niche
science&technology

Good But Not Original:

An Insider's Insight into Plagiarism

The issue of plagiarism is of growing concern within biomedical and life science research. As a result, most major academic publishers now use plagiarism detection software. Importantly, even in cases of coincidental similarity, when detected, the author is always assumed to be guilty of attempted deception.

Can you navigate safely through this important stage of the article submission process? Do you know enough about plagiarism to avoid unwarranted accusations of intellectual theft? Are you at risk of over-zealous plagiarism detection software?

We provide some insights from the Niche medical writing team who have been writing manuscripts for academia and industry since 1998.

Before you start

- Establish the originality of the planned work. You should be particularly aware of possible overlap with your own previously published work
- Know when to quote and how to cite. If you can do these two things, you will be able to use a wide range of different sources in your work without concern
- Read the plagiarism policies of your own institution and any relevant agencies or professional bodies. Definitions vary; if you are writing for a specific journal, you should also read their guidelines placing emphasis on what they say about plagiarism.

Prepare to succeed

- When reading source documents, keep them well-organised and highlight important passages as you go. This will help you to keep track of what you need to cite in your final manuscript
- Never cite articles that you have not read yourself. Always check that the source article aligns with the proposed interpretation if you plan to include an opinion from a review article or online resource that reproduces information from somewhere else
- Avoid relying on just one or two major review articles for your background information. You want to minimise any chance that your narrative appears similar to existing texts

Key Insights

Plagiarism has long been a source of contention in research – Newton and Leibniz spent decades accusing each other of stealing ideas [1]. Time has moved on, but the problem remains. Today, although the incremental advances in science may not be as great, the challenge of plagiarism is bigger than it ever was. In biomedical and life sciences research, almost one in four article retractions are due to plagiarism [2].

Nowadays, global access to digital libraries provides instant access to information on almost any subject. Digitisation has also made it easier for researchers to use scholarly publishing to advance their career and academic standing. However, by using word processing tools to 'convert' the easily accessible information into new 'copy', less scrupulous scientists are able to transform the words of others into 'new' publications with little addition of novel thought, opinion or wisdom. This poses a new challenge to the ever-increasing number of content-hungry journals. On the one hand, journals are aware that their credibility would be significantly damaged if they published an article that incorporated plagiarised content; on the other, they want to avoid the high costs associated with a lengthy, in-depth review process. Furthermore, given that articles appearing online will be subject to many years of in-depth scrutiny, it is highly likely that any plagiarism will be discovered eventually.

Journals have responded to these challenges by employing software that can check the articles they receive for plagiarised content [see Detection of Plagiarism]. The consequences of this for authors can be significant. Suspected plagiarism is automatically considered to be deliberate and the onus is placed on the author to demonstrate their innocence – assuming, that is, that they are given the opportunity to explain. Where plagiarism is reported after publication, a journal will often retract the paper and, because it implies unethical practices, this may have implications for the author's career.

Plagiarism detection software makes its assessments based on the statistical likelihood of a particular combination of words occurring by chance. Sensitivity is balanced between correctly identifying plagiarised texts and finding false positives. Irrespective of the statistical unlikelihood, there remains the possibility, however remote, of an author unwittingly mirroring a combination of words that have been used elsewhere. This possibility can only increase as more and more articles enter the scientific literature.

As well as coincidental similarity, authors also need to be aware of plagiarism resulting from improper quotation and incorrect citation.

> "To steal ideas from one person is plagarism: to steal from many is research."
>
> Steven Wright, Comedian

When to Quote

The most ubiquitous form of plagiarism occurs when an author takes somebody else's words and presents them as their own. This might simply be a consequence of accidental ambiguity (Example 1). Unless you use quotation marks correctly, it is not clear whether or not the words are the author's own even when the original text is cited. For this reason, verbatim text must always be presented in quotation marks.

Depending on the style of the publication, quotes can be inserted in-line (Example 2) or, usually in the case of longer quotes, on their own lines (Example 3). Remember, it is still considered plagiarism to reproduce a quote but with a few word changes. Any sentence that is clearly derived from another constitutes plagiarism whether cited or not (Example 4).

Example 1	It is important for perfumers to remember that a rose by another name would smell as sweet (Shakespeare, 1597).	✗
Example 2	According to Shakespeare (1597), "a rose by any other name would smell as sweet".	✓
Example 3	"A rose by any other name would smell as sweet" – Shakespeare, 1597	✓
Example 4	Perfumers should always remember that, regardless of what you called it, a rose would smell just as good (Shakespeare, 1597).	✗

Rephrasing and Paraphrasing

You are required to provide the source of the original concept even if you use your own words to explain someone else's work (Example 5 and Example 6). Even when the text has almost no words in common with the original quote, it should still be cited (Example 7).

Example 5	The name given to a flower does not affect the odour it produces.	✗
Example 6	According to Shakespeare (1597), the name given to a flower does not affect the odour that it produces.	✓
Example 7	Names are unimportant (Shakespeare, 1597).	✓

What Not to Cite

It is not necessary to provide a citation for well-known pieces of information that are considered common knowledge. A good rule of thumb to remember is that citation is not required for anything that can be reliably found in general, university-level textbooks.

However, we would advise that if you are unsure whether you need to cite, you should always err on the side of caution. If something is truly common knowledge, it should be easy to find citations to support it.

Copyright ©

Plagiarism is generally considered to be an ethical problem and not a crime. However, when it overlaps with copyright infringement, plagiarism becomes a legal matter. A detailed discussion of copyright law is beyond the scope of this guide, but the main thing to remember is that publishers own the copyright of the academic papers that they publish. This means that their permission – and not that of the authors – is required before reproducing any of the content from an article.

A certain amount of quotation tends to be accepted under the academic fair-use rules, although permission must be obtained to reuse items such as figures, diagrams and tables. Remember, this is true even if YOU are the original creator of the diagram! Remember that obtaining permission takes time – you need to build that into your time lines.

Is Plagiarism Ever Acceptable?

Researchers frequently make use of published protocols when conducting their experimental work. When writing up the research, can the existing procedure be reproduced in the methods section of a manuscript? This is especially relevant when the original protocol is written in such a clear, precise way that to rewrite it would make it less accurate. To avoid confusion, we would recommend that the method be rewritten in the author's own words even though direct copying could help the reader and aid experimental reproducibility. There have even been reports of researchers being accused of plagiarism after their Methods section included their own laboratories standard 'boilerplate' methods [3]. The best advice is to show particular care and introduce differences when you are describing similar work.

Self Plagiarism

Our understanding of plagiarism mainly focuses on the uncredited use of other people's work. The idea of self-plagiarism seems counter-intuitive but it is becoming an issue. All researchers face great pressure to publish frequently. Furthermore, if an author is identified as an expert in a particular field they are often hounded by different journals to provide comment and re-comment on a particular topic. In response, it has become common for academics to submit many similar pieces of work to different journals over several years. This sometimes involves re-use of their own data and this so-called 'salami slicing' can sometimes result in authors plagiarising themselves. Self-plagiarism is treated very seriously and over 14% of retractions from the literature in the biomedical and life science fields are the result of what are considered duplicate publications [2].

In one famous case, a researcher at Queen's University in Kingston, Canada was found to have used uncited, self-plagiarised material in at least twenty papers [4]. Although many of these papers were ultimately retracted, the decision was not as clear-cut as would have been expected in a 'standard' plagiarism case. The relevant university authorities were first notified of the duplication in 2005, but it was not until 5 years later that papers started to be withdrawn [5]. In a manner that further demonstrates the confusion that surrounds self plagiarism, several members of staff at the university became embroiled in a dispute that eventually included multiple internal investigations, widespread media coverage and the issuing of a campus banning order to one of the whistleblowers.

The above example should serve as a cautionary tale; if you reuse an idea or some data from one of your previous papers, the original article must be cited. Furthermore, when working on large projects, publications should be planned carefully to avoid overlap. This is particularly important when several collaborators intend to publish their own articles using the results from a single research program. One should also remember that, nowadays, conference proceedings are widely available online and uncited re-publication of their contents in full papers could constitute self-plagiarism; in fact, this issue was central to the case discussed above [4]. We must also caution authors when they are writing about experiments conducted to confirm earlier findings and challenging/or responding to points through the letters sections of journals.

> "I get a lot of big ideas, and occasionally I actually come up with one myself."
>
> Baurard, Author

Detection of Plagiarism

Almost all reputable journals now use plagiarism detection software in the early stages of the submission process. The need for anti-plagiarism software was first identified in the context of school and university coursework. The amount of information available online made it possible for students to simply copy text from the internet. Online communication also made it simple for students at different institutions to share essays with each other. One of the first and most popular solutions to these problems was Turnitin, a service which compares texts with a database of previously scanned texts as well as the wider internet [1].

Following its academic success, Turnitin was adapted for the scholarly publishing world as the commercially available service iThenticate, which compares submitted manuscripts to its database of 52 billion web pages and 42 million books and journal articles [6]. The database behind iThenticate is kept up-to-date with newly published material. After comparison, the software reports any similarities between the tested article and any text in the database. According to its manufacturer, iThenticate is now used by a third of all journals [6].

Research suggests that iThenticate fails to identify a plagiarised text when at least every seventh word is replaced with a synonym (e.g. replacing 'took part" with 'participated") [6]. However, the same researchers also demonstrated that some anti-plagiarism systems could identify duplicate text even after the authors replaced every fifth word [7]. If you are concerned about plagiarism checkers and possibly being wrongly accused of plagiarism, try testing out one of the many free-to-use services that can be found online [see Helpful Resources]. For a more representative result, remember to exclude bibliographies from the submitted text; they are likely to cause false positives because citations are very likely to appear elsewhere in exactly the same format.

Anticipatory plagiarism or false positives?

'Anticipatory plagiarism occurs when someone steals your original idea and publishes it a hundred years before you were born." So says Robert K Merton. To test the potential for software programs to provide false positives we took a tract of text written by one of our medical writers in 1998 but never published or submitted for publication. The author has published over 60 papers in the intervening years and worked on many others. When the unpublished work was tested using an online resource it was identified as having been plagiarised. False positives do happen, the software is not always correct. Be prepared to defend your work by keeping detailed records on what you have done.

Responsibility of Plagiarism

According to the International Committee of Medical Journal Editors, all authors named on a paper 'should have confidence in the integrity of the contribution of their co-authors" [8]. This is as true in the context of plagiarism as it is for any other type of scientific misconduct. You should always make sure that all of your colleagues understand what plagiarism is. Junior co-workers will perhaps require the most instruction but senior co-workers may also need to be reminded of the risks of self plagiarism.

A case from Kansas University illustrates how the actions of one member of staff' can have serious repercussions for many careers [3,9]. A researcher was found to be routinely preparing manuscripts by stitching together extracts from several sources. A more senior member of staff identified the problem and explained the problem with this approach. Unfortunately, his concerns were ignored and an article containing plagiarised material was published under both of their names. Some time later, the original author of some of the plagiarised material alerted the relevant journal, the paper was retracted and both of its authors were publicly censured by their institution. Neither author remains in academia [9].

In this example, the senior member of staff was effectively punished for failing to follow his institution's plagiarism policy, which stated that personnel had a responsibility to report all plagiarism to university administration. By not following this, the author was found to be guilty of plagiarism by association [3].

And finally...

A useful way to avoid accidental plagiarism is to find your own writing style. Although academic writing should principally focus on clarity and content, writing in a voice that comes naturally to you will reduce the likelihood of coincidental similarities with existing work - though you will still be prone to self plagarism.

Remember, if you work carefully and have a working understanding of what plagiarism is, it is unlikely that you will commit it accidentally. Therefore, although the penalties are severe, you should not be afraid of using a wide variety of sources.

References

1. Heather J. Turnitoff: Identifying and Fixing a Hole in Current Plagiarism Detection Software. Assess Eval High Educ 1981; 35:647–660.
2. Fang FC, Steen RG, Casadevali A. Misconduct accounts for the majority of retracted scientific publications. Proc Natl Acad Sci USA 2012; 109:17028–17033.
3. Lushington GH, Chaguturu R. A systemic malady: the pervasive problem of misconduct in the biomedical sciences. Drug Discovery World, Summer 2015. http://www.ddw-online.com/summer-15/p303675-a-systemic-malady:-the-pervasive-problem-of-misconduct-in-the-biomedical-sciences-part-2:-detection-and-prevention.html.
4. Reich ES. Self-plagiarism case prompts calls for agencies to tighten rules. Nature 2010; 468:745.
5. Rosen J. Queen's professor has papers retracted. Queen's University Journal, 28 July 2015. http://www.queensjournal.ca/story/2015-07-28/news/queens-professor-has-papers-retracted [accessed 8 August 2015].
6. Turnittin, LLC. Plagiarism Detection Comparison Database. http://www.ithenticate.com/content [accessed 8 August 2015].
7. Gillam L, Marinuzzi J, Ioannou P. TurnItOff – defeating plagiarism detection systems. In: 11th Higher Education Academy-ICS Annual Conference, University of Durham, 24–26 Aug 2010, UK.
8. International Committee of Medical Journal Editors. Defining the Role of Authors and Contributors. http://www.icmje.org/recommendations/browse/roles-and-responsibilities/defining-the-role-of-authors-and-contributors.html [accessed 8 August 2015].
9. Reich ES. US authorities crack down on plagiarism. Nature News, 11 January 2012. http://www.nature.com/news/us-authorities-crack-down-on-plagiarism-1.9776.

Helpful Resources

www.ithenticate.com – The most widely used plagiarism detection software (subscription required for use)

www.plagscan.com/seesources – A useful, free plagiarism checker

www.grammarly.com/Plagiarism_Detector/ – Another pay-to-use plagiarism detector

www.scanmyessay.com – Offers a free, downloadable plagiarism detector called Viper

Next Steps

We created this Insider's Insight into Plagiarism to share a few helpful tips and pointers that we have gained over the years.

I hope you found our guide useful, if you would to discuss support with any of your publishing challenges please contact me at the email address below.

Dr Justin Cook
Head of Medical Writing
Justin.Cook@niche.org.uk

Get in touch

niche science&technology

+44 (0)20 8332 2588
www.niche.org.uk