

Nipping plagiarism in the bud: Using Turnitin to teach novice science writers how to paraphrase

By Christine Parkhurst and Elizabeth Moore

Introduction

Avoiding plagiarism is a problem for authors trying to publish in the health sciences, especially if they are bilingual students writing in English as an additional language. Learning the skills required to paraphrase and cite correctly should begin early on in the author's education. However, composition courses and writing centers rarely address the needs of novice science writers, and science courses rarely address writing skills such as paraphrasing and citation. Using Turnitin, a computer search engine designed for plagiarism detection (<http://turnitin.com/static/home.html>), can be an effective method to learn and practice paraphrasing skills; this method will be discussed here, and compared with conventional classroom and writing center methods.

We will discuss our experience working with novice science writers whose first or best language is not English, using both conventional teaching methods and Turnitin. This will be based on our respective experience: Christine Parkhurst teaching composition at Massachusetts College of Pharmacy and Health Science (MCPHS,) and Elizabeth Moore as a Writing Center Fellow at Barnard College/Columbia University. First, we will discuss how undergraduate health science and science majors in the U.S. learn to paraphrase and cite through conventional classroom and writing center methods, using both MCPHS's more traditional methods of composition instruction and Barnard College's writing center as examples. Then we will discuss how students at MCPHS learn to avoid plagiarism by doing exercises that use Turnitin. We will also discuss the drawbacks and limitations of using Turnitin for this purpose.

Conventional classroom paraphrasing instruction: MCPHS

In the U.S., most composition classes don't teach novice science writers to write from biomedical journal sources, and most science classes don't attend to students' writing. Science majors often exit their composition course sequences without having practiced writing from biomedical sources. This was originally the case at MCPHS. Now students learn paraphrasing and citation by writing on biomedical topics from journal sources, and by practicing with Turnitin (described below.) However, at MCPHS, our students are all health science majors; the pharmacy program is the largest in the U.S. Over thirty percent of the students are bilingual, although only about four to six percent are

non-U.S. residents with student visas. Students graduate with a professional degree in fields such as Pharmacy, Nursing, and Optometry, or they go on to further study in medical school or other professional schools. Some will publish in their fields, and some publish brief monographs even before they graduate. By the time students enter their final two years of the curriculum, they are expected to be able to write on technical topics in their field. This means that students face problems with genre conventions, and many struggle with paraphrasing.

In my composition classes at MCPHS, originally students learned paraphrasing and citation skills by writing on conventional liberal arts topics. When they wrote research papers from biomedical journal sources in their advanced writing course, many students once again began to copy and paste long chunks of text, and some used sources they didn't attribute accurately, or at all. Now students learn paraphrasing from the start by doing summary or synthesis assignments, and then writing biomedical review articles, all based on medical, pharmacy or nursing journal article sources. (How they practice for this by using Turnitin is discussed below.) If students use sources they don't attribute, by using inaccurate references or no references, professors have to play a cat-and-mouse game to find the source or sources the student has plagiarized. Therefore, students are required to submit copies or links to their sources. If students learn early on that they can't copy language from their source texts, they are more likely to avoid the copy and paste style that leads to unintentional or intentional plagiarism. Most student plagiarism is unintentional, though, and checking the student's too-perfect text against the original text often shows that the student has copied too closely. This can be pointed out by underlining matching text, or discussing the too-close text with students working on early drafts. Students improve by doing multiple revisions that are progressively more correct. This is effective, but time-consuming, so not many assignments can be done in this way.

After students exit the class, consistent source checking usually stops. Clinical professors do not normally have time to check student writing against all its declared and undeclared sources. If a professor recognizes blatant plagiarism, the penalties for academic dishonesty are applied, of course. This requires sufficient familiarity with students' writing styles to spot suspiciously perfect or technical text. Clinical professors may ask students to take papers to the writing center, but these professors usually focus on

Nipping plagiarism in the bud

content rather than attending to student writing per se. Many institutions of higher education in the U.S. now have student- or faculty-staffed writing centers that can provide an additional means of detecting and addressing student plagiarism, as well as serving the more general writing needs of the student population.

Writing Center paraphrasing instruction: Barnard College/ Columbia University

Barnard College of Columbia University seems to be using Turnitin primarily as a tool for guarding against intentional, “malicious” plagiarism, but the problem of unintentional or partial plagiarism—especially in the sciences—remains. Barnard’s emphasis on writing across the curriculum is evidenced by its Writing Fellows (discussed below), who interact with and aid students of every subject. It is Barnard’s strong belief that the ability to write coherently, honestly, and originally is a skill essential to every major, including the sciences.

The Erica Mann Jong Writing Center is set up as a resource for undergraduate students at Barnard College and for Columbia students enrolled in classes at Barnard College. The Center is staffed by Barnard students, all of whom are paid a small annual stipend, and all of whom have shown strength in both writing and peer-to-peer communication. Before working with their peers in the Center, all Fellows are required to take a one-semester course called *The Writer’s Process* on the pedagogy of writing and the philosophy behind peer tutoring (though the term “fellowing” is used instead). Writing Fellows are by no means exclusively English majors; in fact, the directors of the writing center make an effort to recruit strong writers from all majors, including the sciences.

Each semester, Fellows split their time between walk-in hours at the writing center and the specific course to which they are assigned. These courses range from first-year English composition courses to advanced Psychology and Environmental Science classes. The latter tend to require the type of technical writing that invites a certain amount of unintentional plagiarism. Fellows are encouraged to look for language that seems implausible or inconsistent with the student’s speaking ability or with her other work. If a Fellow encounters such language, she will often ask the student to see her sources. Kaitlin Kratter is a former Writing Fellow at Barnard who is now a TA for a 1000-person introductory undergraduate astronomy class at the University of Toronto, where she is a graduate student in astrophysics. When confronted with possible plagiarism, Kratter says, “. . . it is often helpful to explain the place for direct quotes, and to ask students to see if they actually understand the words they are using. I have tried to explain that if they do not understand a phrase, . . . they should definitely not include it in their writing.”

In my experience as a Fellow, the students who were most likely to borrow language were also the students who felt

least comfortable with their own writing ability—typically, non-native speakers. To change or paraphrase was, to them, to risk inaccuracy. Though within the Writing Fellows program there exists no protocol for how to deal with science-specific plagiarism, perhaps this fact in itself is evidence of Barnard’s commitment to teaching strong writing across the curriculum. All plagiarism, whether intentional or unintentional, whether related to humanities or to the sciences, is diagnosable using the same techniques.

Both peer-staffed writing centers and Turnitin can be tools for teaching students—in a nonjudgmental way—what constitutes plagiarism. Although Barnard’s students take its honor code seriously, learning to summarize or synthesize articles while using a minimum of borrowed language is an acquired skill.

Using Turnitin to teach paraphrasing at MCPHS

Class assignments which require students to synthesize and paraphrase journal sources raise students’ consciousness about what does and does not constitute plagiarism, and give students practice paraphrasing while writing in genres they will have to master. When MCPHS acquired a Turnitin site license, new assignments made this process more efficient and effective, although not perfect, as will be discussed. Students can do many more assignments, more autonomously; this additional practice helps them master paraphrasing and citation skills.

For each class, I set up a Turnitin class account with a login and password. Students then access the class account and set up their individual accounts. Students write their summary and synthesis assignments using a word processing program such as Microsoft Word. They then access their Turnitin account and enter their assignment in its Inbox. Any number of assignments can be created. There are menu options for each assignment: for example, a due date can be assigned, and papers can’t be submitted after the due date. There is an option to exclude all material in quotation marks from the percent of matching text, which is an improvement over Turnitin’s previous algorithm. I have students simply cut and paste their assignments in, although they can also be downloaded. When students submit a paper, they click on the Submit icon, and Turnitin’s Web crawlers then compare what the student has submitted to every document in its data base.

The result of the comparison is the Originality Report. Turnitin is careful not to label a paper as plagiarized or not. It simply reports any matching text that it found as a percent match. Papers with more than five percent matching text deserve a second look, depending on length. The Originality Report can be viewed by clicking the Report icon. A split screen comes up, with the student’s text on the left and a reference identifying any matching text on the right. The matching text is flagged by being highlighted in color in the student’s text, with a number corresponding to

>>>

>>> Nipping plagiarism in the bud

the source reference on the right. If there are three sources of matching text, for example, there are three colors and three references. Clicking on the reference for the match will bring up the original text and its source. These can then be viewed on the split screen.

It's possible to cut and paste suspicious-looking phrases into other search engines such as Google. What makes Turnitin's Web crawlers faster than Google's for this sort of detective work, is that the student enters the entire text and then Turnitin's Web crawlers do all the comparisons for the entire text at once. Entering suspicious phrases in Google may bring up a match, but each phrase must be entered individually. The Originality Report in Turnitin is generated quickly, either instantly for short texts or after a few minutes for longer texts. The professor can ask students to submit their assignments, and the next day view a list of all the students' documents and their Originality Reports. Any that have a percent match higher than a few percent can then be viewed. A more important advantage of using Turnitin's algorithm is that its Web crawlers find matching text even if the entire sentence doesn't match. Normally search engines only see identical text as a match; they agree with students who think that changing a word here and there is enough to make the text different.

I use the option that permits students to see their own Originality Reports. This turns out to be a powerful tool for teaching students how to paraphrase. Because students know that they will have their text checked by Turnitin, they do their best to paraphrase. They then get an Originality Report. If half the text appears in color with matches, it's a shock to the student who thought that changing fifty percent of each sentence was enough. The impartiality of the computer means that the student can't complain or try to argue that some words were changed: it's not identical, and that should be good enough. Even students who were clearly in the habit of getting high grades on what were essentially plagiarized papers learn to rephrase completely by the third or fourth assignment. Students internalize the experience of rewriting and trying to change the text completely enough so that next time it doesn't match. This year, for the first assignment done by 44 students, almost a quarter of the papers had matching text that ranged from 20%-50%. By the third assignment, only three students had matching text over 5%, and the highest percent match was 18%. As the assignments become more difficult, the percent of matching text still stays low.

Because students see which strings are close enough matches to be flagged, we can then have more substantive discussions about the paraphrasing spectrum. On one end are chunks of text that match but should not be paraphrased. Some are easily recognized (ex. "Center for Disease Control and Prevention.") Students try to paraphrase other strings that really shouldn't be paraphrased, but students are not yet familiar enough with the genre to recognize that. For example, "new prescriptions for selec-

tive serotonin reuptake inhibitors," flagged as a match, should not be paraphrased "novel prescriptions for medications that prevent the reuptake of serotonin selectively." Although "new" and "novel" generally mean the same thing, in this context "novel" would amount to a claim that the prescription is "interesting and innovative" as opposed to a "new" rather than a "refill" prescription. Students more familiar with the genre would not attempt to paraphrase SSRI, but would recognize it as a compound noun. Further along the spectrum are chunks consisting of lists of technical terms that are symptoms or side effects, for example; these also can't be paraphrased without distorting meaning ("[This may cause] dizziness, drowsiness, dry mouth, nausea or vomiting.") Still further along are chunks of text that should be paraphrased to the extent possible, but are difficult to paraphrase without distorting meaning (for example, methods sections or data-rich results sections.) When students have used Turnitin enough to paraphrase automatically whatever they can easily paraphrase and should paraphrase, the remaining matching text is the correct starting point for a discussion of how to paraphrase and cite correctly, and how to make sure that the resulting text is still an accurate and genre-appropriate representation of the original meaning. This requires interaction with a human, not a computer.

However, there are caveats. Turnitin adds students' texts to its data base, and this is controversial in terms of intellectual property rights. Since I have students do exercises to practice paraphrasing using Turnitin, this seems less like depriving them of intellectual property, and they benefit directly. Also, most of the medical, pharmacy, and nursing journals students work from are not part of the Turnitin data base. For paraphrasing assignments, I use short excerpts from articles (less than 10%,) one time only, and enter those excerpts myself so that I know they will come up as matches. Turnitin permits use of non-technical sources on biomedical topics as a lead-in to more difficult journal sources. Using another search engine such as eTBLAST that is tailored to searching biomedical texts would deal with this issue for more advanced writers. Another problem is that Turnitin is confused by many sentence-level grammar errors. Students' text may not come up as a match when it should, if Turnitin is unable to parse it. Therefore, work with 0% matching text should also be checked to see if sentence-level grammar is the issue. Finally, the content of student paraphrases may no longer be accurate once the language is sufficiently dissimilar. This is, of course, a major problem in science writing, but it can reveal underlying problems with students' reading comprehension which can then be addressed.

Conclusion

Conventional classroom and writing center methods for teaching students to avoid plagiarism can be effective, but are time-consuming. Therefore, students may not get as much practice as they need. Professors' lack of time to do

consistent source-checking, and lack of familiarity with students' actual writing styles, may allow intentional or unintentional plagiarism to slip through unrecognized. This reinforces students' misperception that plagiarism is acceptable, or not likely to be spotted, and they continue to plagiarize difficult text. Using Turnitin for summary and synthesis assignments based on technical sources such as journal articles saves time by doing automatic source-checking for matching text. Because this is much less time-consuming than comparing sources and student text by hand, students can do many more assignments. Turnitin is consistent in flagging matching text. The additional practice helps students learn to paraphrase automatically whatever can be paraphrased easily. More advanced students begin to develop genre awareness and write in a more genre-appropriate style. Discussion of the remaining matching text helps students recognize text that cannot or

should not be paraphrased, understand the issues concerning hard-to-paraphrase text such as symptom lists or methods and results sections, and become conscious of the importance of paraphrasing in a way that doesn't distort meaning.

Christine Parkhurst

Massachusetts College of Pharmacy and Health Sciences, Boston, Massachusetts, USA
Christine.Parkhurst@mcphs.edu

Elizabeth Moore

Morgan Library, New York, New York, USA
lizmoore1234@gmail.com

Christine Parkhurst is an Associate Professor of English and Humanities at Massachusetts College of Pharmacy and Health Sciences, Boston, Massachusetts, where she teaches composition and health communication. Elizabeth Moore was a Writing Fellow at Barnard College, and is currently an Assistant Editor at the Morgan Library, New York, New York, and a free-lance writer.



NAWP CENTENARY ESSAY COMPETITION 2006-2007

A rare opportunity for to put your professional opinions into print

To mark its successful centenary year in 2005, the UK **National Association of Women Pharmacists** is launching a Centenary Essay Competition; open to all pharmacists registered in the UK and Europe. Agreed funding means that only pharmacists can be awarded the financial prizes, **but NAWP also welcomes entries from others individuals with relevant expertise** and they too will have their entries considered for publication and commendation. This will be a once only event, inviting entrants (men and women) to write a literature-based essay of 1200 words on a specified topic. The Pharmaceutical Journal has agreed to consider the prize-winning essays for publication. Winners will be invited to present their work to a NAWP meeting, subject to their willingness and it is planned that all high quality entries will be published by NAWP.

There are two classes of entry: **Academic** and **Professional**.

Entries in the Academic class are to be based primarily on research and review literature published in pharmacy, scientific, medical, nursing and/or behavioural journals.

Entries in the Professional class may draw from these sources if appropriate, but may also draw from professional experience and media sources (magazines, newspapers, TV etc).

Entries are to be submitted by 1st February 2007, in English. The competition will be judged by three judges appointed by, but independent of, NAWP.

The topic of the essay is:

Pharmacy and ageing: Are medicine-taking issues under-recognised in behavioural research?

Background to the topic

The media frequently report study-findings that claim to show that mental function in older people can be influenced by activities such as doing crossword puzzles, taking exercise or socialising. However, pharmacists are well aware that in many countries ageing is associated with increasing use of medication and that this in turn brings benefits and inconveniences. Prevalence of the use of prescribed and non-prescribed medicines, the number of dosage units being taken daily, side effects, and the efficacy of long-term medication are issues that are natural concerns of pharmacists, but awareness of these points comes less naturally to some other disciplines concerned with health and social issues. Authors are therefore asked to question whether behavioural studies investigating mental function in older people adequately recognise the consequences of high levels of medicines usage by those over the age of 60.

For further details and an entry form, contact enquiries@nawp.org.uk or write to Mrs Brenda Ecclestone, Hon Sec. NAWP, Princess Royal Cottage, Butterow West, Rodborough, Stroud, Glos. GL5 3UA