

Common Misconceptions about Text Recycling in Scientific Writing

CARY MOSKOVITZ, SUSANNE HALL, AND MICHAEL PEMBERTON

Among the fundamental principles of scientific publishing, originality is one of the most important. Every manuscript is expected to offer a unique contribution, something clearly different from what has already been published. We typically think about originality in terms of a paper's content: What does this manuscript add to the knowledge of the field? An article may offer some fundamentally new idea or evidence that substantially alters the field, but more often, advances are incremental.

An example of such incremental advances is a series of articles investigating a new vaccine. Many papers are published before a promising vaccine gets to the stage of clinical trials. Then there will be many more studies: safety studies; pilot studies; studies on different populations, such as adults and children; studies about the efficacy of varying dosages; and so on. If done well, each new paper will offer important insights and inform future research. But from one study to the next, some things will stay the same: the essential problem being studied, the relevant prior research, the biochemistry of the vaccine, the method of vaccine delivery, and so on.

This overlap in content raises a question about writing that has troubled both researchers and editors: When is it acceptable for researchers to repeat material from their own prior papers? Prior to the digitization of written communication, the duplication of material between

scientific articles must often have gone unnoticed. It certainly wasn't a topic of discussion in the realm of research ethics. But with the growth of digital communication came online plagiarism detection tools, and these identify overlapping material regardless of whether the prior work was written by the same authors or someone else. Concern over what came to be called *self-plagiarism* was amplified by the companies selling these tools. They could market the software to publishers to flag papers with overlapping text and then sell the same tool to universities for researchers to use to identify (and therefore change) overlapping material before submitting the manuscript so it wouldn't get flagged (Moskovitz and Colton 2021).

Text recycling in the scientific process

Experienced scientists know there is often a need to repeat some content from their papers, especially when the same methodological approach, experimental apparatus, or statistical analyses are used in related studies. Researchers may also have reason to repeat some background material, such as discussion of prior research or theoretical frameworks.

Reusing material from one's published article in a new article is one kind of *text recycling*. Others include reusing material from a published article in one's dissertation, reworking a conference paper into a journal

article, and translating one's work into a different language. Some kinds of recycling are widely considered appropriate and even standard practice. Others, especially manipulating text to disguise a previously published paper and submitting it as a new work, are universally condemned and may lead to disciplinary action.

Given the wide variety of ways that scientists might recycle text, it isn't surprising that they are often unsure about what is and isn't appropriate. The situation has been aggravated by the circulation of incorrect or partially correct ideas about proper practice through online discussion boards, word of mouth, and even well-intentioned editorials.

Misconceptions about text recycling

We started the Text Recycling Research Project (TRRP) in 2017 with a grant from the US National Science Foundation to study text recycling and develop materials to help the scientific community better understand the practice and establish consensus where possible. In this Viewpoint, we discuss some of the most common misunderstandings that we've come across during our research on this topic.

Misconception 1: Text recycling is widely considered unethical. Some scientists have been told that text recycling is fundamentally unethical and should always be avoided. It is certainly true that some uses of text recycling are

unethical. But most editors and editorial board members from top journals believe that some uses of text recycling have a place in research writing, even if they hold differing opinions about the limits of the practice (Hall et al. 2018, Pemberton et al. 2019). Our research shows that the acceptability of text recycling depends on factors such as the quantity and rhetorical nature of the material (e.g., describing equipment versus presenting results), whether the source is published, and whether the authors have been transparent with the editors throughout the submission and publication process.

Misconception 2: Text recycling is a form of plagiarism. Reusing material from one's own prior documents is often called *self-plagiarism*. Plagiarism is never ethical, so if text recycling were indeed a kind of plagiarism, it would always be unethical too. But text recycling is not plagiarism. Text recycling doesn't appropriate the intellectual property of others or deprive them of credit for their work. And unlike plagiarism, text recycling is often acceptable and sometimes even desirable. That's why organizations like COPE (Committee on Publication Ethics) and the TRRP prefer the neutral term *text recycling* instead.

Misconception 3: Respectable scientists never recycle text. Given the controversy surrounding text recycling over the past decade, one might think it is uncommon. But text recycling is not unusual in scientific writing. Scientists routinely recycle from conference posters when writing articles, and they frequently recycle large parts of conference papers into articles as well—often at the invitation of their professional societies. Recycling is even common—although in much smaller amounts—from one published research paper to another. In fact, most recent papers published under National Science Foundation grants in the United States contain some recycled text (Anson and Moskovitz 2021). Of course, being common doesn't make something right, but it does

suggest that some uses of text recycling are generally accepted within the scientific community.

Misconception 4: Text recycling infringes copyright. Many people believe that recycling text from published papers should always be avoided because it infringes copyright and could therefore lead to lawsuits. The legal status of recycling text in research contexts is rather complicated. Under US law, the most common instances of text recycling in scientific writing are allowed as a *fair use* (Hansen and Moskovitz 2021). And although the laws of other countries vary, limited amounts of recycling from one published paper to another are not likely to be legally problematic—especially because the practice is so common that any publisher considering legal action would likely have published recycled material as well. If authors or editors are concerned that the amount or kind of material being recycled could be legally problematic, they can always obtain permission from the rights holder—just as they would when using an image or extended passage from another author.

Misconception 5: Authors can recycle their own previous writing without limits. Some authors mistakenly believe the opposite of misconception 4—that, as the author, they have the legal right to reuse anything they themselves wrote. That may be true for most unpublished material (although not if the work was originally produced as *work for hire*), but for published work, authors may surrender their rights of reuse when they sign a publishing contract. And just because an author may have the *legal* right of reuse, that doesn't mean it's *ethical*—especially if the author is trying to deceive editors or readers about the originality of the “new” work.

Misconception 6: Authors can avoid concerns about text recycling by rewording the material or putting it in quotes. Researchers are sometimes told that they can sidestep any ethical or legal concerns about recycling merely by rewording

the material. However, this is not inherently more ethical, especially if the aim of rewording is to deceive editors or readers about the originality of the work (Anson et al. 2020). Similarly, superficial changes in wording such as substituting synonyms or rearranging clauses don't eliminate concern about copyright infringement. If the revised version is “substantially similar” to the prior version, it would legally be considered the same. There are practical downsides as well: Superficial edits make it more difficult for editors to determine whether any material has been reused, even if they are using plagiarism detection software, and can also make it harder for readers to tell exactly how a new study compares with the authors' prior studies.

Some researchers have been told that there is one single set of correct writing practices for all types of text reuse, no matter the context. According to this belief, authors must put *all* reused material in quotation marks—even single sentences or phrases—just as they would when reusing the words of others. In reality, there are no universal ethics of writing practices; the context always matters. And those who advocate for quoting rather than recycling don't realize that quotation marks also draw readers' attention to the specific words used. For most instances of text recycling in research papers, readers would find quotation marks distracting. Furthermore, scientists rarely recycle entire blocks of text verbatim; instead, they reuse parts of paragraphs or sentences, editing the material as needed to make it appropriate for the new study. The resulting passages—a mix of recycled and new material—would require multiple sets of quotes, which would not only confuse readers but be unacceptable to most science editors.

Better terminology leads to better policy

Some of the confusion around text recycling has resulted from problematic terminology. Publishers, professional organizations, and scholarly societies have frequently used many of the same terms—*self-plagiarism*,

text recycling, redundant publication, duplicate publication—to mean different things, sometimes even within the same organization (Moskovitz 2021). One of the TRRP's initial priorities was to develop a set of terms to clarify the ethics and legalities for different types of text recycling, as follows:

Developmental recycling. Recycling material from unpublished documents produced as part of the research and writing process (e.g., reusing material from a conference poster in a journal article).

Generative recycling. Recycling a limited amount of published material in a new work that offers a substantive and original intellectual contribution (e.g., reusing the description of a sampling procedure from one's published article in a new article).

Adaptive publication. Recycling the entirety or core content of a published work—but for a different readership, genre or context (e.g., translating one's published article into another language).

Duplicate publication. Recycling the entirety or core content of a published work for the same audience and genre, which is nearly always unacceptable.

This taxonomy can reduce much of the confusion around text recycling, because naming the different practices allows us to more carefully describe the ethical norms and legal issues that apply to each. That is why we explain the terms and use them in our own educational documents and our new model policy.

Text recycling isn't always okay

Recycling large amounts of material or reusing the central parts of a published article might be unethical or illegal.

Duplicate publication—deceiving editors and readers by reworking prior publications to make them appear to be new works—is especially unethical and usually illegal. Authors who are uncertain as to whether their particular use of text recycling is acceptable should consult with the editors.

And let us be especially clear about this: We are *not* encouraging plagiarism—the reuse of *other people's* work without proper attribution. Text recycling is the reuse of material by the same author (or team). The ethical norms for recycling are fundamentally different from those for plagiarism, which is always unethical and widely considered academic misconduct at research institutions.

How to learn more

The TRRP has produced a number of research-based resources for the scientific community which are freely available on our website (textrecycling.org). These have been thoroughly vetted by our advisory board, which includes leaders from COPE and the Council of Science Editors, editors from leading science publishers, and research integrity officers from government and universities. Researchers will find our brief *TRRP Best Practices for Researchers* and the more in-depth *Understanding Text Recycling: A Guide for Researchers*. Editors will find our *Understanding Text Recycling in Research Writing: A Guide for Editors*, the *TRRP Model Text Recycling Policy*, and the *TRRP Guide to Developing Text Recycling Policies*.

Acknowledgments

This article was produced with support from the US National Science Foundation Cultivating Cultures for Ethical STEM program, award no. 1737093.

Project website

Readers can access all TRRP resources and full-text versions of our published papers at textrecycling.org

References cited

- Anson I, Moskovitz C. 2021. Text recycling in STEM: A text-analytic study of recent NSF-sponsored research reports. *Accountability in Research* 28: 349–371.
- Anson C, Hall S, Pemberton M, Moskovitz C. 2020. Reuse in STEM research writing: Rhetorical and practical considerations and challenges. *AILA Review* 33: 120–135.
- Hall S, Moskovitz C, Pemberton M. 2018. Attitudes toward text recycling in academic writing across disciplines. *Accountability in Research* 25: 142–169.
- Hansen D, Moskovitz C. 2021. Text Recycling in Research Writing: U.S. Copyright Law and Fair Use. Text Recycling Research Project. <https://textrecycling.org/files/2021/05/TRRP-White-Paper-Text-Recycling-in-Research-Writing-U.S.-Copyright-Law-and-Fair-Use.pdf>
- Moskovitz C. 2021. Standardizing terminology for text recycling in research writing. *Learned Publishing* 34: 370–378.
- Moskovitz C, Colton A. 2021. Self-plagiarism, fraud, and iThenticate: A complicated relationship. *Inside Higher Ed* (5 March 2021). www.insidehighered.com/views/2021/03/05/resources-avoiding-self-plagiarism-are-scarce-and-problematic-opinion
- Pemberton M, Hall S, Moskovitz C, Anson C. 2019. Journal editors' views on text recycling: An interview-based study. *Learned Publishing* 32: 355–366.

Cary Moskovitz is affiliated with the Thompson Writing Program, at Duke University, in Durham, North Carolina, in the United States. Susanne Hall is affiliated with the Division of Humanities and Social Sciences at the California Institute of Technology, in Pasadena, California, in the United States. Michael Pemberton is affiliated with the Department of Writing and Linguistics at Georgia Southern University, in Statesboro, Georgia, in the United States.

<https://doi.org/10.1093/biosci/biac090>