Guide to Using and Citing Sources When Writing in the Life Sciences

by Tamily Weissman, Tom Torello, Natalie Nannas, and John Dowling

In your scientific writing it is important to provide credit to previous scientists, whether you are putting together a senior thesis, term paper, or lab report. It is thus essential that you clearly understand how to incorporate sources correctly. Harvard College has a strict policy on plagiarism, which includes inappropriate paraphrasing, and students sometimes have difficulty applying the rules they learn in Expository Writing to the sciences. While writing in other disciplines is based largely on ideas, theories, and original wording, writing in the sciences often requires you to virtually restate specific results from previous research studies without paraphrasing too closely. An important part of your undergraduate training is learning when and how to appropriately summarize and cite previous data in your scientific writing. Read this document carefully and refer to it when you are writing any life sciences document.

1) When and why is a citation necessary in my scientific writing?

It is important to acknowledge any work done by previous authors in your writing. You should thus include a citation for any statement of a fact or theory not considered to be common knowledge. For a more detailed description of what constitutes common knowledge, see Section 1 below. When in doubt, it is always safe to support your statement by citing a recent review article or textbook chapter.

2) How do I summarize scientific text without paraphrasing too closely?

Your description of someone else's findings or ideas should be put into your own original words. It is often necessary, however, to paraphrase or even restate partial phrases, especially when summarizing detailed data or technical terminology. You should repeat only the crucial parts of any original text (typically key scientific phrases that are commonly used in the field) while completely restructuring and rewriting the other parts of the sentence. For specific examples, see Section 2 below.

3) When do I cite a review versus the primary literature?

As a rule, if you describe data from a primary source (that you learned about in a review), you should **read** the primary article and **cite** the primary article. If however, you depend heavily upon the review, either for argument or overall structure, **you must also cite the review**. For more detail, see Section 3.

4) When should I use quotation marks?

Quotes are not typically appropriate in scientific writing, as they are in other disciplines, and can often indicate a lack of understanding. However, when in doubt, **it is always *safe* to use quotation marks with citations**. See Section 4 below for more information about the use of quotes.

5) Which publication should I cite?

When several authors have published a similar finding over the years, it is most appropriate in the life sciences to cite the *earliest* publication. This gives credit to the scientists who first made the discovery.

6) What about citing information from the web?

"Primary" online sources are **acceptable** as references (e.g.. tools that exist solely on the web and not elsewhere, such as sequence databases, interactive anatomical maps, etc.).
"Secondary" online sources are **not acceptable** as references (e.g. newspaper articles, online encyclopedias, pop science reviews, etc.) unless they report information never put forth elsewhere.

7) What about plagiarism in problem sets?

Unless it is already stated clearly in the syllabus, ask your TF or professor whether it is acceptable to work in groups. If so, then ordinarily you must still turn in an assignment that is written in your own words – not one that is identical to a study partner's. You must also acknowledge any collaboration: simply write a note on your assignment listing any study partners' names.

More Information on Using and Citing Sources in the Life Sciences

The *Harvard Guide to Using Sources* considers plagiarism to "draw any idea or any language from someone else without adequately crediting that source". This includes copying original text or even paraphrasing original text too closely with an appropriate citation. Since writing in the sciences often focuses on specific data from previously published research studies, it is particularly important to know how to summarize and reference the appropriate literature. Here we expand upon topics from on page 1.

Section 1. What constitutes common knowledge?

In your scientific writing, you should include a citation for any statement of a fact or theory not considered to be common knowledge. For example, it is considered common knowledge that mitosis is one stage of the cell cycle, and thus this fact need not be cited. But where does one draw the line on what is considered common knowledge? Although there is no strict rule, common knowledge can be considered to be: 1) facts known by the general public; 2) facts well known in vastly different fields; or 3) basic facts learned in high school. When in doubt, it is always safe to cite a recent review article or textbook chapter, especially if the information is particularly cogent to your argument. In some cases you may be assisting the reader if s/he is interested in further literature on the topic.

Section 2. Examples of appropriate paraphrasing

The challenge when paraphrasing is to retain a statement's original meaning while exhibiting clear mastery of the detail by using your own wording and overall structure. See the following examples.

Example One, original text: "The mutation is predicted to result in an arginine-to-histidine substitution (R553H) in the forkhead DNA-binding domain of FOXP2." (Lai et al., 2001)

Quality of paraphrasing	Rewritten text	How exactly was the text rewritten?
Inappropriate	The mutation is predicted to result in an arginine-to-histidine substitution (R553H) in the forkhead DNA-binding domain of FOXP2 (Lai et al., 2001).	Copied exact words from original text. Cited source.
Inappropriate	The mutation may result in an arginine-to-histidine substitution (R553H) in the FOXP2 forkhead DNA-binding region (Lai et al., 2001).	Used the same words or phrases with a few replacement synonyms. Cited source.
Acceptable	Lai et al. (2001) proposed that the FOXP2 mutation is due to an arginine-to-histidine substitution (R553H) in its forkhead DNA-binding region.	Changed words, but used the same structure of sentence or paragraph. Cited source.
Excellent	Point mutations can often affect overall protein function, for example one arginine-to-histidine mutation in the FOXP2 gene (Lai et al., 2001) is thought to block its ability to bind DNA.	Restructured overall format and cited source. Concept completely rewritten, perhaps synthesized with other points and sources. Retains and often adds to original meaning.

Example Two, original text: "In recent years, innovations in three areas have contributed to the success in localizing genes for dyslexia and other language-related traits. These are quantitative trait locus (QTL) mapping, phenotypic dissection and high throughput genome-wide scanning." (Fisher & DeFries, 2002)

Appropriate	Quantitative trait locus (QTL) mapping, high throughput genome-wide scanning, and phenotypic dissection have been crucial to the identification of dyslexia and additional language-related genes (Fisher & DeFries, 2002).	Sentence re-written and restructured but maintains key scientific phrases. Source cited.

Paraphrasing TIP: First, read, digest, and fully understand the data. Then wait a few minutes and begin a fresh sentence, explaining the findings in your own words without referring to or rearranging the original text. Students get into trouble when they attempt to paraphrase a sentence while looking at the original – it is difficult to be creative and write something in your own words if you are using the original text as a template. If you simply cannot restate a finding without referring to the original, you may not completely understand it.

Paraphrasing TIP: To avoid inadvertent plagiarism, be careful with note taking. If you type notes directly into the computer when reading articles, highlight any verbatim text in a distinctive font so that you can later recognize and rewrite those sections. Even better, paraphrase/rewrite immediately as you are taking notes, but always make sure you don't change the meaning.

Section 3. When do I cite a review versus the primary literature?

One stumbling block for life sciences students is understanding how review articles should be cited. Review articles are excellent resources early in your literature search as you identify primary research articles you may wish to read. *As a rule, if you describe data from a primary source (that you learned about in a review), you should **read** the primary article and **cite** the primary article.* If however, you depend heavily upon the review, either for argument or overall structure, *you must also cite the review article*. This is one of the most common mistakes made by life sciences students. *Students who rely upon a review article and do not cite it are sometimes accused of trying to hide the review as a source*. One option is to cite both the primary article/s as well as the review, for example: "(Cardon et al., 1994, Lai et al., 2001, as reviewed in Fisher & DeFries, 2002)." If you cannot locate a referenced original paper through the Harvard library system but would like to include its data in your paper, you may cite it as follows: "(Cardon et al., 1994, as cited in Fisher & DeFries, 2002)".

Section 4. When should I use quotation marks?

As stated in A Student's Guide to Writing in the Life Sciences, quotes are not typically appropriate in scientific writing, as they are in other disciplines. In some fields, guidelines suggest putting phrases of three words or more into quotes, but in the sciences it would become tedious if terms like "central nervous system" were placed into quotes throughout a document. While it is not appropriate to place quotes around most scientific phrases, occasionally you may wish to use them, in particular if you repeat someone's spoken words, or if you cannot find a way to restate the data. If you do not fully understand data you are describing, and thus have difficulty rewriting it in your own words, it may be safest to either use quotes or leave the data out. *In the worst-case, last-minute scenario, if you are late for a deadline and do not have time to rephrase the original text, use quotes and cite the source. Never assume that your word-for-word similarity will go unnoticed.* Using quotes may hurt your grade, but at least it is legal. Many students have been required to withdraw from the College after just one plagiarism mistake.

Additional Resources

If you would like further advice about how to incorporate citations into your scientific writing, ***ask for help***. Your professors, TFs, and advisors are there to help you navigate this confusing topic as you learn to become a good scientific writer. These guides may also help:

A Student's Guide to Writing in the Life Sciences. This document provides helpful tips for how to write in the life sciences including how to find and incorporate references. Online pdf (linked here); printed copies in the Life Sciences office.

Harvard Guide to Using Sources. This document provides a thorough overview of how to cite all sources of information. Particularly relevant information can be found in the "Avoiding Plagiarism" section.

References

Cardon LR, Smith SD, Fulker DW, Kimberling WJ, Pennington BF, DeFries JC (1994). Quantitative trait locus for reading disability on chromosome 6. Science, 266: 276-279.

Fisher SE & DeFries JC (2002). Developmental dyslexia: Genetic dissection of a complex cognitive trait. Nature Reviews Neuroscience, 3: 767-780.

Lai CSL, Fisher SE, Hurst JA, Vargha-Khadem F, & Monaco AP (2001). A forkhead-domain gene is mutated in a severe speech and language disorder. Nature, 413: 519-523.