CSIRO PUBLISHING

Pacific Conservation Biology, 2019, **25**, 219–221
https://doi.org/10.1071/PCv25n3_ED

Editorial

Preventing information distortion in the conservation literature

Mike Calver

Environmental and Conservation Sciences, Murdoch University, Murdoch, WA 6150, Australia.
Email: m.calver@murdoch.edu.au

At the end of each teaching semester at my university (and no doubt many others) comes the ritual of student evaluations of teaching and units. These evaluations provide not only summary numbers of student opinions expressed on Likert scales but also qualitative written comments from students on the learning and teaching experience. While there are plenty of critics of the value of these evaluations in assessing the quality of teaching, many institutions do take them seriously and ask academics to include them in tenure and promotion applications. Academics vary greatly in how they use the evaluations. Some cherry pick them for good results, quoting favourable comments and statistics to showcase the best in their teaching as perceived by students. Others are more reflective, using negative comments to develop plans for improvement. Still others blend the results from the student evaluations with other commentary such as peer feedback, employer comments or reflections from graduates on how their experience did (or didn't) prepare them for the workforce.

A similar diversity of approaches can be used when assembling evidence for academic writing. Authors can assemble data or literature selectively to make a case while discarding alternative opinions or evidence, or acknowledge alternative views or multiple strands of evidence before reaching conclusions. In any controversy, authors tend to argue that their opponents take the first approach, but may conveniently forget that they are following it themselves. I am no more immune to this than anyone else, but I will take the risk of being called a hypocrite by opening the issue of distortion in data representation and in the use of the scientific literature.

What is distortion?

Bella (1996) introduced the term 'systemic distortion' to describe the processes by which organisations or research groups select and promulgate information that supports the agendas of the organisation or research group, while marginalising, ignoring or suppressing unfavourable information. Here, I broaden the scope to include authors preparing manuscripts and deciding on data to include, papers to cite, and opinions to consider. Of course, 100% coverage of relevant data and papers is rarely possible in one manuscript, so selectivity alone is not distortion – the defining feature of distortion is selectivity based on support for a key argument, with unfavourable information discarded.

Types of distortion in scientific reporting

Distortion of evidence may arise in the selection of sources cited in a paper. Citations purportedly show the influences that underpin a scientific paper and give credit to those influences, but MacRoberts and MacRoberts (1996, 2010, 2018) argue that they do not because of: failure to cite all influences/sources, systematic bias in selection of citations (including deliberately selecting or ignoring sources depending on their support for the author's view or prioritising self-citation), a preference for secondary sources (e.g. reviews, textbooks over primary sources – for those who doubt that this happens, ask yourself when you last cited a primary statistics paper over a 'how to' chapter in a textbook), failure to cite or acknowledge informal communications, traditional non-citing of certain communications (e.g. floras), and selectivity arising from databases used or a focus on a subset of mutually supportive literature. While the MacRoberts were mainly concerned with problems that arise when citation counts are used to evaluate the merit of work, other authors across a range of disciplines have turned to problems in the conduct and interpretation of research.

For example, authors may marginalise or ignore particular views when choosing papers to cite. This may be deliberate, or may arise from incomplete literature searches or relying on citation networks of supportive papers (Leng 2018). Authors may also refer to data in other papers but interpret it differently to the original authors – this isn't a problem if the different interpretation is acknowledged, but can be misleading if the implication is that the reinterpretation is the view of the original authors (Greenberg 2009). Other problems include failure to separate hypotheses/opinions (weak evidence) from empirical data (strong evidence) (Greenberg 2009), or a preference to cite studies supporting a view or interpretation and ignore others against it (Fletcher and Black 2007; Misemer *et al.* 2016; Hanson *et al.* 2018; Oldehinkel 2018). The meaning of data may be distorted as well. For example, one medical study evaluated analyses of data from 616 random controlled trials in which there were statistically insignificant results for the primary outcomes, finding that ~40% of the studies they evaluated used '…specific reporting strategies, from whatever motive, to highlight that the experimental treatment is beneficial, despite a statistically nonsignificant difference for the primary outcome, or to distract the reader from statistically nonsignificant results' (Boutron *et al.* 2010, p. 2059). All of

Journal compilation © CSIRO 2019

these issues may contribute to 'network authority' (Greenberg 2009), in which within networks of mutually citing papers critical/negative views are ignored and the importance of supportive/positive studies is amplified by repeated citation (Greenberg 2009).

Reasons for and consequences of information distortion

Boutron *et al.* (2010, p. 2058) claimed that distortion can arise from '…ignorance of the scientific issue, unconscious bias, or wilful intent to deceive.' Greenberg (2009) has argued that authors may be advantaged in publishing, citation and grant success by accepting rather than critiquing a prevailing view, because positive studies are more likely to be accepted.

Whatever the reasons for information distortion, it comes at a cost to a discipline. For example, Hanson *et al.* (2018, p. 1044) concluded:

> *…this case study with atrazine demonstrates that there is bias in the manner in which authors are citing studies, and journals are publishing studies, regardless of quality, in favour of those that report 'significant' adverse effects within ecotoxicology. This lack of engagement by authors and journals with the complete body of knowledge undermines ecotoxicology as a discipline and imperils environmental protection.*

These are strong reasons to tackle information distortion.

Guarding against information distortion

There are habits that we can all develop that may help us in recognising and removing information distortion. They include:

Using multiple databases for literature searches – most of us have our favourite databases, which we may use uncritically to search the literature. If a favourite database has a bias towards particular journals, is geographically or temporally limited in its range, or is likely to exclude significant grey literature such as government reports, important information will be missed (Calver *et al.* 2017). Such biases may be overcome by searching in multiple databases.

Be aware of the grey literature and unpublished data – grey literature (government reports, theses and the like) can include vital empirical data, yet it is often covered poorly in the major databases. In some cases, grey literature can be located using special features of a database (e.g. the 'secondary documents' feature in Scopus retrieves sources not in Scopus that have been cited by documents that are in Scopus). Unpublished data may also be relevant, especially in the context of augmenting work already published (Côté *et al.* 2011). Harding (1998) notes that indigenous knowledge may be an important subset of unpublished data. Google Scholar can be good for searching for grey literature (Harzing and van der Wal 2008). RAC (1993) and Côté *et al.* (2011) give good frameworks for locating relevant grey literature on a topic. Network contacts and personal approaches can be used to obtain unpublished information, which can only be used with permission (Côté *et al.* 2011).

Consider systematic reviews – systematic reviews describe clearly the methods used to search the primary literature, evaluate the papers retrieved for relevance to the topic, and choose papers for detailed consideration. Therefore others can replicate the approach. Often, bibliographies of all literature retrieved are presented as well as the criteria for choosing papers for more detailed examination, with a list of the subset selected. Decisions for inclusion might be based on experimental design, empirical papers versus reviews, sample sizes and so on. The critical point is that the decision path is clear, so there can be no accusations of hidden selectivity (Côté *et al.* 2011; Curtis *et al.* 2011).

Consider whether or not to base a decision on a P-value alone – there is a lively literature on whether or not the null-hypothesis significance testing model (NHST), which has guided research for nearly a century, should be replaced as an evidential standard (e.g. Ho *et al.* 2019). Perhaps rather than torture data using the NHST if authors feel that this is not providing the necessary nuance, other options could be considered (e.g. Harlow *et al.* 1997; Burnham and Anderson 2002).

Acknowledge data limitations – Underwood (1997, p. 484) expressed this clearly:

> *Above all, be self-critical. If some experimental test of an hypothesis is problematic because of difficulties with controls, independence, replication and other aspects of design, the person most responsible for explaining this is the person who did the experiment. Then an argument can be advanced to explain why the results and interpretation should be accepted, despite the problem. This argument will rest on ancillary evidence, inductive notions based on experience, analogy, etc. Some of it may be compelling. All of it needs to be aired. Otherwise, readers, referees, editors, etc., elsewhere in the world all working on different problems are entitled (and have a duty) to reject the findings.*

Presenting all the relevant detail behind a decision allows others to evaluate the decision from a position of full information.

Acknowledge alternative views – in many cases, following the steps above will lead to authors uncovering alternative opinions to those they wish to present. I suggest that in these cases authors acknowledge the alternative views before explaining why they disagree and advancing their own position. Simply ignoring the difference of opinion/interpretation may lead to a tightly cohesive paper, but it misrepresents the range of views.

To me, striving to present conclusions in the context of all relevant data, interpretations and opinions strengthens, not weakens, conclusions. It's worth striving for such a goal.

References

Bella, D. A. (1996). The pressures of organizations and the responsibilities of university professors. *Bioscience* **46**, 772–778. doi:10.2307/1312854

Boutron, I., Dutton, S., Ravaud, P., and Altman, D. G. (2010). Reporting and interpretation of randomized controlled trials with statistically nonsignificant results for primary outcomes. *Journal of the American Medical Association* **303**, 2058–2064. doi:10.1001/JAMA.2010.651

Burnham, K. P., and Anderson, D. R. (2002). 'Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach.' 2nd edn. (Springer-Verlag: New York.)

Calver, M. C., Goldman, B., Hutchings, P. A., and Kingsford, R. T. (2017). Why discrepancies in searching the conservation biology literature matter. *Biological Conservation* **213**, 19–26. doi:10.1016/J.BIOCON.2017.06.028

Côté, I. M., Curtis, P. S., and Rothstein, H. R. (2011). Gathering data: searching literature and selection criteria. In 'Handbook of meta-analysis in ecology and evolution.' (Eds J. Koricheva, J. Gurevitch, and K. Mengersen.) Chapter 4, pp. 37–51. (Princeton University Press: Princeton and Oxford.)

Curtis, P. S., Mengersen, K., Lajeunesse, M. J., Rothstein, H. R., and Stewart, G. B. (2011). Extraction and critical appraisal of data. In 'Handbook of meta-analysis in ecology and evolution.' (Eds J. Koricheva, J. Gurevitch, and K. Mengersen.) Chapter 5, pp. 52–60. (Princeton University Press: Princeton and Oxford.)

Fletcher, R. H., and Black, B. (2007). "Spin" in scientific writing: Scientific mischief and legal jeopardy. *Medicine and Law* **26**, 511–525.

Greenberg, S. A. (2009). How citation distortions create unfounded authority: Analysis of a citation network. *BMJ* **339**, b2680. doi:10.1136/BMJ.B2680

Hanson, M. L., Deeth, L. E., and Prosser, R. S. (2018). Evidence of citation bias in the pesticide ecotoxicology literature. *Ecotoxicology (London, England)* **27**, 1039–1045. doi:10.1007/S10646-018-1918-4

Harding, R. (Ed) (1998). 'Environmental Decision-making: the Roles of Scientists, Engineers and the Public.' (The Federation Press: Sydney.)

Harlow, L. L., Mulaik, S. A., and Steiger, J. H. (Eds) (1997). 'What If There Were No Significance Tests?' (Lawrence Erlbaum Associates: Mahwah, New Jersey.)

Harzing, A. W. K., and van der Wal, R. (2008). Google Scholar as a new source for citation analysis. *Ethics in Science and Environmental Politics* **8**, 61–73. doi:10.3354/ESEP00076

Ho, J., Tumkaya, T., Aryal, S., Choi, H., and Claridge-Chang, A. (2019). Moving beyond P values: data analysis with estimation graphics. *Nature Methods* **16**, 565–566. doi:10.1038/S41592-019-0470-3

Leng, R. I. (2018). A network analysis of the propagation of evidence regarding the effectiveness of fat-controlled diets in the secondary prevention of coronary heart disease (CHD): Selective citation in reviews. *PLoS One* **13**, e0197716. doi:10.1371/JOURNAL.PONE.0197716

MacRoberts, M. H., and MacRoberts, B. R. (1996). Problems of citation analysis. *Scientometrics* **36**, 435–444. doi:10.1007/BF02129604

MacRoberts, M. H., and MacRoberts, B. R. (2010). Problems of citation analysis: A study of uncited and seldom-cited influences. *Journal of the American Society for Information Science and Technology* **61**, 1–12. doi:10.1002/ASI.21228

MacRoberts, M. H., and MacRoberts, B. R. (2018). The mismeasure of science: Citation analysis. *Journal of the Association for Information Science and Technology* **69**, 474–482. doi:10.1002/ASI.23970

Misemer, B. S., Platts-Mills, T. F., and Jones, C. W. (2016). Citation bias favoring positive clinical trials of thrombolytics for acute ischemic stroke: A cross-sectional analysis. *Trials* **17**, 473. doi:10.1186/S13063-016-1595-7

Oldehinkel, A. J. T. (2018). Editorial: Sweet nothings – the value of negative findings for scientific progress. *Journal of Child Psychology and Psychiatry, and Allied Disciplines* **59**, 829–830. doi:10.1111/JCPP.12952

RAC (1993). Ecological impacts of forest use: a survey of completed research. Resource Assessment Commission, Research paper no. 9, Canberra.

Underwood, A. J. (1997). 'Experiments in Ecology: Their Logical Design and Interpretation Using Analysis of Variance.' (Cambridge University Press: Cambridge.)