expected to be more accurate than other citations, since authors are unlikely to misinterpret findings from publications that they themselves were involved in writing. Our results suggested that there was no significant relationship between whether a citation was a self-citation or not, and the accuracy of the citation. However, three out of the 22 self-citations were rated as having no support. This was surprising, as one would expect authors to be familiar with their own work and one therefore would not expect any of the self-citations to be inaccurate. This result is of some concern, since it may indicate willful misconduct in the citation practices of some physical geographers. One possible reason for authors misciting themselves is pressure to accumulate citations of manuscripts. The pressure to publish scientific results and to obtain citations of such work is well known (Lawrence, 2003) and competition for limited numbers of jobs and funding has resulted in scientists increasingly being assessed by their productivity and impact (Nicolini and Nozza, 2008). The common use of citation metrics (e.g. h-index) may therefore lead to the temptation amongst scientists to cite their own work, irrespective of the relevance.

The only statistically significant, but weak relationship we found in our analyses was between number of citations with clear support and the impact factor of the primary journal. This relationship was positive, indicating that miscitations were more likely in articles published in journals with lower impact factors. A potential explanation is that

higher impact factor journals apply more stringent reviewing and editing practices. The relationship between journal impact factors and article citations is weakening and may lead to a reduction in the use of impact factor as a bibliographic tool to assess research output quality (Lozano et al., 2012). We did not consider adding a temporal component to our analyses, and so it would be interesting to assess if the strength of the relationship between citation accuracy and journal impact factor reported here has shown any signs of change over time.

Scientific misconduct is difficult to quantify, but can have major impacts on the image and the credibility of the scientific community at large. Unfortunately, scientific misconduct appears to be quite a common phenomenon, with examples of falsified scientific results, plagiarism and other forms of fraud abounding in all fields of science (Lacetera and Zirulia, 2009). Like other forms of scientific misconduct, the accuracy of citation practices in science affects the amount of trust one can place in the article being read, as well as the conclusions one can draw from it. We believe that the quality of citation practices is a key part of the overall integrity and quality of the science undertaken and reported, and should be monitored accordingly. Todd et al. (2010) suggested a number of remediation steps to improve citation practices amongst scientists, but emphasized that the main responsibilities of good citation practices lie with the authors themselves, and not the journals publishing their work. We agree with

this sentiment and encourage authors to be particularly careful to employ accurate citation practices. However, more directed measures at combatting poor citation practices by scientific journals may help prevent willful miscitation practices. Such suggestions have been summarized by Todd et al. (2010) and include publishing citation errors identified by readers (de Lacey et al., 1985), as well as instigating a process whereby authors are required to submit appropriate sections of text from cited material as requested by the editor, similar to the approach reported by Asano et al. (1995). We believe such practices should form part of a standard peer-review process. While such suggested measures will necessarily require additional time and input from editors, it could improve journal reputations if advertised appropriately.

V Conclusions

The results presented here indicate that at least 19% of the citations in the physical geography journals examined do not provide adequate support for the statement made. The commonness of miscitations in physical geography publications is concerning, as it could damage the reputation of the overall quality and trustworthiness of the work produced. We believe that while the responsibility for accurate citation practices lie primarily with the scientific authors themselves, some additional measures may be implemented by journal editors to encourage accurate citation practices.

Funding

This research was supported by a University of Pretoria Research Development Programme (RDP) grant [grant number AOT380].

References

Al-Benna S, Rajgarhia P, Ahmed S and Sheikh Z (2009) Accuracy of references in burns journals. *Burns* 35: 677-680.

Asano M, Mikawa K, Nishina K, Maekawa N and Obara H (1995) Improvement of the accuracy of references in the *Canadian Journal of Anaesthesia*. *Canadian Journal of Anaesthesia* 42: 370-372.

Bar-Ilan J (2008) Which h-index? - A comparison of WoS, Scopus and Google Scholar. *Scientometrics* 74: 257-271.

Buchan JC, Norris J and Kuper H (2005) Accuracy of referencing in the ophthalmic literature. *American Journal of Ophthalmology* 140: 1146-1148.

Conger AJ (1980) Integration and generalization of Kappas for multiple raters. *Psychological Bulletin* 88: 322-328.

de Lacey G, Record C and Wade J (1985) How accurate are citations and references in medical journals? *British Medical Journal (Clinical Research Edition)* 291: 884-886.

Evans JT, Nadjari HI and Burchell SA (1990) Quotational and reference accuracy in surgical journals. *JAMA* 263: 1353-1354.

Fenton JE, Brazier H, De Souza A, Hughes JP and McShane DP (2000) The accuracy of citation and quotation in otolaryngology/head and neck surgery journals. *Clinical Otolaryngology* 25: 40-44.

Fleiss JL (1971) Measuring nominal scale agreement among many raters. *Psychological Bulletin* 76: 378-382

Gavras H (2002) Inappropriate attribution: The 'lazy author' syndrome. *American Journal of Hypertension* 15: 831.

Goldberg R, Newton E, Cameron J, Jacobson R, Chan L, Bukata WR and Rakab A (1993) Reference accuracy in the emergency medicine literature. Annals of Emergency Medicine 22: 1450–1454.

Gosling CMR, Cameron M and Gibbons PF (2004) Referencing and quotation accuracy in four manual therapy journals. *Manual Therapy* 9: 36-40.

Hansen ME, McIntire DD (1994) Reference citations in radiology: accuracy and appropriateness of use in two major journals. *American Journal of Roentgenology* 163: 719-723.

Lacetera N and Zirulia L (2009) The economics of scientific misconduct. *The Journal of Law, Economics, and Organization* 27: 568-603.

Lawrence PA (2003) The politics of publication. *Nature* 422: 259-261.

Lozano GA, Larivière V and Gingras Y (2012) The weakening relationship between the impact factor and the papers' citations in the digital age. *Journal of the American Society for Information Science and Technology* 63: 2140-2145.

Lukić IK, Lukić A, Glunčić V, Katavić V, Vučenik V and Marušić A (2004) Citation and quotation accuracy in three anatomy journals. *Clinical Anatomy* 17:534–539.

Nicolini C and Nozza F (2008) Objective assessment of scientific performances world-wide. *Scientometrics* 76: 527-541.

R Development Core Team (2012) R: A language and environment for statistical computing. Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org/.

Schuermans N, Meeus B and De Maesschalk F (2010) Is there a world beyond the Web of Science? Publication practices outside the heartland of academic geography. *Area* 42: 417-424.

Schulmeister L (1988) Quotation and reference accuracy of three nursing journals. *Journal of Nursing Scholarship* 30: 143-146.

Sin SJ (2011) International co-authorship and citation impact: A bibliometric study of six LIS journals, 1980-2008. *Journal of the American Society for Information Science and Technology* 62: 1770-1783.

Slyder JB, Stein BR, Sams BS, Walker DM, Beale BJ, Feldhaus JJ and Copenheaver CA (2011) Citation pattern and lifespan: A comparison of discipline, institution and individual. *Scientometrics* 89: 955-966.

Todd PA, Guest JR, Lu J and Chou LM (2010) One in four citations in marine biology papers is inappropriate. *Marine Ecology Progress Series* 408: 299-303

Todd PA, Yeo DCJ, Li D and Ladle RJ (2007) Citing practices in ecology: Can we trust our own words? *Oikos* 116: 1599-1601.

Viera AJ and Garrett JM (2005) Understanding interobserver agreement: the Kappa statistic. *Family Medicine* 37: 360-363.