Copyediting's Role in an Open-Access World

by **Sanford G. Thatcher** (Director Emeritus, Penn State Press, 8201 Edgewater Drive, Frisco, TX 75034-5514; Phone: 214-705-1939) <sandy.thatcher@alumni.princeton.edu>

In an earlier article in *Against the Grain* titled "The Value Added by Copyediting" (September 2008), I raised the question of whether the move toward self-archiving of less-than-final versions of articles carried a risk of corrupting scholarship by tempting students and scholars to rely on imperfect versions of reported research because of their greater ease of access. Green OA has much to be said for it as a vehicle for more widespread and rapid dissemination of research, but it is not, I suggested, a perfect solution in every respect. What compromises would we be making by too readily accommodating ourselves to a new world of multiple variants of articles, I wondered?

My suspicions were based on experiences I had had early in my career in publishing when I worked on the staff of **Princeton University Press** as a copyeditor. I gave examples in the earlier article of some perhaps extreme cases of scholarly slovenliness, ranging from bad prose to incorrect citations to inaccurate quotations. During my later years as an acquiring editor and director, I continued to be exposed to the shortcomings of scholarly writing, but did not have direct responsibility for repairing them, relying on the seasoned copyeditors we had on staff or hired as free-lancers to take care of the problems. My commitment to copyediting as a valuable contribution of publishers remained strong, however.

But beyond those older anecdotes I had no evidence to offer of the problems as they exist today. Thus I decided to enlist the help of a few colleagues with copyediting expertise at other university presses who volunteered to assist me in conducting a preliminary assessment of the nature and scope of the risks that reliance on Green OA might entail for scholarship's integrity: **Jenny Hunt**, Assistant Production Manager at **Baylor University Press**; **Sylvia Hunter**, Editorial Services Supervisor in the Journals Division of the **University of Toronto Press**; and **Wayne Larsen**, Project Editor at **Southern Illinois University Press**.

For convenience, and because of its high profile, we focused on articles posted at **Harvard University's** DASH (Digital Access to Scholarship at Harvard) Website, which now has more than 5,000 articles in its repository. Many of the articles posted there have links to the final versions as published, and it was therefore relatively easy for us to compare the Green OA versions with the versions of record to see how much and what kind of copyediting was done, since many of these journals are accessible through the subscriptions that the libraries at our universities have to them in digital form. While the copyediting done for the published versions caught most of the errors in the Green OA versions, we discovered that some errors remained in the versions of record.

Collectively, we covered a range of journals in the humanities and social sciences: *American Economic Review* (2004), *The Journal of Consumer Affairs* (2010), *Philosophy and Public Affairs* (1989), *Political Theory* (2007), *Proceedings of the British Academy* (2007), and *Psychological Science* (2010). Two articles from edited volumes, published by **Edward Elgar** (2006) and **Russell Sage** (2008), also were scrutinized. The authors included both junior and senior faculty. Three of the articles were written by more than one author. We made no attempt to be "scientific" in this selection, since such a small number could not constitute any kind of statistically significant sample. Our findings must therefore be considered as illustrative only. If there is any bias in the sample, it was purely unintentional.

Two of my colleagues attempted to quantify the results of their inspections. **Wayne Larsen** looked at the two articles included in edited volumes, but did not compare those Green OA versions with the versions appearing in the books and did not attempt to check the accuracy of quotations or citations. For one he counted 15 errors in grammar and 49 in style; for the other, 3 in grammar and 85 in style. For the latter, he noted, "inconsistencies or errors in punctuation and spelling style constituted the greatest number. These included spacing problems, such as quotation marks not closed up to words; omission of commas between complete clauses in compound sentences; placement of commas where they are inappropriate." He also observed that "two reference entries mention only the author names (and, in one case, a publication year) and end with ellipses."

Sylvia Hunter provided a more detailed breakdown in her report. For the article on French history in the *Proceedings of the British Academy*, she enumerated 4 grammatical errors (2 of faulty parallelism, 1 subject/verb disagreement, and 1 other), 10 stylistic infelicities, 1 spelling error, 4 errors in quotation, 2 citation errors, and 6 errors in tables, figures, etc. Problems with this manuscript included the author's indecisiveness about whether to use U.S. or UK formatting conventions and the unhelpful lack of specific textual references to the figures accompanying the article. The tally for the article in the *American Economic Review* included 31 grammatical errors (5 of faulty parallelism, 3 subject/verb disagreement, 1 dangling modifier, and 22 others), 23 stylistic infelicities, 6 spelling errors, and 7 errors in citations. The authors of this article frequently used acronyms in the Green OA version without providing the full names of the organizations to which the acronyms pertained; this problem was corrected in the version of record. One work cited in the bibliography as published in 1996 did not match the citation in the text, which was inexplicably changed from 1994 in the Green OA version to 1997 in the version of record.

Jenny Hunt did not quantify her findings but presented examples in considerable detail. The article published in *Psychological Science*, for instance, "did not include [in its Green OA version] important information such as the identity of the corresponding author, acknowledgments, funding information, and declaration of conflicting interests." Also, "the figure captions changed very significantly between the Green OA and published versions. The published captions were much longer and more detailed in their information/explanations and included the definitions for the error bars." Many of the problems were minor formatting errors or inconsistencies, but in a number of instances the copyeditor had improved the style by reformulating sentences for greater clarity

against the grain
people profile

Sanford G. Thatcher
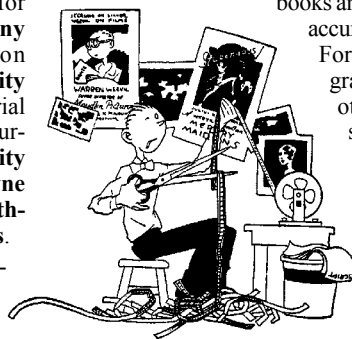
or fluency. The copyeditor also corrected some inconsistencies between citations in the text and entries in the bibliography, in names or dates, but also missed a couple of these, which remained as mistakes in the version of record. The results for the article published in *The Journal of Consumer Affairs* were very similar, with the majority of problems being minor errors of formatting that were corrected for the version of record. **Ms. Hunt** also discovered a couple of mistakes in quotations and references that were not caught by the copyeditor and remain in the version of record.

My contribution was to compare the Green OA and published versions of two articles in political philosophy, one by a senior professor of philosophy appearing in *Philosophy and Public Affairs* and the other by a junior professor of political science in *Political Theory*. As I subscribe to both journals, I was able to pull the issues off my shelf and read them in parallel with the versions posted at DASH. In the case of the former journal, which I helped found at **Princeton University Press**, I even had the advantage of knowing who the copyeditor was, as he had been a former colleague of mine at the **Press** who now teaches philosophy at **St. John's College** in New Mexico; he was not only a very fine copyeditor but also had a Ph.D. in philosophy, giving him the ability to catch errors of substance that might elude a less well-trained mind. I would myself give high marks to these two authors for the clarity of their prose and organization of their argument. Thus the need for copyediting was minimal. Interestingly, neither copyeditor chose to make the effort to introduce the that/which distinction into the author's writing, which would have resulted in a lot of additional changes beyond the ones they did make.

For the senior author, the copyediting amounted to a very few minor stylistic improvements plus a number of basic formatting changes. I checked the quotations she drew from classic works by **David Hume** and **Henry Sidgwick**, and they were accurate — though the latter was truncated by a few words left out without ellipses indicating any omission. A somewhat longer quotation from **John Stuart Mill's** *On Liberty* differs from the original edition by omitting or changing six punctuation marks, mostly commas, but these do not change the meaning at all, nor can I be sure that the edition the author consulted (published by **Bobbs-Merrill**) had not already changed this punctuation silently for its edition, which I do not have at hand.

For the junior author, copyediting was also done very lightly, with most of it aimed at converting everything to the publisher's house style. No changes of any substantive kind were made, nor any affecting meaning. More problematic for this author, however, was his accuracy in quotation. This article focused on the work of **John Rawls** and quotes frequently from the revised edition of *A Theory of Justice*. I checked every quote from that source against the original and found a number of

Director Emeritus, Penn State University Press
8201 Edgewater Drive, Frisco, TX 75034
Phone: (214) 705-1939 • <sandy.thatcher@alumni.princeton.edu>
http://www.psupress.org/news/SandyThatchersWritings.html

BORN AND LIVED: Washington, DC 8/4/43. Lived in NE Pennsylvania from age four until college.

EARLY LIFE: Attended **Wyoming Seminary** (prep school) in Kingston, PA, and **Princeton University**; took two years of graduate study, in philosophy, at **Columbia** and then **Princeton** again.

PROFESSIONAL CAREER AND ACTIVITIES: 22 years at **Princeton U.P.** (MS Editor, Social Science Editor, Assistant Director, Editor-in-Chief), 1967-1989; 20 years at **Penn State U.P.** (Director), 1989-2009

FAMILY: Wife **Robin**, who has three children, and two children of my own from a previous marriage.

IN MY SPARE TIME: Masters swimming, jazz/rock drumming, sailing.

FAVORITE BOOKS: David Copperfield; *Brothers Karamazov, Lorna Doone*; *The Passions and the Interests* (**Albert O. Hirschman**); *Personal Destinies* (**David L. Norton**); Greek mythology; *Iliad, Odyssey*.

PET PEEVES: People who say "less" when they should say "few."

PHILOSOPHY: Sound mind in a sound body (the Greek ideal).

MOST MEMORABLE CAREER ACHIEVEMENT: Winning the **AAUP's Constituency Award** for my work on copyright.

GOAL I HOPE TO ACHIEVE FIVE YEARS FROM NOW: Making it into the Top Ten national ranking in at least one masters swimming event.

HOW/WHERE DO I SEE THE INDUSTRY IN FIVE YEARS: Moving ever more quickly toward OA (as PDA hastens the death of market-based university press publishing). 🌳

errors, which included giving the wrong page number in one instance, eliding two quotations that are separated in the text, omitting a phrase from another quotation without using ellipses, dropping a word from another quotation, using "affective" instead of "effective" (though the original text should have used "affective" as that was the meaning intended), quoting "lack certain fundamental attitudes" when the original text reads "lacks certain natural attitudes," adding a comma in two instances where none exists in the original, and dropping a word at the beginning of a sentence. All of these errors remain in the version of record because the copyeditor had not compared the quotations in the article with **Rawls's** book. It must be admitted that copyeditors rarely do take the time to check the accuracy of quotations in this way unless they have special reason to be suspicious, though with the availability of many works in the public domain now through **Project Gutenberg**, **Google**, and other readily searchable sites the effort needed to do so is far less than it used to be in the pre-Internet age.

What may we conclude from this analysis? By and large, the copyediting did not result in any major improvements of the manuscripts as they appear at the DASH site. As with the technical editing done for STM journals by

people like **Joe Fineman**, the vast majority of changes made were for the sake of enforcing a house formatting style and cleaning up a variety of inconsistencies and infelicities, none of which reached into the substance of the writing or affected the meaning other than by adding a bit more clarity here and there. Thus it would appear that the DASH versions are probably "good enough" for use by scholars seeking new ideas and information and wishing to keep up with the literature and by teachers who may want to have their students read these versions as class assignments.

More problematic from the viewpoint of scholarly rigor are the errors in citation and inaccuracies in quotation. As noted in my earlier article, mistakes like these have a way of being repeated since people will often not take the trouble to go back to the original texts but merely trust the authors of these articles to have given the right information about page numbers, dates of publication, authors' names, and the like and to have accurately transcribed passages from the sources used. Unfortunately, these are just the kinds of errors that are seldom caught by copyeditors either because the extra work involved in discovering them to be errors is usually not considered to be justified by

the publishers who pay the copyeditors. The errors of this kind I caught in the two articles in political philosophy remain in the versions of record.

Whether the same would be true of the copyediting done on monographs I am less sure because my own experience suggests that more, and more thorough, editing is done on this type of work than on a journal article. And I wouldn't be surprised if articles in the humanities and social sciences receive slightly more extensive editing than that done for the typical STM journal, as described by **Joe Fineman**. Scrutiny of some of the articles in our little survey suggests as much. But I will admit that the differences between the DASH versions and the versions of record are not so great as I had anticipated they would be, and I fear the ill effects of an extended Green OA period less than I did before I conducted this analysis. The worst faults in the system, it seems, remain the shortcomings of scholars themselves that only copyeditors who go the extra mile are likely to be able to help them identify and correct.

Stevan Harnad suggests that in an online world it may make sense to offer copyediting as "an optional extra service" that an author can ask a publisher to provide for a fee, as he envisages that the demand for it will continue to decline, especially if — as **Joe Fineman** argues — "the idea of a house style ought to be abandoned." The question then becomes, what will authors be willing to pay for? Peer review is so essential to the system that it will continue to be required. What benefits of copyediting are most worth preserving? Most authors probably care little, or not at all, about such niceties as using "that" instead of "which" where a restrictive pronoun is the form that most style guides still prefer, or "few" instead of "less" when a number, not quantity, of items is discussed (a nicety that **Mr. Fineman** himself thinks "not worth saving"). Authors should be concerned that their meaning is clear, that their references to sources are consistent, that their English usage is grammatically sound, and that they are spelling words and names correctly. Even more important for the integrity of their scholarship, I would argue, is accuracy in page citation (or whatever begins to take its place in born-digital documents) and in quoting from sources. This type of editing is not done today in STM publishing, as **Fineman** and **Harnad** both observe, and it is rarely done in any journal editing, it appears. That it was not done when publishing was in print form only is understandable, since the extra hours involved in checking such details could easily double the time spent and thus the cost (as free-lance editing is generally charged at a per hour rate). But as mass digitization advances and the ease of checking original sources increases along with it (in the way **Mr. Fineman** indicates), this may well be a role copyediting should begin to play more, as it will help ensure that the integrity of scholarship is maintained at a high level in versions of record. 🌳