

Forgive me for repeating myself: Self-plagiarism in the medical literature

Stephen Gilliver

Center for Primary Health Care Research, Sweden

Correspondence to:

Stephen Gilliver,
Science Editor, Center for
Primary Health Care
Research, Malmö,
Sweden
stephen.gilliver@med.
lu.se

Abstract

While plagiarism of others' work is universally condemned, authors' reuse of their own words and data (so-called 'self-plagiarism') is a far more contentious issue. The recycling of one's own text, in particular, polarizes opinion: some consider it unacceptable, whereas others don't see anything wrong with it at all. This being so, it is unsurprising that there are no widely adopted guidelines outlining which (if any) and how much text may be recycled. My aim in writing this article is to briefly introduce the different types of self-plagiarism; to present the views of journal editors and other interested parties and describe ways in which the former are combating abuses; and to highlight some of the steps authors can take to avoid trouble.

Keywords: Plagiarism, Self-plagiarism, Text recycling, Data recycling

While tales of students and researchers passing off others' words as their own are commonplace, the issue of scientists plagiarising their own work hasn't created anything like the same hoo-ha. Indeed, some question whether there's anything wrong with it at all.

'Self-plagiarism' means different things to different people. While to some it is the republication of one's published data in a modified or unmodified form (so-called 'data recycling'), others would include the reuse of one's old text ('text recycling') in their definition.

An editorial in *The Lancet* from 2009¹ makes a clear distinction between data recycling and text recycling, referring to the former as 'unacceptable' and the latter as 'less of a crime'. However, responding in the same journal, Iain Chalmers² of James Lind Library, Oxford rejected the idea that reuse of one's own words is necessarily a bad thing, claiming

that getting an important message across outweighs the interests of editors and publishers.

Others seem to share this view. In a 2001 survey of 195 health education staff at US universities, nearly two-thirds of respondents were of the opinion that inclusion of the same section of text in multiple articles was acceptable.³

Unacceptable practices

In a recent editorial, the editorial board of *ACS Nano* describe data recycling in strong terms – 'fraud', no less.⁴ The authors rail against the waste of peer reviewers' time, warn of the loss of reputation and likelihood of getting caught, and lay the blame squarely on pressure on academics to publish.

The *Lancet* editorial identifies deception as the central issue here. The authors of the *ACS Nano* article concur, opining that it 'comes down to the central issue of deception – were the authors trying to deceive the editors, the referees, and the readers [by] presenting recycled data, text, and figures as entirely new material'?

The consequences of deliberate attempts to mislead by recycling data or large amounts of text can be serious – retraction, submission bans, getting grassed up to one's more senior colleagues – and rightly so.

Text recycling

The reuse of one's own words is a far greyer area. When an author replicates descriptions of methods or other text from similar studies, it is perhaps because (s)he does not consider rephrasing to be a worthwhile exercise. Why waste time rewording perfectly written text merely to avoid the charge of self-plagiarism?

Stuart White⁵ wonders as much in a letter of apology to *Anaesthesia*, written after he got into a bit of bother for publishing related (but different) articles with the same title in different journals.



© Anders Holmqvist, 2012

He goes on to bemoan the lack of guidance for authors in his position, and argues that it is up to journal editors to decide what constitutes self-plagiarism.

But couldn't they use some guidelines too?

While the International Committee of Medical Journal Editors (ICMJE) guidelines⁶ touch on specific topics (such as the publication of important medical guidelines in multiple journals in order to reach a wider audience), they do not address all forms of self-plagiarism. A far more useful resource is Miguel Roig's guide to ethical writing⁷, essential reading for anyone concerned about any aspect of plagiarism. Roig defines what he considers to be the major types of self-plagiarism (see Table 1),

Table 1: Forms of self-plagiarism

Data augmentation*	Publication of old data with new supporting data as a new study
Duplicate publication*	Submission of essentially the same article for publication in two different journals
Redundant publication*	Publication of previously published data (with or without new data) with a new angle or focus
Salami slicing	Publication of different results from a study as separate papers when they would best be presented together**
Text recycling	Reuse of published text in a new publication

Adapted from⁷

*Data augmentation, duplicate publication, and redundant publication are all forms of data recycling.

**It is generally assumed that the motivation for this practice is to maximize the number of publications obtained from a single study.

explains why they are a problem, and provides helpful advice on maintaining high ethical standards.

Nonetheless, the apparent lack of official guidelines covering text recycling makes it hard for authors, editors, and readers alike to judge what is acceptable.

Staying out of trouble

Copyright is an obvious practical issue to consider. The authors of an accepted manuscript are often required to sign over copyright to the publisher. Subsequent reproduction of parts of the manuscript may constitute a breach of copyright. (Different publishers have different rules governing the amount of text that may be reused without permission.) A number of journals do, however, allow authors to retain copyright, and the 'fair use' clause – which permits limited reproduction of one's own work for specific purposes – affords some room for manoeuvre.

One way to avoid self-plagiarism in methods sections is to describe the procedures briefly and provide references to previous articles in which they are described in full. However, this is not an entirely satisfactory solution as it risks inconveniencing the reader (who may be forced to refer back to, and perhaps purchase, these previous articles).

Some consider methods to be a special case. *Anesthesia & Analgesia*⁸, for example, permits verbatim copying of descriptions of methods, but not other text, by the original author. Other journals, however, do not. In short, there is no consensus.

The ACS *Nano* article quotes the ethical guidelines of the American Chemical Society,⁹ according to which appropriate citation and use of quotation marks is necessary and sufficient to legitimize text recycling. However, convention dictates that it is not okay to present a whole page of methods in quotation marks. Roig, meanwhile, advocates the '[use] of quotations and proper paraphrasing'.

A declining problem?

Plagiarism in general has never been easier to detect. A range of detection software is now available including eTBLAST, a free tool available from the Virginia Bioinformatics Institute website,¹⁰ and journals and publishers are waking up to the benefits of these new resources.^{11(a)} *Anaesthesia*, for example, now uses proprietary software to analyse every submitted article for evidence of plagiarism.^{12(c)} (Describing this new practice in 2010, Editor-in-Chief Steve Yentis reported that his journal had directly rejected 4% of submitted articles because of

plagiarism in the year since it was introduced, but sadly failed to pinpoint the precise grounds on which decisions to reject are made.)

There are signs that efforts to tackle the duplication of manuscripts may be working. The number of new articles deposited in Déjà vu – an online database of Medline articles that are ‘highly similar’ to other Medline articles¹³ – fell by approximately half in relative terms between 2006 and 2008.¹⁴ Whether this change reflects better detection of self-plagiarism by journal editors or increased wariness on the part of potential offenders is open to speculation.

How much is too much?

Not everyone would be overjoyed if I were to take a paper I had published and create a new one by merely replacing the data for one disease with those for another – as Andrzej Jendryczko more or less did in a notorious case that came to light in 1997.^{15,16(3)} But how much repeat text *is* okay?

In the absence of established guidelines, Drs Richard Kravitz and Mitchell Feldman of the University of California polled a number of experts for their opinions.¹⁷ While many considered 10% an acceptable amount of recycled text, none felt that anything above 30% was reasonable. Similarly, ‘some editors’ have operated on the principle that ‘overlap of more than one-third of the material’ in review articles is too much, according to a World Association of Medical Editors (WAME) report from 2004.¹⁸ Earlier sources quoted recycled text limits of 10¹⁹ and 30%.²⁰

Wherever one draws the line, consideration should perhaps be paid to the background of the author. For a non-native English-speaker who had difficulty describing something first time around, finding a second set of words to describe the very same thing may be an insurmountable challenge.

Conclusion

Both authors and editors would benefit from a clear set of guidelines. The former would know how to avoid trouble; the latter would know when to take action and what action to take.

Acknowledgement

My thanks to Miguel Roig for his helpful comments on an earlier version of this article and for directing me to some very useful information sources.

Notes

(1) In researching this article, I tested whether two free Google search-based plagiarism detection tools – Article Checker (<http://www.articlechecker.com>) and Dupli Checker (<http://www.duplichecker.com>) – could recognize abstracts retrieved from PubMed. While Article Checker struggled to determine the origin of any of the abstracts I threw at it, Dupli Checker spotted signs of plagiarism in most cases, but produced different results when performing identical searches. A third plagiarism checker, available at <http://plagiarisma.net>, was far more effective (flagging almost every sentence of every abstract as unoriginal), but free use is limited to five searches per day.

(2) CrossCheck, available to members of academic publisher organization CrossRef. Users pay a per-document fee and an annual administration charge.

(3) In fact, the paper Jendryczko ripped off wasn’t even his; it had been written 12 years previously by fellow Pole Tatiana Gierek and her colleagues. Remarkably, Gierek herself appears to have borrowed excessively from her own work on occasion (<http://spore.vbi.vt.edu/dejavu/duplicate/67233>).

References

1. Self-plagiarism: unintentional, harmless, or fraud? *Lancet* 2009;374(9691):664.
2. Chalmers I. Intentional self-plagiarism. *Lancet* 2009; 374(9699):1422.
3. Price JH, Dake JA, Islam R. Selected ethical issues in research and publication: perceptions of health education faculty. *Health Educ Behav* 2001;28(1):51–64.
4. Bonnell DA, Hafner JH, Hersam MC, *et al.* Recycling is not always good: the dangers of self-plagiarism. *ACS Nano* 2012;6(1):1–4.
5. White SM. Self-plagiarism. *Anaesthesia* 2011;66(3): 220–1.
6. International Committee of Medical Journal Editors. Uniform requirements for manuscripts submitted to biomedical journals: writing and editing for biomedical publication. Available from: http://www.icmje.org/urm_full.pdf. Accessed 2012 March 13.
7. Roig M. Avoiding plagiarism, self-plagiarism, and other questionable writing practices: a guide to ethical writing. Available from: <http://ori.hhs.gov/sites/default/files/plagiarism.pdf>. Accessed 2012 March 13.
8. 2010 Anesthesia & Analgesia Guide for Authors. Editorial Board, Anesthesia & Analgesia. Available from: <http://www.anesthesia-analgesia.org/content/111/2/525.full#content-block>. Accessed 2012 March 27.
9. Ethical Guidelines. ACS Publications. Available from: <http://pubs.acs.org/userimages/ContentEditor/1218054468605/ethics.pdf>. Accessed 2012 March 13.
10. eTBLAST. Virginia Bioinformatics Institute. Available from: <http://etest.vbi.vt.edu/etblast3>. Accessed 2012 March 13.

11. Butler D. Journals step up plagiarism policing. *Nature* 2010;466(7303):167.
12. Yentis SM. Another kind of ethics: from corrections to retractions. *Anaesthesia* 2010;65(12):1163–6.
13. Déjà vu. Available from: <http://spore.vbi.vt.edu/dejavu>. Accessed 2012 March 13.
14. Garner HR. Combating unethical publications with plagiarism detection services. *Urol Oncol* 2011;29(1):95–9.
15. Marshall E. Medline searches turn up cases of suspected plagiarism. *Science* 1998;279(5350):473–4.
16. <http://spore.vbi.vt.edu/dejavu/duplicate/2977>. Accessed 2012 April 4.
17. Kravitz RL, Feldman MD. From the editors' desk: self-plagiarism and other editorial crimes and misdemeanors. *J Gen Intern Med* 2011;26(1):1.
18. Self-plagiarism of textbook chapters. World Association of Medical Editors. Available from: <http://www.wame.org/ethics-resources/self-plagiarism-of-textbook-chapters>. Accessed 2012 March 14.
19. Iverson C. *AMA manual of style: A guide for authors and editors*. 9th ed. Baltimore: Williams & Wilkins; 1998.
20. Samuelson P. Self-plagiarism or fair use? *Commun ACM* 1994;37(8):21–5.

Author information

Stephen Gilliver studied for a PhD in Cell Biology at the University of Manchester. After working as a postdoc at the same institution and an associate lecturer at Manchester Metropolitan University, he became a freelance copy editor. He is now the science editor at the Center for Primary Health Care Research in Malmö, Sweden.