Nurse Author & Editor

Volume 20 - March 2010, Issue 1

Detecting and Preventing Plagiarism in Publishing

Charon A. Pierson

Despite its fairly straightforward dictionary definition, plagiarism is a complex concept. Words or ideas are plagiarized if they have been copied from one source and presented as original from another source. Martin (2004) identifies four distinct types of plagiarism: ideas, sources, authorship, and words. Plagiarism of ideas occurs when one fails to acknowledge the origin of a thought, idea or invention and claims the idea as one's own. Plagiarism of sources occurs when one does not attribute the correct sources for a piece of writing or an argument. Plagiarism of authorship occurs when one claims authorship of an entire work without actually fulfilling the criteria for authorship. Finally, verbatim or word-for-word plagiarism occurs when one copies or uses portions of text written by another without clearly demonstrating appropriate quotation and citation standards.

With the advent of electronic access to published material, the emphasis has largely been on detection of verbatim or word-for-word plagiarism with plagiarism detection software. Comparison of a single written document with thousands of articles, websites, newspapers, books, and assorted written works is now possible within minutes. Results are only a "mouse-click" away for most users. Large scale programs that have been used in academia for several years are now emerging in the corporate world. Publishing is among the emerging markets for plagiarism detection software and many nurse authors will discover their manuscripts have been submitted to one of these programs. What does this new technology mean for the average nurse author, reviewer, and editor?

For authors, a thorough understanding of proper paraphrasing and careful attention to details such as the mechanics of proper citation are essential. Citation protocols are used in various referencing styles for the purpose of assisting readers to find the original sources of cited material. Of course citations credit the originator of the thought or the words, but an underlying purpose is to direct readers to the source. Failure to include quotation marks, page numbers, correct author and journal citation data can all lead to a suspicion of plagiarism. Well-written text available on websites, blogs, and other Internet sites may seem a free and easy source of material for a school paper or a manuscript, but these are among the most easily detected writings for plagiarism detection software. An author once sent me a manuscript for this newsletter on plagiarism that was largely taken verbatim and without attribution from websites advertising plagiarism detection software.

For reviewers, the job of detecting plagiarism remains a high priority. Verbatim plagiarism is easily detected by scanning a manuscript with a plagiarism detection software program. What is lacking in algorithms that search the electronic sources is the ability to detect plagiarism of ideas. For this difficult task, we rely on our expert reviewers who know the literature in the field and are familiar with the important ideas that frame the debates. Whether it is a failure to cite relevant research or ideas, the incorrect interpretation of those ideas, or claiming credit for an idea that is already in the literature, expert reviewers are our best resource for maintaining the integrity of our science.

Finally, for editors the task of assuring that plagiarism does not go undetected or unacknowledged is paramount. Articles have been retracted in major journals for duplicate publication, plagiarism, and other ethical infractions. Guidance on how to do a retraction is available from the Committee on Publication Ethics (COPE; Retractions, 2009). COPE, with its current membership of 4770 individuals, corporations, and journals, provides valuable assistance with educating, monitoring and supporting editors, journal staff, authors, and others interested in the ethics of conducting and reporting scientific inquiry. Plagiarism remains a serious ethical issue for journal publishers and many have instituted the use of plagiarism detection procedures.

As an editor, I submit manuscripts to a plagiarism detection program, iThenticate, early in the process before assigning reviewers. The iThenticate program reports a "Similarity Index" as a percentage of total text (see also article by Galliard in this issue). For example, I submitted this article to the program and received a similarity index of 1%; the only duplications were eight words from the reference to COPE above that were found in an on-line journal report. The decision process for the editor is to determine how much similarity is acceptable; this takes time and careful thought. I've found, for example, that research articles may include some questionable sections because the methodology has been reported in the past and it is difficult to create new ways to explain the process. So, a certain amount of similarity is acceptable and unavoidable and the programs can be modified to some degree based on journal preferences. For example, it's possible to exclude references from the analysis because they will always be highlighted and counted in the index. It's also common for terms with multiple words, such as explanations of genetic sequences, to be counted in the index; however, excluding small matches from the report may avoid that problem.

There are no absolute guidelines for what is an acceptable degree of similarity – that is always a judgment call on the part of the editor; nevertheless, I find the software helpful. Although my concern begins at around a 25% similarity index, I review the report to verify that references, small matches, and direct quotes are excluded. For a manuscript with a relatively high similarity index where much of the unattributed text comes from Internet sources, a few words are changed randomly, and complete paragraphs of verbiage are highlighted in the report, my inclination is to

reject immediately without sending the paper out for review. When the similarity index exceeds 50%, I automatically reject. When the report highlights self-plagiarism, most commonly in the methods section, I will continue with the review but highlight the problem in my letter to the author requesting revisions. In review articles, when the similarity index is relatively high and authors fail to add their own interpretations or critiques of the studies, my sense is that the authors have not added anything new to the literature and I provide that feedback to the authors.

For those who wonder if I accuse authors of plagiarizing - I do not. I send a rejection letter to all authors on the paper simply stating my findings and attaching a brief copy of the similarity report as in the following example: "I have rejected your submission without review based on your extensive unattributed verbatim use of material detected by the iThenticate plagiarism detection program we use to scan all submissions. The similarity report, a screen shot of which is attached to this email, found a 63% similarity, which is far too high for journalistic standards." To date, no authors have contacted me to dispute my decisions.

References

Martin, B. (2004). Plagiarism: policy against cheating or policy for learning? Accessed 02/27/10 on-line at **http://www.uow.edu.au/~bmartin/pubs/04plag.pdf**

Retractions: Guidance from the Committee on Publication Ethics (2009). Accessed 02/27/10 on-line at
http://publicationethics.org/files/u661/Retractions_COPE_gline_final_3_Sept_09__2_.pdf

Charon Pierson, PhD, GNP-BC, FAANP, is a Clinical Associate Professor and Director of the Center for Aging at the University of Texas El Paso School of Nursing and the Editor-in-Chief of the *Journal of the American Academy of Nurse Practitioners*. She is also the Editor of *Nurse Author & Editor*.

Print this article   Email it to a friend

Back to Table of contents | **View all articles in this issue**