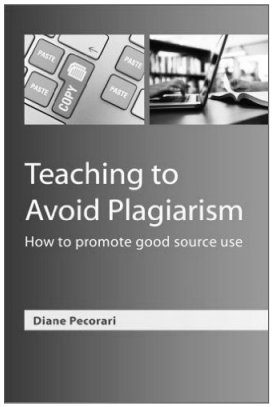
Book review

Teaching to avoid plagiarism: how to promote good source use

Diane Pecorari, McGraw Hill, 2013, ISBN 978-0-33-524593-2



Reviewing this book for *European Science Editing* feels rather unfair since it makes no claims to be aimed at journal editors. However, since editors should be concerned about plagiarism, and because I (and perhaps other EASE members) offer training for authors on this topic, I was interested to read it. The introduction states that this book is aimed at university staff to help them understand plagiarism and develop their skills in handling it. One strength of the book is that the author, Diane Pecorari, mentions her own research into plagiarism and also her own experiences of working with students. Her main argument is that students should be taught about citation and the use of sources and she attributes much student plagiarism to poor understanding of academic norms. Given this premise (ie that much student plagiarism may be inadvertent), teachers need to decide when plagiarism is, instead, deliberate, aimed to deceive, and therefore misconduct.

The use of text-matching software (such as Turnitin, which uses the same software as CrossCheck used by many journals) is discussed in detail. Pecorari urges caution and emphasises the limitations of such systems, while rather grudgingly accepting that they may be useful.

One aspect editors might empathize with, and which I thought was well covered, was the isolation and lack of support that teachers (and, I dare say, editors) may feel when faced with student (or author) plagiarism. Pecorari describes the effects of confronting plagiarism and the reasons why teachers may be tempted to overlook cases. She suggests that institutions could provide confidential advisors to help teachers handle specific cases – this reminded me of the origins of COPE (the Committee on Publication Ethics) which started, and still functions, as a 'self-help group for editors' where troublesome cases can be discussed in confidence.

One thing I felt was lacking was a distinction between the needs and behaviours of undergraduate and postgraduate students (or any discussion about whether this distinction was relevant). Editors may vary their response, especially to more minor or borderline cases of plagiarism, depending on the seniority of the author and whether he or she 'should have known better'. I was therefore surprised that this book contained no discussion about the difference between copying in an undergraduate essay (especially if it does not count directly towards a degree) and plagiarism in dissertations.

The book contains a short chapter on "Differences across academic subjects" but I would have been interested to have more examples of such differences. Pecorari notes "At the risk of generalising, the STEM field deals in objective fact, and it is the facts themselves which matter; the form in which they are expressed is much less important". While most editors only have to work within a single discipline, and therefore understand its particular conventions, this doesn't help much for people in cross-disciplinary areas and seems to downplay the problem of plagiarism in the hard sciences.

While each chapter ends with helpful suggested activities and "Questions for discussion or reflection" I felt the overall ending of the book was rather weak. Rather than attempting to synthesize the arguments, the final chapter covers plagiarism in non-academic writing and I found it disappointing.

My overall conclusion is that, while this book may be helpful for university teachers, especially those who lack confidence in approaching cases of plagiarism, it doesn't contain much of interest to journal editors. Finally, firmly wearing my editor's and writing trainer's hat, I couldn't help but notice some horribly long sentences (containing 50-65 words). I also spotted some wordy phrases and use of jargon which would have benefited from careful editing.

Elizabeth Wager
Publications Consultant, Sideview, Princes Risborough, UK
liz@sideview.demon.co.uk

EASE-Forum Digest: March to June 2015

You can join the forum by sending the one-line message "subscribe ease-forum" (without the quotation marks) to majordomo@helsinki.fi. Send in plain text, not HTML. Details at www.ease.org.uk/node/589.

Do UK researchers have or need in-house writing support?

Several US universities have in-house publications departments that help researchers with manuscript writing. Some non-Anglophone European university departments provide similar support through contracted/freelance editors. Valerie Matarese had read a survey of UK researchers about technical support which did not appear to include editorial support (http://www.kcl.ac.uk/sspp/departments/management/people/academic/lewisgatsbyreport.pdf). She was curious to know if UK universities offer any writing assistance to their researchers.

In Alan Hopkin's experience, any help given in the UK was informal, eg a senior colleague reading the paper. Indeed, many scientists did not want assistance, especially if editing delayed submissions. By contrast, statistical support was welcomed because it was thought to avert reviewers' comments. Andrew Davis had found both writing and statistical support were only common in commercial institutions in the UK. Researchers in academia resisted writing assistance because as native English speakers they did not recognise that their papers were poorly written. His suggested solutions were either to provide manuscript writing training courses, or support as in the commercial world. He preferred the first solution, as did Marge Berer, who found that most papers her journal received needed editing.

Kersti Wagstaff lamented the time when writing was 'taught' by a process in which dissertations were passed back and forth between student and supervisor until the required standard had been attained. Supervisors no longer have time, so theses are edited commercially, with the ethical implications hotly debated on copy editors' listservers. Ultimately, it came back to money, "The universities don't have anything extra to spend on teaching writing, and the pressure to publish is so great that the journals have got the upper hand on their authors and can force the cost of achieving a readable text back down on them." Chris Sterken thought lack of time was dire when scientists confined their reading to online sources and never went to the library to read well-written and structured 'old' documents that explained valuable basic science. More money spent on outsourcing meant less was available for hiring in-house staff, resulting in an overworked faculty.

Are deadlines for manuscript revisions useful?

"Why should it be a problem if an author takes 6 months to revise a manuscript following review?" asked Aleksandra Golebiowska. Because, Pippa Smart listed, the research may become out of date, the journal may in the meantime publish another article on the same topic and not want to repeat it, or the editorial policies may change. Stuart Handysides added, to get the information into the public domain as soon as possible, to tie in with related papers or other articles in a forthcoming issue. Helle Goldman found some authors were aware that their research may become irrelevant so did not require a deadline. Prompt publication was, however, often requested because the authors were due for promotion, not because the research was groundbreaking. Eric Lichtfouse had experienced authors who had taken advantage of a long deadline time to make the corrections and submit to a higher impact journal.

An intriguing argument for keeping deadlines short came from some of Helle's editors—reviewers will have forgotten the manuscript and be less inclined to agree to check the revisions. Aleksandra thought reviewers often approved a new version without looking at the changes, but more because they did not want to spend time on middling papers than because of lapse of time.

Andrew suggested deadlines were given to allow negotiation with the printers on issue size, when editors needed to be able to estimate the time it would take for accepted manuscripts to come back from revision. Aleksandra questioned if journals did plan issues, as not all revised manuscripts would be immediately accepted for publication. Angela Turner agreed, which paper goes into which issue is of little importance with online first publication because papers are published online as soon as they are ready. An issue is produced on a certain date in the month from whichever papers are published online. Angela allowed a month for revisions but authors could always ask for more time.

Fake email addresses: combating peer review scams and false signatures on authorship forms

Incidents of authors using non-institutional email addresses to trick journals into appointing them to review their own papers have been discussed on the forum before (see ESE vol 41(1):21-22). However, such addresses can also be used for other reasons. Researchers from Asia and Africa have reported on the WAME listserver using them because institutions did not provide addresses or the addresses did not work well. Udo Schuklenk, an American, said on that list that he uses Gmail because data would be more easily transferrable should he change institutions. PhD students might not have an institutional address between jobs. EASE forum participants agreed it would be unfair to penalise scientists by insisting on an institutional email address. Although caution was needed, use of non-institutional accounts, eg Yahoo, Gmail, could not be the sole reason for excluding a reviewer (see Marcin Kozak and colleague's article "Do researchers provide public or institutional E-mail accounts as correspondence E-mails in scientific articles?" http://onlinelibrary.wiley.com/doi/10.1002/asi.23401/abstract).

Liz Wager asked how journals should deal with false author signatures. She gave an example of a journal

that had received online signature forms confirming authorship from each author of a Chinese paper within a few minutes. The email addresses were from free email services, which is common for authors from China. Michael Altus on the WAME listserver (Liz cross-posted) believed the publication process should be halted while the corresponding author was asked for an explanation. He thought that in general corresponding authors forwarding forms from each of the co-authors was acceptable. On our forum, Angela reported that Elsevier's new online submission system, Evise, made it mandatory to include an email address for every co-author, so that they can receive a copy of emails sent to the corresponding author. But, she added, that doesn't prevent people making up an address. Likewise, falsification of signatures on author forms cannot be prevented. An author's complaint that her signature had been falsified on a form I received by post while managing editor of *Diabetologia* in the 1990s resulted in the police seizing the file.

Forum participants were unanimous in their view that editors should check the identity of reviewers and authors, and agreed this was more difficult with non-institutional email addresses. Helle's journal checks potential reviewers using the Web of Science and institutional websites and ensures they do not share an institutional affiliation with the authors and haven't co-authored papers with them. Andrew suggested the journal Liz mentioned should get a Chinese person to help check whether the authors named had ever been at the institution where they claimed to have done the work. He had worked in China where he found most researchers had institutional email addresses and believed reputable research institutes would not prevent former PhD students from retaining their institutional email addresses while working to publish work they had done there.

Use of reviewers suggested by authors

Use of reviewers suggested by authors was also discussed in relation to *Nature*'s peer-review scam article (vol. 515, pp. 480-482) mentioned in ESE's earlier issue previously cited. Opinions varied from Marcin who would never use referees suggested by authors to Marge Berer who uses them all the time along with reviewers she chooses herself. She sees little alternative when the subject is narrow and the suggested reviewers are virtually the only scientists with the requisite expertise. The difficulties editors face in finding willing peer reviewers are explained in Helle's article series (http://www.editage.com/insights/series/delays-in-peer-review). Tom Lang posted a link to WAME's position statement on author-suggested peer reviews. http://www.wame.org/about/policy-statements#Best%20Practices%20for%20Peer%20Reviewer%20Selection.

Ownership of biological material

Elisabeth Heseltine posed an interesting question from an author who had received a request to share primary cells from human tissues used in research the author had published in *PLoS One* a few years before. The author could not provide the same cells but was currently using cells established similarly from different human tissues for a future publication that would meet their requirements. She asked if the sharing of this unique material warranted authorship and whether she was obliged to provide either the same published material or the new material.

Tom Lang was certain that providing cells did not warrant authorship, only an acknowledgement (ICMJE recommendations www.icmje.org). He was not sure about the second question as there would be privacy issues, patent issues, cost issues, and pragmatic issues in sharing cell lines. Examples of where sharing was required were *Nature* http://www.nature.com/authors/policies/availability.html and the NIH's Principles and Guidelines for Reporting Preclinical Research http://www.nih.gov/about/reporting-preclinical-research.htm. However, as pointed out by Andrew, *PloS One* does not oblige authors to provide their original material to others. Nevertheless, Andrew thought they should do so unless there were good reasons not to, eg confidentiality or the original informed consent for taking the cells not allowing use by others. The same applied if only substitute cell lines were available. Most major journals publishing this kind of material recommended the deposition of cell lines in a suitable facility from which they could be accessed by others. This was also beneficial for authors because the repository maintains the cell lines and sends them to researchers who require them without cost to the original authors. Paola De Castro posted a link to COBRA (Citation of Bioresources in Research Article, http://www.biomedcentral.com/1741-7015/13/33) on the standard citation of bioresources in journals. A workshop is being organised by EASE and the BRIF (Bioresource Reasearch Impact Factor) for Toulose on October 9, titled "Editors as promoters of good practice in bioresource research", details about which can be found on the EASE website.

Reporting p values

Valerie had noticed that although many guidelines (including SAMPL http://www.equator-network.org/2013/02/11/sampl-guidelines-for-statistical-reporting/) recommend reporting exact p values, *Science* uses cut-offs, eg $p<0.05$ and $p<0.001$ and even "ns". She wanted to know why. She was referring to articles in biology with common tests like *t* and ANOVA and very small samples. Kersti was also used to seeing cut-offs rather than exact p values in clinical medical literature.

Andrew considered p values useless unless the test, exact model, degrees of freedom and sample sizes were reported for the reader to assess whether that test and that p value showed an interesting relation. However, he thought there were worse things than using cut-offs for p values, and gave the example "... ANOVA p=0.0165". He was more concerned, however, by Valerie's mention of very small samples because the power of any test was likely to be too low to be trustworthy.

Tom, who is more familiar with the symbol p in clinical medicine, said using exact p values (or "p values expressed as equalities" because an "exact" number has a specific meaning in mathematics) was recommended in clinical journals, although confidence intervals are increasingly being reported. He thought the reason for this was that "an estimated effect size (the result of the study) is clinically interpretable in clinical medicine, and using the p value for interpreting the result has, and continues to be, horribly abused because the effect size is often ignored when p is less than 0.05. The confidence interval is a measure of precision for the estimated effect size and so keeps the interpretation focused on the medicine. In basic science, the biological implications of an effect size are often unknown, so the only thing left is to say the difference was statistically significant or not (hence, "ns" is as good as $p<0.05$). It's a rough measure that something is happening or not, and that may be useful in basic science. The p value is a mathematical concept of chance as an explanation and has no biological interpretation; the alpha level (often the 0.05) is an arbitrary number, and p values of 0.051 and 0.049 should be interpreted pretty much the same way, so I always recommend reporting the exact p value, never ns, and in most cases (there are some statistical exceptions) ask for confidence intervals."

From the responses, Valerie noted that clinical (and related biomedical) fields prefer exact p values, usually with three significant digits and no smaller than 0.001, while p values down to almost infinitesimally small values (with at least three significant digits) were reported in genomics. In more traditional biological fields, one-significant-digit values were preferred either with cut-offs (eg ns, <0.05, <0.01, <0.001) as in *Science* or a bit more precise with p values given (eg $p<0.04$, $p<0.002$, $p<0.0008$) as in *Nature*. These experiments typically had 5 or 10 mice per group. The reported effects (mean and standard mean of error, SE) were large and could be 'seen' without p values. She posted a link to an article that criticises top journals like *Nature* and *Science* for poor statistical reporting: http://www.nature.com/neuro/journal/v14/n9/full/nn.2886.html. Andrew assured that *Nature* had since adopted a more rigorous policy towards statistics. Also, rather than ways of reporting, the article was critical of incorrect procedure. He gave an example: "if p for the comparison Control against Treatment A is less (say p=0.0157) than the p for the comparison Control against Treatment B (say p=0.0202), Treatment A is reported as being significantly different from Treatment B. This is bad experimental design. A better test would be the correct generalised linear model of the three classes Control, Treatment A and Treatment B followed by appropriate contrasts and post hoc tests between classes."

Tom joined that basic science has a history of reporting the SE, which defines a confidence interval around the estimate of about 68%. Debate continues as to whether reporting should be changed to the more correct SD. Marcin said the *Annals of Applied Biology* accepted SEs but not confidence intervals. SEs might be better because readers could choose the particular confidence level. Tom was concerned that SEs were often used to describe statistics rather than as estimates.

Farrokh Habibzadeh advised researchers to use SD for expressing the distribution of data in the sample (and as a good estimation of the distribution in the population) and confidence interval (which is calculated from SE) for showing how precisely they have measured an estimate (say mean age of patients) in the population. He thought SEs were not useful but were attractive for authors because they are always less than SDs, implying (incorrectly) that the dispersion of data was lower than when SDs are reported. Discussants agreed that there was no difference between using SE, SD or confidence interval provided the reader knew what each meant and how to interpret them. However, as most readers and indeed many authors are not proficient in statistics, the critical message to editors from this discussion is that their instructions to authors need to explain how authors should describe their samples and choose the most appropriate measures to represent the sample distribution, ie confidence interval.

Journals publishing articles simultaneously in two languages

Pippa asked for examples of journals that publish entire articles in two languages. Kersti said *Deutsches Ärzteblatt* does this and gave a link (http://www.aerzteblatt.de/int/archive/article?id=58636) mentioning other medical journals that publish bilingually. Marge Berer, founding editor of *Reproductive Health Matters*, said that journal published 7 different language editions. A small editorial team in each country manages the process and translates the articles most relevant to their regions from the main English edition. Each group has to include a list of common items in the prelim pages. The different versions are not technically new publications, which meant it had been more complicated than expected for the publisher, Elsevier, to put them on their website.

Other forum topics

The forum also discussed levels of editing. Space does not allow me to report on this topic but anyone interested is welcome to contact me for a copy of the original contributions.

Elise Langdon-Neuner (compiler)
a.a.neuner@gmail.com

Discussion initiators

Valerie Matarese: vmatarese@uptoit.org
Aleksandra Golebiowska: algol@ciop.pl
Liz Wager: liz@sideview.demon.co.uk
Elisabeth Heseltine: e.heseltine@gmail.com
Pippa Smart: pippa.smart@gmail.com