Redefining 'plagiarism'

Roy Jensen, M.Sc., Ph.D.

On first glance, plagiarism is a simple concept with the goal of ensuring that credit is given when information is taken from other sources. Indeed, plagiarism is often defined simplistically and absolutely:

> **plagiarism** 1. the passing off of the ideas or words of another as one's own work
> 2. to use another's work without crediting the source
>
> (adapted from the Merriam-Webster online dictionary)

This simplistic definition is applied in contexts where the expectation and established practice is to cite the source of all third-party information. Persons preparing scholarly articles, news reports, and students learning these skills are subject to this definition. However, the question of what needs to be cited is often unclear. Additionally, this simplistic definition fails to consider the numerous contexts where different citation expectations have been established:

- in preparing policies (and laws), existing policies from other organizations are routinely reviewed. Since the intent is usually to establish a similar policy, sections are often copied verbatim or with minor modification, but rarely are these sources cited.
- executives and politicians routinely sign their name to letters prepared by others.
- in legal decisions, judges often use the arguments of the prosecution and defence in their decision without citation.
- journalism is changing to a model where articles no longer directly reference the source of material. An increasing number of articles now contain the statement, "with files from ___"
- celebrities and publishers hire writers to write books — autobiographies and novels — and the writer is not credited for their work (ghostwriting)
- academic instructors routinely create assignments and exams using questions taken from textbooks. The questions are often used with little or no modification, but are rarely cited.
- academic textbooks often present real-world examples to illustrate the importance and applicability of the material, but rarely are the sources cited.

The simplistic definition would find that all of these actions constitute plagiarism, yet they occur commonly. Clearly, a better understanding of plagiarism and a more robust definition of plagiarism is required to confidently and accurately understand plagiarism in different contexts.

This article explores the legal foundation of plagiarism and then identifies several factors that must be considering when determining if plagiarism has occurred. In conclusion, a refined definition of plagiarism is proposed that addresses the limitations of the simplistic definition.

* This document is based on my exploration into copyright, copyright infringement, and plagiarism. I am not a lawyer, just an interested academic. I can be contacted at r.jensen@consol.ca

Legal foundations of plagiarism

Plagiarism is not defined in law. In practice, it is seen as a breech of moral and ethical codes of conduct. In law, plagiarism derives its existence primarily from copyright law. *Copyright* is the legally enforceable right of a person — the creator — to claim ownership of a literary, dramatic, musical, or artistic *work*. The creator has the sole legal authority to determine how, where, and by whom a work may be used. This authority is subject to the *fair use* provisions within copyright legislation. *Copyright infringement* occurs when someone violates the creator's copyright. *Plagiarism* is a form of literary copyright infringement.

A second legal basis for plagiarism is *fraud*. Wikipedia defines *fraud* as "an intentional deception made for personal gain or to damage another individual." The definition outlines the key features of fraud: there must be false information or a false action, the action must be intentional, there must be intent to deceive, and there must be either personal gain or damage to another individual. In this context, *plagiarism* occurs when a person knowingly misrepresents someone else's literary work as their own for personal gain.

A third legal basis for plagiarism is the law of *passing off*, which is a component of unfair competition. The Duhaime Legal Dictionary defines *passing off* as making a false representation likely to induce a person to believe that the goods or services are those of another. Underlying passing off is intent to deceive the consumer. Passing off requires there to be an established connection between the product and the manufacturer that is commonly known by the consumer, and there must be a high probability of confusion when a third party replicates the product. In the context of plagiarism, passing off is the reverse of fraud: producing a literary work that looks like that of a recognized author so that consumers will purchase it. (Ghostwriting is an example of passing off.)

US lawyer Jon Siegel is a law professor at George Washington University specializing in intellectual property law. In an article entitled *Plagiarism's Defenses*, Siegel argues that the law distinguishes between negligence and intentional wrongdoing. For example, in property law, it isn't theft if you didn't intend to take someone else's property.

> *Imagine being in a store and absentmindedly put an object in your pocket while talking with a friend or while trying to console your child. Security stops you outside the store. It's embarrassing and humbling, but it isn't theft since it wasn't your intention to steal.*

Siegel argues that the same rationale should apply to plagiarism. While we strive to be perfect, we do make mistakes. If it can be shown or suggested that a person was negligent in failing to cite a source, it's embarrassing and humbling, but it isn't plagiarism because it wasn't intentional.

The *Plews* v. *Pausch* case (2006 ABQB 607) specifically addressed academic plagiarism. Paragraph 50 of the decision proposes a test to determine if a person (the defendant) is guilty of plagiarism. The test is adapted here for general use.

Once all the evidence has been submitted,

1. If the tribunal believes the defendant, then the defendant is not guilty.
2. If the tribunal does not believe the defendant, but if the defendant has raised sufficient questions about the complainants' evidence so that the credibility of the evidence is in doubt, then the defendant is not guilty.
3. If the tribunal does not believe the defendant, the complainant must still satisfy the tribunal that their version of events is more probable, or the defendant is not guilty.
4. Only if the tribunal believes the complainant and is satisfied that the complainant's version of the events is most probable is the defendant guilty.

This test clearly puts the burden of proof on the complainant to convincingly prove the defendant knowingly plagiarized, which is consistent with established legal principles and practices: innocent until proven guilty.

Given these legal foundations, it is clear that the simple, absolute, definition of plagiarism on page 1 is inconsistent with established legal principles and practices.

Considerations

Plagiarism is considered by most to be an egregious violation of moral and ethical codes of conduct. Either because of or despite this belief, most organizations give little consideration to what constitutes plagiarism and do not state much more than the common, absolute, definition. Yet the examples presented above provide numerous contexts where this definition is inadequate within those same organizations. The Wikipedia entry on plagiarism is probably most accurate in stating that the concept of plagiarism is "problematic with nebulous boundaries".

The American Historical Association (AHA) recognizes the nebulous nature of plagiarism in their Statement on Standards of Professional Conduct.

> [Section 4] The expropriation of another author's work, and the presentation of it as one's own, constitutes plagiarism and is a serious violation of the ethics of scholarship.
>
> …
>
> Of course, historical knowledge is cumulative, and thus *in some contexts — such as textbooks, encyclopedia articles***,** broad syntheses, and certain forms of public presentation — the form of attribution, and the permissible extent of dependence on prior scholarship, citation, and other forms of attribution will differ from what is expected in more limited monographs. *As knowledge is disseminated to a wide public, it loses some of its personal reference.* What belongs to whom becomes less distinct. *But even in textbooks a historian should acknowledge the sources of recent or distinctive findings and interpretations*, those not yet a part of the common understanding of the profession. (emphasis added)

The AHA document contains several important points:

- citation expectations differ depending on the type of document being written
- work in collective documents should be distinctive to require citation
- as information becomes increasingly disseminated, it loses its distinctiveness and the requirement for citing the work decreases

In attempting to frame plagiarism, I identified five factors that must be considered when determining if plagiarism has occurred:

- common knowledge
- originality
- non-copyrightable information
- established practices
- intent

Common knowledge

The Wikipedia entry on *common knowledge* states,

> **common knowledge** is knowledge that is known by everyone or nearly everyone, usually with reference to the community in which the term is used. … Often, common knowledge doesn't need to be cited.

Common knowledge depends on the community. The breadth of knowledge possessed by grade school students is limited. Thus, grade school students are expected to cite information that more senior students are expected to know intuitively. With continuing education and increasing breadth of knowledge, the information that is expected to be cited changes. This changing expectation exemplifies the nebulous boundaries of plagiarism. Additionally, common knowledge in one community (biology, for example) is not likely common knowledge in another community (physics, nursing, etc.). This latter observation makes it difficult to bring in people from other disciplines to determine if plagiarism has occurred — they do not have an appropriate reference. Even within a community, when a concept becomes common knowledge is nebulous, so there is a transition period when some sources cite the information and others do not.

For example, every concept now in textbooks without citation was, at one time, at the cutting edge of research and requiring citation when referenced by others. For example, Einstein's famous equation, $E = m\,c^2$, was originally published in *Annalen der Physik*, **1905**, *18*, 639. At the time, Einstein's paper would have been cited because of the new information he discovered. With increasing verification and promotion, $E = m\,c^2$ has become common knowledge. $E = m\,c^2$ is presented in most first-year physics and chemistry textbooks without citing the 1905 paper.

For example, consider the statement, *The sun is a blackbody radiator.* If you (the reader) have a university chemistry or physics background, you will probably know this as common knowledge. If you have a different background, you may not know the scientific definition of *blackbody*, *radiator*, or know that the sun is a blackbody radiator — this is not common knowledge to you and you would be expected to cite this information.

As previously noted, the AHA document explicitly states that collective works, such as textbooks and encyclopaedias, have different citation expectations from scholarly works. No one would reasonably believe that the author(s) of a textbook or encyclopaedia had themselves done all the fundamental research presented in the textbook or encyclopaedia. The material in these collective works is common knowledge within the specific discipline, with the exception of recent or distinctive information that is not yet common knowledge. However, as documented in the ***Established practices*** section below, even these are not commonly cited in textbooks.

Originality

In the context of plagiarism, two definitions of *original* must be considered:

> **original** 1. created for the first time
> 2. presentation in a unique, creative, or distinctive manner

Pertaining to the first definition, copyright protects the *original* work. Only the creator can claim copyright infringement. As information becomes increasingly disseminated, originality is lost and the information becomes common knowledge.

Pertaining to the second definition, there must be something unique, creative, or distinctive about the work to warrant protection by copyright. The AHA document uses the term 'recent or distinctive'.

Consider the following quotes from *Hamlet*, by William Shakespeare.

> I am very glad to see you.
>
> To be, or not to be: that is the question.

There is nothing original about the first quote. This phrase is part of regular speech and does not require citation. The second quote is original, by both definitions. While *Hamlet* was written over 400 years ago, most academics would agree that Shakespeare should be credited when the second statement is used in scholarly articles. However, if this sentence becomes part of common language, it would no longer require citation, and there would be a transitionary period where some would cite the quote and others not. Additionally with literary works, as more of a work is reproduced, it becomes increasingly creative and distinctive, thereby requiring citation.

The CCH Canadian v. Law Society of Upper Canada decision considered the concept of *original*. From the Wikipedia entry on the Canadian Copyright Act,

> The CCH Canadian case … found that for a work to be original it must be the result of the exercise of "skill and judgment". More specifically: skill, meaning the "use of one's knowledge, developed aptitude or practiced ability in producing work", and judgment, meaning the "use of one's capacity for discernment or ability to form an opinion or evaluation by comparing different possible options in producing the work".

If can be shown that negligible skill and judgement were required to create the work, then it is *not* copyright infringement nor plagiarism if another author presents the same or similar information without citation. Assessment of 'negligible skill and judgement' could include considering if the presentation is unique, creative, or distinctive and considering how an equally qualified person

would communicate the information. Generally, when information is presented in clear, concise language — as is common in textbooks and encyclopaedias — that presentation is probably not unique, creative, or distinctive.

Non-copyrightable information

The Feist v. Rural decision established that data — facts, discoveries, and processes — is not protected by copyright. This case centered on the entries in a telephone book, but is generally applicable to all facts, discoveries, and processes. In Canada, the Tele-Direct Inc. v. American Business Information Inc. (1997) 76 C.P.R. (3d) 296 (F.C.A.) reached a similar decision to the Feist ruling.

From the Wikipedia entry on Feist v. Rural,

> For example, a recipe is a process and not copyrightable, but the words used to describe it are…. Therefore, you can rewrite a recipe in your own words and publish it without infringing copyrights. [Even if] you rewrote every recipe from a particular cookbook, … the decisions suggest that it is unlikely [you have infringed the original author's copyright] unless there is some significant creativity carried over from the original presentation.

When considering plagiarism, the statement applying the decisions to cookbooks is critical: if a person can reproduce an entire cookbook, keeping the same sections, same recipe order, same ingredient quantities, only rewriting the instructions in their own words, and not infringe copyright, then the second cookbook cannot be plagiarized from the first. Textbooks are like cookbooks. Within textbooks on the same topic, you also will find the same concepts presented in similar orders, similar styles, and using the same or similar examples and similar visual representations. Textbook questions are often very similar and occasionally identical between textbooks because the same concepts are being taught. A person outside the discipline may cry plagiarism without realizing that the information is presented in this manner because of pedagogy and established practices. Very likely, the first textbook is not *original* by either definition: the example, image, or question is not original to the textbook author nor is it presented in a unique, creative, or distinctive manner.

Unless there is significant originality (definition 2) in how data are presented, it is not plagiarism if they are reproduced from a source document without citation. (Citation does, however, add credibility to the information that is presented.)

Established practices

The established practice for persons preparing scholarly articles is to cite information taken from other sources. The citation expectations in other contexts differ from this rigorous requirement. Consider the examples given at the beginning of this article:

- Policies on a given subject — harassment, plagiarism, safety — contain similar language in a similar fashion. The administrators preparing policies routinely review existing policies

from other organizations. Since the intent is to establish similar guidelines, sections are often copied verbatim or with minor modification, but rarely are these sources cited.

- It is an established practice for administrative assistants, public relations personnel, and human resources personnel to write letters that are signed by executives and politicians. Never is the true author acknowledged in the normal course of business.
- The common practice of attorneys is to prepare their arguments in such a way that the judge can readily use their arguments in the judges decision. Regularly, the judges' decision is composed of arguments written by attorneys, used with little or no modification and used without citation.
- Ghostwriting is common in publishing houses, where writers write novels that are published under the name of a popular author. Since this author has a public following, the book can be sold at a premium price and the sales volume are higher — all in the name of profit. Celebrities hire writers to write their memoirs and autobiographies. Ghostwriting also occurs in academia: ghostwriting of student papers, of scholarly articles, or putting another scientists' name on a scholarly article is considered misconduct; ghostwriting of letters and speeches signed and given by administrators is acceptable.
- Academic instructors routinely create assignments and exams using questions taken from textbooks. The questions are often used with little or no modification, but are rarely cited. Two primary reasons for this are to build student confidence and to test if students actually completed the homework by providing them with examples similar or identical to those done in the textbook.
- Textbooks present information in a manner conducive to learning, which routinely leads to the same information being presented in a similar manner with similar examples. The use of similar (often very similar) text, figures, and examples occurs to convey the same information in a straightforward, logical manner that is conducive to learning. This is critical to establishing a foundation for future learning on a given subject. For example, in chemistry, acid-base equilibria is almost universally taught using acetic acid as the first example; benzene is the first example of aromaticity; in physics, projectile motion is presented as a ball being throw or shot, and there are many other examples in all scientific disciplines.

All of these are *established practices* in their respective areas.

The *Plews* v. *Pausch* decision (para. 51) states that "the most reliable basis for credibility findings is in independent, contemporary materials." Identifying the established practices in a given context is important to determining if plagiarism has occurred.

For example: textbooks contain numerous real-world examples to illustrate the importance and applicability of the material. Rarely are the sources cited. Given that this information is usually recent or distinctive, it should be cited, but the established practice is to not cite this information. To illustrate the nebulous nature of what constitutes plagiarism, I reviewed one chapter from the 6th edition of a first-year chemistry textbook and identified over twenty sections that would require one or more citations if presented in another context. Had the information been presented by the same author in a scholarly article or by a student using the information in a report, citation would have been required. Yet in this and the majority of science textbooks, the information goes uncited.

Intent

The fraud and passing off arguments in the ***Legal foundations of plagiarism*** section require there to be an intent to deceive. *Intent* points to the underlying motives and goals of the author, and is a critical consideration when determining if plagiarism occurred. Specifically, Siegel argues that the law distinguishes between negligence and intentional wrongdoing. Siegel identifies several factors that hint to the authors' intent: the amount of information copied, the nature of the information copied (is it common knowledge? original? distinctive?), the past actions of the writer (plagiarists are often serial offenders), and the mental state of the writer.

Alternate intentions were presented in the previous section to explain the established practices that seemingly constitute plagiarism.

Conclusion

Plagiarism, like any crime, isn't black and white. Plagiarism is a nebulous concept with numerous factors that must be considered when determining if plagiarism has occurred: common knowledge, originality, non-copyrightable information, established practices, and intent. Some of the complications and considerations include

- the lack of a reasonable definition of plagiarism.
- that the common definition of plagiarism is inconsistent with established legal principles and established practices in many contexts.
- that reviewers may not understand what constitutes plagiarism in a given context.
- that reviewers are often from different disciplines, so they do not have the same foundational knowledge of the discipline.
- that citation requirements are dependent on the context where the information is presented.
- that citation requirements change as information becomes common knowledge and the transition period to common knowledge is often long.
- determining the intent of the author.

When determining if plagiarism has occurred, a test in the *Plews* v. *Pausch* case puts the burden of proof on the complainant to convincingly prove the defendant knowingly plagiarized.

Once all the evidence has been submitted,

1. If the tribunal believes the defendant, then the defendant is not guilty.
2. If the tribunal does not believe the defendant, but if the defendant has raised sufficient questions about the complainants' evidence so that the credibility of the evidence is in doubt, then the defendant is not guilty.
3. If the tribunal does not believe the defendant, the complainant must still satisfy the tribunal that their version of events is more probable, or the defendant is not guilty.
4. Only if the tribunal believes the complainant and is satisfied that the complainant's version of the events is most probable is the defendant guilty.

The simplistic, absolute definition of plagiarism must be replaced with a definition(s) that recognizes that complex, nebulous nature of plagiarism. The following is proposed:

> **plagiarism** to knowingly pass off the original and distinctive ideas or words of another author as ones own work without crediting the author and when the context of such use expects the work to be credited

One final example of the nebulous nature of plagiarism: the definition of plagiarism used by the AHA in their Statement on Standards of Professional Conduct is almost identical to what is printed in *Histology News Network*, 2007, para 1 (citation found on page 3 of the document *How To Avoid Plagiarism: The Scourge Of The Academe*, by the Knowledge Growth Support company). Does this constitute plagiarism by the AHA or *Histology News Network*, or is this an example of acceptable use for the reasons given above?

Acknowledgements

I thank Travis Huckell, B.A., LL.B., for his suggestions to improve this document and for directing me to the *Plews* v. *Pausch* decision.

Works cited

American Historical Society *Statement on Standards of Professional Conduct* www.historians.org/PUBS/Free/ProfessionalStandards.cfm (17 November 2011)

CanadaLegal.info *Intellectual Property Law* www.canadalegal.info/ref-intellectual-property/ip-it-law.html (17 November 2011)

CCH Canadian v. Law Society of Upper Canada, [2004] 1 S.C.R. 339, 2004 SCC 13.

Duhaime Legal Dictionary *Passing Off* www.duhaime.org/LegalDictionary/P/PassingOff.aspx (20 February 2012)

Feist Publications, Inc., v. Rural Telephone Service Co., 499 U.S. 340 (1991).

KGSupport.com *How To Avoid Plagiarism: The Scourge Of The Academe* www.kgsupport.com/How_to_Avoid Plagiarism: The Scourge of the Academe.pdf (26 November 2011)

Merriam-Webster.com *Plagiarize* www.merriam-webster.com/dictionary/plagiarize (20 August 2011)

Plews v. Pausch, 2006 ABQB 607.

Siegel, Jon *Plagiarism's Defences* jsiegel.blogspot.com/2010/02/plagiarisms-defenses.html (20 August 2011)

Wikipedia *Common knowledge* en.wikipedia.org/wiki/Common_knowledge (20 August 2011)

Wikipedia *Copyright Act of Canada* en.wikipedia.org/wiki/Copyright_Act_of_Canada (20 August 2011)

Wikipedia *Fraud* en.wikipedia.org/wiki/Fraud (11 January 2012)

Wikipedia *Feist v. Rural* en.wikipedia.org/wiki/Feist_Publications_v._Rural_Telephone_Service (20 August 2011)

Wikipedia *Plagiarism* en.wikipedia.org/wiki/Plagiarism (17 November 2011)