Avoiding plagiarism in academic writing: a guide on what to cite and how

KG Behrens
Steve Biko Centre for Bioethics, Faculty of Health Sciences, University of the Witwatersrand, Parktown, South Africa.

Correspondence: Dr Kevin Gary Behrens, Steve Biko Centre for Bioethics, University of the Witwatersrand, Private Bag 3, Wits 2050, South Africa. e-mail: kevin.behrens@wits.ac.za

ABSTRACT
This article defends the conventional view that plagiarism in academic writing is a form of publication misconduct. I emphasise that what constitutes plagiarism is presenting the intellectual product of another as if it is one’s own. Common knowledge does not need to be cited, but verbatim text, data, results, distinctive arguments, organisational structures and ideas, do. What academic authors would like to have protected are their distinctive or novel contributions to their fields of knowledge, whatever form those take. Those who plagiarise act unethically because they seek to obtain credit that is not due to them and they deny credit to those to whom it is due. Self-plagiarism is another form of academic misconduct that entails seeking to obtain credit for the same work more than once. As with plagiarism, it is seen as dishonest and fraudulent. Avoiding accusations of plagiarism can be easily achieved by following a few simple conventions of referencing and punctuation, and academic authors are advised to protect their own reputations by following these conventions.

Keywords: self-plagiarism, scholarly ethics, publication misconduct, fraudulent writing

INTRODUCTION
The incidence of discovered plagiarism in published academic work has grown significantly in recent years. Most cases could easily have been avoided if the authors had simply followed some basic conventions. Since the consequences of being identified as a plagiarist can be very serious, it is important that all authors – novices and the more experienced – understand what plagiarism is and how to avoid it. In this article, I give a brief account of what plagiarism is, what sorts of things authors should cite, and how they should do so.

The Committee on Publication Ethics (COPE) is an international association of editors with a membership exceeding 9 000. Its main purpose is to advise publishers and editors on publication ethics and on how to handle cases of misconduct. COPE has produced a set of international standards for authors of research publications. Section 4 of the Standards document deals with originality and includes the following standards:

“Relevant previous work and publications, both by other researchers and the authors’ own, should be properly acknowledged and referenced. The primary literature should be cited where possible... Data, text, figures or ideas originated by other researchers should be properly acknowledged and should not be presented as if they were the authors’ own. Original wording taken directly from publications by other researchers should appear in quotation marks with the appropriate citations.”

This represents the conventional view that authors who present the data, words or ideas of others as if they were their own, are guilty of publication misconduct. This kind of misconduct is usually termed plagiarism. It also reflects the standard view that authors should not improperly re-use their own work unless this is correctly acknowledged and referenced. This kind of misconduct is often referred to as self-plagiarism.

COPE has also produced international ethical standards for editors. In this document they assert that “Editors should respond to all allegations or suspicions of research or publication misconduct raised by readers, reviewers, or other editors” and that “beyond the specific responsibility for their journal’s publications, editors have a collective responsibility for the research record and should act whenever they become aware of potential misconduct if at all possible.” With respect to plagiarism in particular, they recommend that editors screen submissions for possible plagiarism and redundant publication, making use of appropriate software tools and that, where this is detected, the matter should be pursued with authors. In cases of very serious plagiarism they even recommend retraction of articles.

COPE’s position reflects the conventional view that plagiarism is a form of academic misconduct and that it is regarded as a serious breach of scholarly ethics. Authors who are found guilty of plagiarism face many possible consequences, including having their misconduct exposed publicly, having to revise their work, and even having their articles retracted. Plagiarism can also negatively affect decisions about hiring, promotion and tenure, and always casts a shadow over the integrity and character of the perpetrator. It is very likely that a great deal of plagiarism goes undetected and that even when it is found, some authors are protected by institutions and editors who prefer not to expose the misconduct. But it is a risk that is not worth taking and authors are advised to follow standard academic conventions and avoid opening themselves to accusations of plagiarism.
It is a matter of some dispute whether plagiarism is increasing or whether it is simply being detected more frequently than in the past. What is clear, though, is that more and more journals and institutions are deliberately taking steps to avoid publishing plagiarised work, and that the widespread use of plagiarism detection software tools is making it more difficult for authors to get away with plagiarism.

WHAT IS PLAGIARISM?

For authors to successfully avoid plagiarism, it is obviously necessary for them to understand what it is. There are many different definitions for plagiarism and it can be difficult to understand a concept when there are varying ways of defining it. Recognising this problem, Helgesson and Eriksson have researched the notion of plagiarism and analysed its essential nature. They propose the following definition: “Plagiarism... [is] an instance of using someone else's intellectual product (such a texts, ideas, or results), thereby implying that it is [one’s] own”\(^5\). This definition implies that the essential characteristic of plagiarism is that an author presents the work of another in a way that implies that it is his or her own. In turn, this implies that it is not plagiarism to use the work of another, as long as it is done in such a way as to make it clear to the reader that this work has been authored by someone else.

The other key element in this definition is the use of the phrase “someone else’s intellectual product”\(^5\) (my emphasis). This is exceptionally important in determining when it is necessary to cite. The phrase suggests that the kind of work of another author that must be correctly acknowledged is work that is the product of that author’s independent intellectual labour. What counts is that the author has applied his or her intellectual acumen to some field of knowledge in a way that makes a novel or unique contribution to that field. If all that an author has done is to rehash or describe existing knowledge, there is no need to cite that author. No intellectual labour is expended in merely relaying what is already known or describing what has already been written. Consequently, there is no need to cite what might be said to be “common knowledge” in a particular field. For example, it is now common knowledge that individuals whose work exposes them to asbestos over a prolonged period of time are likely to develop asbestosis. If you are writing an article about asbestosis, and you have read another article that makes this claim, it is not necessary to cite that article. After all, the author of the article you have read has simply made a claim based on existing, common knowledge. On the other hand, if it is known who did the original research that established a causal link between occupational exposure to asbestos with the disease we now routinely call asbestosis, both you and the author of the article you read that made this claim ought to acknowledge the intellectual product of those scientists who originally made the association. This is not difficult to do. All you need to do is identify the researchers who made the association by name and reference their relevant publication(s). Thus, the basic principle that needs to be applied is that one needs to cite anything that is distinctive to a particular source.\(^6\) What counts as distinctive is anything novel, any intellectual invention, innovation, original discovery, previously unknown data, unique argumentative strategy, distinctive organisational structure, new analytical approach, or distinctive expression. All of these things are the results of the intellectual labour of some other person. They are the intellectual products of another and must therefore be properly acknowledged. What is common knowledge does not need to be cited because it is not the intellectual product of any specific person.

Helgesson and Eriksson unpack the notion of intellectual product as consisting of “texts, ideas or results”.\(^5\) In what follows, I discuss the need to cite each of these, reversing their order.

CITING RESULTS

The most obvious thing that needs to be cited is distinctive data (in Helgesson’s and Eriksson’s definition of plagiarism, this is covered by the use of the word “results”). This is the kind of knowledge that can only have been acquired by means of some kind of empirical research. It is obvious that research results are the “intellectual product” of some person. Thus, all newly discovered facts, statistics, evidence-based conclusions, tables and graphs (and conclusions drawn from these) should be accompanied by a reference.\(^6\) Where the data are the product of the author’s own research which is being reported in a manuscript, the report should indicate how this information was discovered through the research, and give an account of the evidence upon which it is based. The source of any data that is not the result of the author’s own research
must be appropriately acknowledged. One reason for this is that it is plagiarism not to reference the source. But there is another equally important reason: when an author relies on facts or hard data from other sources, citing the source provides secondary “evidence” and lends authority to the information. Readers realise that the proffered data is not based on the author’s own research and might like to verify it. A reference enables them to do so.

CITING IDEAS

The necessity to cite distinctive data is relatively uncontroversial. Less so is the need to cite distinctive ideas. There has been much debate about whether it is appropriate to regard ideas as intellectual property.\(^5\) After all, it is quite possible that different people could come up with the same ideas quite independently of one another. Ideas are also not as tangible as data. Yet, new ideas undoubtedly qualify as someone’s “intellectual product” and, as such, it is regarded as plagiarism to pass off someone else’s ideas as one’s own. Clear cases of ideas that should be cited include a new evidence-based theory, a novel treatment protocol, or a new scientific method. It is easy to identify these ideas distinctive of a particular source. But there are other kinds of ideas that are equally the intellectual product of another, even though they are not as easy to identify as such. For instance, an author might have come up with a distinctive intellectual organisational strategy, such as a novel approach to the classification of occupational diseases. Another example is the Nuffield Council on Bioethics’ Public Health “Interventions Ladder” that assists in the classification of various kinds of public health interventions based on the extent to which these interventions limit the freedom of citizens.\(^7\)

“. . . authors should not improperly re-use their own work unless this is correctly acknowledged and referenced.”

A novel, robustly-defended argument can also be a distinctive idea. Helgesson and Eriksson argue against the conventional view that what makes plagiarism wrong is that it entails stealing the work of another.\(^5\) They point out that when something is stolen, the original owner is deprived of that thing. If I steal your copy of a book, I now have it and you do not. This is not the case when I present your intellectual product as my own. I do not deprive you of that product; but I deprive you of the acknowledgement that it was your intellectual labour that resulted in the product. They go on to argue that what is essentially wrong with plagiarism is that it is a form of dishonesty. On presenting someone else’s work as one’s own, one is misleading others and taking undue credit.\(^5\) Hegelsson and Eriksson’s argument that plagiarism is not really theft is a distinctive argument, the result of the authors’ intellectual labour, and should therefore be cited.

Ideas do not necessarily have to represent a quantum leap in knowledge, or a groundbreaking new discovery, to be distinctive. In fact, most often, new ideas represent no more than a small, incremental shift from, or refinement of, existing knowledge. For example, an author might argue for the need for another rung to be added to the Nuffield Council’s ladder. This would represent little more than minor refinement of an existing framework, but it would still be the consequence of the author’s intellectual labour and, as such, should be cited by others who might wish to use the idea.

CITING TEXT

The need to cite the words of others is generally much more contentious, particularly in the scientific disciplines. Researchers in the sciences often question the need to cite the words of others at all. They claim that they would have no objection to other authors using their words, as long as they acknowledged and cited their data and findings. I grant that, in some disciplines, the intellectual product that researchers most want to protect is their findings. After all, in these fields, the novel contribution made to existing knowledge is primarily the new data that are obtained from empirical research. In contrast, in other disciplines, such as the humanities, it is sometimes precisely their words that authors want protected. In these fields, Chrousos et al. point out that “the wording is the essence of the novelty”.\(^8\) It seems fair enough to hold the view that, to some extent at least, the importance of citing the words of others is contextually dependent on the academic field. However, even in the sciences, the careful construction of lucid, unambiguous text often takes hard work, and scientific researchers should be as anxious to protect this intellectual product as any other authors. Be that as it may, the standard conventions of academic writing dictate that, if you use the words of another, verbatim, you should enclose those words in quotation marks (a simple convention, relying on nothing more than some punctuation), and you should cite the source. Doing this does no harm but does a great deal in protecting you from accusations of plagiarism.

My own view is that it matters little whether you agree with the standard conventions of academic writing. Most academic publishers these days routinely screen submissions by means of software that checks for similarities in text. Where authors use the words of others, verbatim, without using quotation marks and providing references, they open themselves to accusations of plagiarism. Given that the consequences of this can cause embarrassment and result in a stain on their academic integrity, for purely pragmatic reasons, authors should avoid this.

I should also point out that it is not only the use of word-for-word passages of text without proper attribution that is problematic. Sometimes authors try to get around this by using a sentence or two from another source, retaining exactly the same sentence structure and simply replacing a few of the words with synonyms. This is known as word substitution or thin paraphrasing, and is also regarded as a form of plagiarism, especially when the source is not cited. It is best to avoid this practice. If you paraphrase someone
else’s text, it should be substantially re-written in your own words. You should still cite the source if there is anything distinctive in the ideas of the original author.

**SELF-PLAGIARISM**

Many have pointed out that self-plagiarism is something of a misnomer. If plagiarism is presenting the work of another as your own, then self-plagiarism seems impossible. However, even if the term is questionable, the basic idea that there are conventions that apply to the re-use of one’s own work is not. Since academic status is measured in terms of the number and quality of published works, authors who publish substantially the same work on multiple platforms may be regarded as dishonestly trying to get more credit than is due to them. Re-use of already published work should be kept to a minimum and you should be transparent about the fact that this is re-use. If you use verbatim passages from your own work, quote and cite yourself as you would any other author.

Another ethical concern with self-plagiarism arises in contexts (such as South Africa) where subsidies are paid to institutions for articles published in accredited journals. In such situations, an author who publishes the same work in more than one journal is not only unfairly seeking to get credit for the same work more than once, he or she is also guilty of a form of financial fraud, in that more than one amount of subsidy is claimed. Furthermore, academic journals are essentially in the business of publishing original research. Consequently, they typically require authors to attest to the fact that their submission has not been published before and is not under consideration by any other journal. Thus, authors who do publish substantially the same work more than once are also guilty of having made a false declaration to at least one of the journals to which they submitted their work. Even more concerning is that journals typically rely on the good will and collegiality of reviewers who review submissions without remuneration. They do this in order to contribute to the advancement of knowledge in their fields. An author who submits the same work to more than one journal is surely abusing the generosity of colleagues who are prepared to give freely of their time and expertise only because they think they are contributing to the advancement of scholarship. Self-plagiarism that involves duplicate publication of the same work or re-using very large sections of text verbatim also constitutes the infringement of copyright which is not only unethical but also unlawful.

**CONCLUSION**

Earlier I referred to Helgesson and Eriksson’s argument that what is wrong with plagiarism is that it entails academic dishonesty, misleading others so as to obtain undeserved credit. I think this is correct. What is also wrong with it is that it deprives other authors of credit that is due to them. Academic authors are measured in terms of their intellectual product. As a matter of collegial respect it is important that we acknowledge one another’s distinctive contributions.

There are a number of simple conventions of punctuation and referencing that need to be followed to avoid plagiarism. My advice to authors is just to do it right.

**DECLARATION**

There are no conflicts of interest to declare.

**REFERENCES**