

对日趋严重的剽窃问题和正确使用科学引用的讨论

石理察 杨毅欣 (安玻利亚州立大学生物科学系 美国堪萨斯州)

John Richard Schrock and Yixin Yang (Department of Biological Science, Emporia State University, U. S. A.)

在几乎所有的发达国家,高中和大学教师发现学生从书籍和互联网摘抄资料而不注明出处现象显著增加。剽窃是作者把别人的文字和或者思想据为己有的现象。尽管生物学教师反复告诫学生不要剽窃,可是剽窃现象依然有增无减。产生这种现象的原因可能是文字很容易从互联网上粘贴过来。这种剽窃有时被称作“科技促成便捷剽窃”,或者称作“粘贴复制剽窃”。

1 正确的科学引用

当教师让学生写一个科学题目的报告时,在绝大多数情况下,学生用来整合的材料和信息不是来自于学生自己在实验室从事科学研究得到的数据和成果,而是来自于其他科学工作者发表的工作成果。因此,学生需要在报告中引用别人的陈述和观点。直接引用和解释(改说)是两种正确的科学引用和形式。

(1)直接引用 是直接其他资料来源摘抄的陈述。这些陈述要放在双引号中间,并且要清楚地表明资料的来源,我们用 Wilmott 和 Harrison 于 2003 年在《生物教育杂志》发表的一篇题为“指导生物科学专业学生如何避免剽窃的一个练习”的文章^[1]来举例。下面这句话使用了双引号,并且在语句的结尾完整地标明这句话的文章来源,这样的引用是合理合法的:

“对剽窃产生新的兴趣的一个原因很明显是互联

网越来越多地成为材料的来源”^[1]。

如果作者的原文非常精辟而恰当,或者当需要去反驳作者的观点而使得引用作者的原话显得非常重要的时候,直接引用无疑是最好的策略。

(2)解释(改说) 当资料来源的作者的陈述不见得是最恰当最精练的时候,或者只需要简略地说出资料的大意,那么学生作者可以只总结原资料的陈述,但是要注意的是,语句的前后不要加双引号,而且仍然需要标明文章的出处。借用上例举例如下:

Wilmott 和 Harrison 认为互联网是剽窃行为的一个促成因素^[1]。

2 关于辨别剽窃的一个练习

Wilmott 和 Harrison 用一个文字表格来帮助学生理解正确引用及解释(改说)跟剽窃之间的区别^[1]。在安玻利亚州立大学,我们用经过修改后的类似的文字表格训练生物学专业的学生和实习教师。

在表格 1 我们提供了一个类似的练习。在把上述的正确引用及解释(改说)和剽窃之间的区别介绍给学生之后,提供给学生一段来自原始资料的原文以及由原文演变而来的例句。然后请学生判断这些例句是剽窃,还是可以接受的用法,或者是完全正确的用法。学生完成练习表格之后,教师组织学生进行课堂讨论确保所有的学生理解这些区别。

表 1 加深学生对剽窃认识的练习(据 Wilmott 和 Harrison 的文字表格修改而成)

| 学生论文中的语句 | 剽窃? 是/否 |
|---|---------|
| (1)“紫杉醇被美国新化合物实体称为近 20 年以来最伟大的发现之一,它具有独特的微管稳定作用而被普遍用于乳腺癌、前列腺癌、卵巢癌、头颈部肿瘤、非小细胞肿瘤以及类风湿性关节炎的治疗”(×××, 2006)。 | |
| (2)紫杉醇被美国新化合物实体称为近 20 年以来最伟大的发现之一,它具有独特的微管稳定作用而被广泛用于乳腺癌、前列腺癌、卵巢癌、头颈部肿瘤、非小细胞肿瘤以及类风湿性关节炎的治疗(×××,2006)。 | |
| (3)紫杉醇被美国新化合物实体称为近 20 年以来最伟大的发现之一,它具有独特的微管稳定作用而被广泛用于乳腺癌、前列腺癌、卵巢癌、头颈部肿瘤、非小细胞肿瘤以及类风湿性关节炎的治疗。 | |
| (4)紫杉醇被美国新化合物实体称为近 20 年以来最伟大的发现之一,它具有独特的微管稳定作用而被普遍用于乳腺癌、前列腺癌、卵巢癌、头颈部肿瘤、非小细胞肿瘤以及类风湿性关节炎的治疗。 | |
| (5)紫杉醇被美国新化合物实体称为自上个世纪九十年代以来最重要的发现之一,它具有其他有机物少见的微管稳定作用而被普遍用于治疗乳腺癌、前列腺癌、卵巢癌、头颈部肿瘤、非小细胞肿瘤以及类风湿性关节炎。 | |
| (6)紫杉醇作为新药被药品研发机构称为最伟大的发现。它能够稳定在有丝分裂中起重要作用的细胞微管结构,从而抑制肿瘤细胞的有丝分裂。现在已被普遍应用于各种癌症的治疗(×××,2006)。 | |
| (7)紫杉醇作为新药被美国药品研发机构称为自上个世纪九十年代以来最伟大的发现之一。和其他有生理活性的生物碱相比,它能够稳定在有丝分裂中起重要作用的细胞微管结构,从而抑制肿瘤细胞的有丝分裂。现在已被普遍应用于某些癌症的治疗(×××,2006)。 | |

附资料原文:紫杉醇被美国新化合物实体称为近 20 年以来最伟大的发现之一,它具有独特的微管稳定作用而被广泛用于乳腺癌、前列腺癌、卵巢癌、头颈部肿瘤、非小细胞肿瘤以及类风湿性关节炎的治疗(×××,2006)。

在课堂讨论中,我们仿照 Wilmott 和 Harrison 的文章中的解释方法去阐述剽窃以及如何正确使用原始资料。

(1)这个用法是正确的。引号表明所有的文字都来源于原文,而且原文的来源出处也被明确地标明。当然,我们假定学生的论文在末尾的参考文献部分列出了这篇原文的出处。

(2)这个用法是剽窃。学生作者用了资料文章的原文,但是却没有使用双引号。虽然学生作者标明了原文的出处,但是这种做法却会让人错误地认为学生作者已经改写了原文的论述。

(3)这种用法是明显的剽窃。学生作者既没有使用双引号,也没有标明原文的出处,而是直接使用了资料文章的原文。这种做法使人误认为这句话表达的思想是学生作者的创作。

(4)这种用法还是剽窃。虽然学生作者改变了原句中的几个词,但是没有使用双引号,也没有标明原文的出处。这种做法也会让人误认为这句话表达的思想是学生作者自己的创作。

(5)和上句相比,这句话中学生作者改变了原句中更多的词语。但是作者并没有用自己的话来改写原文。在这种情况下,如果作者没有标明这句陈述的来源,也是剽窃行为。

(6)在这句陈述中,作者用自己的话来改写原文,并标明了原始资料的出处。但是问题在于,学生作者除去了原文陈述中的限制性用语,而使得改写后的陈述结论不完全符合原始资料的观点。

(7)这种用法是正确的。学生作者正确地理解原文的思想和概念,并用自己的方式来陈述。同时也明确地标明了原文的出处。

学生经常会问,“每行中使用几个与原始资料陈述中相同的词会被算作剽窃?”其实,没有一个明确的数字来回答这个问题。但是,按照相同的顺序每多使用一个原文中的词,都会使改写后的句子让人感觉更不是原创。我们教学生在写作时要采用诚实的态度,而不是教他们去数修改几个词会使改写后的句子合理合法。在这个问题上有这样一个原则可以遵循,那就是,如果一个学生在想,“我必须换原文中的多少词可以使我的陈述看上去不像在抄袭”,那么他(她)就是在剽窃了。作为教师,我们要给学生灌输在写作中很清楚地注明陈述和观点出处的观念,这种观念被称作“鸣谢知识援借”。

3 侦测剽窃手段

教师有时会怀疑学生论文中使用的一些词汇和语句对于一个生物学初学者来说太优美和精炼了。许多

学生会用一些常用的搜索引擎在互联网寻找相关资料,因此教师可以在搜索引擎作为关键词输入学生使用的词汇来寻找这些原始资料。目前,一些美国的教师开始使用商用软件工具来检测整篇学生论文以找到互联网剽窃的蛛丝马迹。这些商用软件工具包括 Turn-It-In, EVE2, PalgiServe 和 WordCheck。这些软件工具可以进行在线网上搜索,也可以在软件公司建立的文献数据库中进行广泛地搜索。在美国,一些商业机构向学生兜售他们所谓的“研究性”论文,并声称只有学生而没有教师读过这些论文。事实上,这些论文是为那些想买论文并以此冒充为自己的论文来交给老师的学生设计的。剽窃侦测软件使得教师能够探测到这些商业论文库中的文章以及互联网上的原始资料。

4 剽窃不是一个侵犯著作权的问题

有人提出抄袭和剽窃只是一个侵犯著作权的问题,而教师惩罚剽窃的目的只是保护刊物和书籍的商业所属权。其实侵犯著作权不是自然科学教师惩罚剽窃的主要原因。标明所有引用的出处以及鸣谢知识援借是从事科学研究的一个至关重要的部分,所以也应该成为所有自然科学专业学生早期训练的重要内容。下面的例子帮助说明引用在证实科学发现是如何的重要,以及为什么我们必须追溯科学文献中陈述和观点的产生源头。

(1)一个研究人员声称,他听说有调查研究在美国和中国的高中毕业生的比较中发现,中国学生的数学成绩要比美国学生好很多。

(2)一个同事想知道这个调查研究的更多细节。但是如果这项研究报告的文章来源,就得不到有关科研方法的细节,那么那个研究人员的断言的正确性就不得而知。

(3)那个研究人员提供了研究报告的文章来源。通过阅读这篇引用文章,我们发现这项研究比较的是中美两国所有的高中毕业生。而我们知道,大多数的美国学生会接受高中教育,而中国只有学习较好的那一半学生接受高中教育。

(4)因此,这个调查研究中的比较是有缺陷的。比较应该在中国高中学生和美国排名前一半的高中生中进行。

这个例子说明了能够追溯陈述和观点源头的重要性。追溯陈述和观点源头是科学不断纠正自己并得到发展的方法之一。在这个例子里,著作权并不是问题所在。

诚然,如果一个作者在写书时从其他作者的作品中摘抄了文字却没有加引用注释,这两个作者会陷入著作权的争吵。但这不是要求加引用注释的主要原

因。只有明确地说明产生当前论文的一系列已有的研究成果,才能使证实在论文中论述的每一步产生的结论成为可能。科学研究都是建立在前人研究工作的基础之上,因此我们要鸣谢“知识援借”。当我们训练学生做论文的时候,我们必须让他们知道,作为一个未来的科学研究人员,他们必须严格明确地执行对“知识援借”的鸣谢,即注明前期研究成果的所有者。

中国和西方国家著作权都确实得保护作者的文字。然而,和西方国家类似,中国 1990 年颁布的著作权法 22 条规定,“以批评或者评论为目的而作的引用不违反版权法”^[2]。在科学著作中,经过详细注释的引用被广泛使用,这种使用是介绍一系列科学发现的过程,或者用来批评和修订结论的核心部分,是完全合理合法并且必要的。但是,这里需要注意的一点是,以评论和综述为目的的摘录不能用被引用文章的全文所构成。

5 来美国留学的亚洲学生的剽窃问题

虽然美国学生的剽窃问题也是日益严重,但是大学教授发现来自亚洲的留学生有时候更可能忽略剽窃问题。这个现象可能有如下几个因素促成。

在印度、中国和其他亚洲国家,老师往往充当“师傅”的角色,而学生只是重复老师的话和想法。能够最好地拷贝师傅的学生就是优秀的学生。

在亚洲国家,教育着重于训练学生如何顺利地通过关键性的大考或者是毕业考试。因此,对正确答案的死记硬背经常被看作是学习的唯一目的。对问题所有学生都有一个完全相同的标准答案。另外,虽然学生可以在学习上帮助他人,但是学生的功课都以个体的形式被单独批改。

在中国和其他亚洲国家,有抱负学生、上进的学生人数众多而师资力量不足,因此学生面对很大的家庭和社会压力,迫使他们的一些依靠死记硬背才能得到高分的考试中取得成功。学生没有时间也没有理由去产生自己的观点想法,或是找到新的或者不同的方法去解答一个问题。

当这样的学生来到美国,他们发现老师会让他们阅读若干文献资料,然后整合产生出自己的观点和想法。他们对这样的课业要求很不适应,会习惯性地到文献中寻找“标准答案”交给老师。

另外,美国的学生经常会被要求几个人合作完成一个作业方案。教师会根据合作成果的质量为这个组的所有学生给出一个共同的成绩。亚洲学生对这种课业模式很不熟悉,他们经常会理所当然地假定提交和组里其他同学相同的作业是被允许的。

6 什么是“自我剽窃”

在过去的几年,西方国家的科学期刊的编辑们开始用剽窃侦测软件检查科研人员提交的科学论文,并发现了极个别科学工作者不注明引用出处的案例。令人惊讶的是,最常见的剽窃方式竟然是剽窃者使用他们本人以前发表过的论文中的词语,句子,或者基于这些论文的改写。因为科学期刊通常要求作者提交新的科研成果,因此提交以前发表过的研究成果,或者提交通过大规模摘抄以前的工作成果而合成的“成果”,都是违反期刊杂志的文章发表规则的做法,也是对金钱的一种浪费(同样的研究被发表两次)。这种剽窃称作“自我剽窃”。现在发现这种剽窃比想象中的更为普遍。大学教师承受着发表尽可能多的论文的压力。这可以被大家理解,但是不能作为剽窃的借口。自我剽窃发生在所有的文化背景环境中。在课堂上,如果教师在布置论文作业时没有清楚地说明提交的论文必须是一篇新的,以前未曾提交的文章,学生也会自我剽窃。美国的大学对教授的剽窃他人成果行为的惩罚是极为严厉的,教授经常会因此丢掉工作。与此相比,自我剽窃只被当作是小过失。即使被杂志期刊的编辑发现,作者所需要面对的也只不过是论文的提交被拒绝,而没有其他更严重的后果。

7 间接引用和对有限制性结论的引用的危险性

学生或者研究人员有时会从一本书上看到一个引自另一本书的思想或者观点。在写一篇新的论文时,作者会使用书中的词语和思想,并把对文献的引用照搬过来,而不回到最原始的被引用文献查看其原文和上下文的语言环境。这种做法很普遍,也不违背任何规则和行为准则。但是,这种被许多作者一系列地逐层引用,并且每一个作者都把引用内容放在自己的语言环境中的做法,很容易扭曲或者偏离被引用内容在最原始文献中的真正含义。这种间接引用的做法至少是当前作者惰性的表现。

在最坏的情况下,间接引用可能无意识地成为完全不合理的误用,并可能作为一个似是而非的弱论据去支持作者文章的结论。Simkin 和 Roychowdhury 使用数学方法,根据引用中的错误的重复率,推测出“只有大约 20% 的文献引用者真正读过原始文献”^[3]。

在引用其他科学工作者著作的结论时,一定要注意作者对结论所加的限制。比如说,研究表明一种药物在减少某些类型的肺肿瘤的疗效上有百分之四十有效率。学生在报告这个研究结论时可能把它引用或者改写成“对肺肿瘤有疗效”,而没有注明“百分之四十”、“某些类型”这些限制。一个研究人员曾经讨论过这个问题:“当科学工作者引用或者用自己的话改写以往研究的结论时,他们经常改变用来描述结论不确

定性的限制性用语,从而改变了以往研究结果的不确定性的事实”[4]。

8 明确说明各个作者对文章的贡献

现代的科学研究越来越依赖若干个研究人员的团队工作。因此,现在的科学论文比以前更可能拥有许多合著者。然而,随着一些科学欺诈行为被发现和报道,人们发现虽然一些人的名字作为合著者出现在文章中,他们对发表的科研成果所知甚少,甚至一无所知。文章的原创者经常把帮助申请科研经费的行政管理人员,帮助分析数据的数学专家,以及其他一些没有真正参与实验工作的人的名字列在合著者的行列。Sigma Xi 科学研究学会于1984年出版的《科学的荣誉》一书中讨论了这个问题,并给出如下建议:①所有的合著者应该在文章提交发表之前阅读原稿。②所有的合著者都应该对文章的内容负责,即能够证实数据的真实和准确性。③当不同的作者贡献不同的专业技能时,在“鸣谢”部分里应该写明各个作者在研究中所起的作用。举例说明:“史密斯采集数据,琼斯分析数据,布朗饲养动物”[5]。

虽然上述建议已经提出二十多年,科学工作者和学术期刊很少会采用这个方法去阐明各个合著者的作用与贡献。当数据被发现有问题或者存在欺诈行为时,合著者不能够为发表的文章辩护。对每一个合著

者在文章中的作用做出说明,像本文在下一个段落中所做的那样,正在重新成为论文发表的一个重要要求。这个要求有利于发现像韩国科学家黄禹锡在干细胞克隆研究中造假这样的伪科学行为。

(说明:本文第一作者石理察是本刊的海外编委,他起草了这篇文章的英文原稿“The growing problem of plagiarism and the correct use of science citation”。这篇文章的内容来自于石理察于2006年在广西师范大学,陕西师范大学,以及北京师范大学开设的讲座。杨毅欣为本文提高学生对剽窃辨别能力的中文范例重新设计制作了 Wilmott - Harrison 文字表格,并把全文译成中文)。

主要参考文献

- [1] Wilmott CJR. and Harrison TH. 2003. “An exercise to teach bioscience students about plagiarism”. *Journal of Biological Education*, 37 (3): 139 - 140
- [2] Qu SQ. 2002. *Copyright in China*. Beijing: Foreign Languages Press
- [3] Simkin, MV. and Roychowdury VP 2002. Read before you cite! Condensed Matter website; <http://www.arxiv.org/abs/cond-mat/0212043>
- [4] Horn K. 2001. The consequences of citing hedged statements in scientific research articles. *BioScience* 51(2): 1086 - 1093
- [5] Sigma X. 2000. Honor in Science. Research Triangle Park NC

Shengwuxue-6

The Growing Problem of Plagiarism and the Correct Use of Science Citation

by John Richard Schrock and Yixin Yang

Department of Biological Sciences, Emporia State University

In all developed countries, teachers at both high school and college level have seen a dramatic increase in students copying materials from books and the internet without crediting the source.

"Plagiarism" occurs when a writer represents some else's words or ideas as his or her own. In spite of biology teachers spending more time warning students not to plagiarize, the practice appears to be increasing rapidly, and is probably due to the ease by which text can be copied from internet sources, sometimes referred to as "technology driven plagiarism" or cut-and-paste plagiarism.

CORRECT SCIENCE CITATION

When a teacher asks a student to write a report about a science topic, in nearly all cases, the information the student must assemble will not come from laboratory research the student has conducted in the lab, but will come from the ideas generated from other researchers and reported in print.

Direct Quotation: A statement can be directly taken from another source when the words are enclosed in quotation marks and the source is clearly cited. For instance, Chris J. R. Wilmott and Tim H. Harrison wrote an article titled "An Exercise to Teach Bioscience Students About Plagiarism" in the *Journal of Biological Education* in 2003. It is legitimate to quote the following line as long as I use quotation marks and fully cite the article in the references at the end, as we do here:

"One reason for the renewed interest in plagiarism has clearly been the rise of the Internet as a source of material" (Wilmott and Harrison, 2003).

A direct quotation may be the best strategy when that author has "said it best" or when there is a need to dispute the author's statement and it becomes important to use the author's exact words.

Paraphrase: When a source author has not "said it best" or when the student writer only needs to refer to the general idea briefly, it is appropriate to summarize the source statement but there still must be a citation. In a paraphrase, the exact words of the source author have been generalized by the student and no quotation marks are used:

Wilmott and Harrison (2003) believe the Internet is a factor in plagiarism.

AN EXERCISE TO TEACH ABOUT PLAGIARISM

Wilmott and Harrison (2003) describe a worksheet they provide to students to help them understand the difference between correct quotation and paraphrasing, and plagiarism. At Emporia State University, we have modified the worksheet and used it in training both biology majors and biology student-teachers.

In Table 1, we provide a similar exercise in Chinese. After the above distinctions are described to students (quotation, paraphrase, plagiarism), the students are provided with an original source wording, and then with samples from student essays. Students are asked to decide whether the essay usage is plagiarism, or fair and correct use. After all students have completed their worksheets, class discussion is used to make sure all students understand the distinction.

Table 1. Student Plagiarism Worksheet (after Willmott and Harrison, 2003)

Source Wording:

Pacilitaxel is a new chemical entity that is believed in the United States to be one of the greatest discoveries in the recent twenty years. Due to its unique function of stabilizing microtubule, Pacilitaxel has been widely used to treat cancers of breast, prostate, ovary, brain and head, non-small cell tumor, and rheumatoid arthritis.

Student Essay Wording:

Plagiarism?

Yes/No

| | |
|---|--|
| 1. "Pacilitaxel is a new chemical entity that is believed in the United States to be one of the greatest discoveries in the recent twenty years. Due to its unique function of stabilizing microtubule, Pacilitaxel has been widely used to treat cancers of breast, prostate, ovary, brain and head, non-small cell tumor, and rheumatoid arthritis". (Chu, 2006) . | |
| 2. Pacilitaxel is a new chemical entity that is believed in the United States to be one of the greatest discoveries in the recent twenty years. Due to its unique function of stabilizing microtubule, Pacilitaxel has been widely used to treat cancers of breast, prostate, ovary, brain and head, non-small cell tumor, and rheumatoid arthritis. (Chu, 2006) | |
| 3. Pacilitaxel is a new chemical entity that is believed in the United States to be one of the greatest discoveries in the recent twenty years. Due to its unique function of stabilizing microtubule, Pacilitaxel has been widely used to treat cancers of breast, prostate, ovary, brain and head, non-small cell tumor, and rheumatoid arthritis. | |
| 4. Pacilitaxel is a new chemical entity that is believed in the United States to be one of the most important discoveries in the recent twenty years. Due to its unique function of stabilizing microtubule, Pacilitaxel has been extensively used to treat cancers of breast, prostate, ovary, brain and head, non-small cell tumor, and rheumatoid arthritis. | |
| 5. Pacilitaxel is a new chemical entity that is believed in the United States to be one of the most important discoveries since 1990s. Due to its activity of stabilizing microtubule, which most of other organic compounds lack, Pacilitaxel has been extensively used to treat cancers of breast, prostate, ovary, brain and head, non-small cell tumor, and rheumatoid arthritis. | |
| 6. Pacilitaxel is claimed by drug development agencies as the greatest discoveries. It can stabilize microtubules that play important roles in mitosis, thus inhibiting the mitosis of tumor cells. Pacilitaxel has been now extensively used to treat all types of cancers. (Chu, 2006) | |
| 7. Pacilitaxel, as a new drug, is claimed by American drug development agencies as the greatest discoveries since 1990s. It can stabilize the structure of microtubules that play important roles in mitosis, a function that most of alkaloids lack, thus inhibiting the mitosis of tumor cells. Pacilitaxel has been now extensively used to treat some types of cancers. (Chu, 2006) | |

In class discussion, we follow the rationale similar to that described by Willmott and Harrison.

1. The first version is correct. The quotes indicate all words come from the source, and the source is cited, and we assume fully cited in the references at the end of the essay.
2. This version is plagiarism. The writer has used exactly the same words as the source without quotes. Although the citation is given, we would wrongly assume the writer paraphrased the statement.
3. This version is plagiarism. The writer has used exactly the same words as the source without quotes, and without any citation is given, we would wrongly assume the writer originated this idea him/herself.
4. This version is also plagiarism. Although the writer has changed a few words, without quotes or any citation, we would wrongly assume the writer originated this idea him/herself.
5. Even more words have been replaced, but the writer has not put it in his/her own words. Without

- a citation, this is probably plagiarism.
6. This version is a paraphrase that is cited, but the paraphrase removes the "hedge" used by the original author to indicate that there is a limitation to the effect. The paraphrase makes the conclusion seem unlimited and certain. (See discussion of indirect and hedged citations below.)
 7. This version is correct. The writer makes a good attempt to understand the concept and put it in his/her own words, and also references the source(s) of the writer's synthesis.

It is common for a student to ask, "How many identical words/Chinese characters in a row constitute plagiarism?" There is no set number of words, but each additional word in same sequence as another source, makes it more unlikely the wording is original. What we are teaching is an attitude of honesty, not a legalistic counting of words. A guideline can be provided on this issue. If a student is thinking "how many words must I change to make this statement not appear to be copied?"—the student is plagiarizing. As teachers, we are asking students to be clear in identifying the sources of their words and ideas, the concept called "acknowledging intellectual debt."

HOW IS PLAGIARISM DETECTED?

A teacher may suspect that some words in a student's essay are too elegant or refined to have been written by a beginning biologist.

Since many students use common search engines to locate materials on the internet, it is often possible for an instructor to merely enter the written excerpt into the search engine and locate the source.

However, to check all segments of an essay for internet plagiarism, U.S. teachers are now using commercial software tools such as Turn-It-In, EVE2, PalgiServe, WordCheck, and others that conduct extensive searches of both online text and databases of papers archived by the software companies. In the U.S., there are commercial agencies advertising their services to students and claiming to provide "research" papers merely to be read by the student. However, it is recognized that they are designed to be bought and submitted as the student's own essays. Plagiarism software allows teachers to detect these commercial essay-bank papers as well as online materials.

PLAGIARISM IS NOT A COPYRIGHT ISSUE

It is sometimes suggested that copying is merely a copyright issue, and that teachers who penalize for plagiarism are only acting to protect commercial property rights of journal and book publishers. This is not the reason science teachers penalize plagiarism. Crediting all sources, or acknowledging intellectual debt is a vital part of advancing research, and is a practice that must begin in the early training of all science students. The following example helps explain how citations are critical in confirming science discoveries, and why we must be able to trace the origin of statement made in the science literature.

1. A researcher asserts that he has heard that in a comparison of U.S. high school graduates and Chinese high school graduates, the Chinese students scored much higher in mathematics.
2. A colleague asks for details. If there is no research citation, there are no further details on the methodology and the discussion ends. No one can be sure if the assertion is correct.
3. The researcher provides a citation. Reading the cited research, we discover that they compared all graduating students in high schools in both countries. We realize that most U.S. students attend high school but only the upper half of China's students attend senior high school.
4. Therefore the study is a flawed comparison. The study should be re-done and compare the upper half of U.S. students with the Chinese students.

Note how it is critical to be able to trace the source of the statements in science. This is one way that science constantly corrects itself and advances. Nowhere in this example is "copyright" at issue.

It is true that sometimes two writers will engage in a copyright dispute when one has copied text from another in writing a book, and has failed to give credit. But that is not the primary reason for requiring citation. By clearly showing the line of research that has led up to the current paper, it is possible to confirm the conclusions presented at each step. Since research builds on the work of others, this is called: acknowledging "intellectual debt." When we train students to write papers, we must teach them how, as future researchers, they must practice this clear acknowledgment of "intellectual debt" or credit to those whose work has come before.

Copyright does protect the words of an author in both China and the Western countries. However, similar to the West, Article 22 of China's Copyright Law 1990 states "No infringement exists in the use for purposes of criticism and review" (Qu, 2002). The common use of fully attributed quotations in science work is an integral part of documenting the chain of discovery, or criticizing and modifying conclusions, and is fully legal and necessary. However, such excerpts for review should not constitute the total text of the cited work.

PROBLEMS WHEN ASIAN STUDENTS COME TO THE U.S.

Although a growing number of American students are committing plagiarism, College professors are noticing that students from Asia are sometimes more likely to not recognize this problem. There appear to be several factors involved.

In Indian, China and other Asian cultures, the teacher has often served the role of "master" and the student is to repeat the master's words and ideas. An excellent student is the student who can copy the master.

Education is usually centered on training to pass a critical final test or leaving examination. Therefore recitation and memorization of the correct answer is often seen as the only goal of learning, and the perfect answer can be the same exact words used by all students. While students may help each other study, all student work that is graded is individual.

Because of the high number of aspiring students and low number of teachers, there is great family and social pressure to score high on rote exams, and there is no time nor reason to ask students to synthesize their own ideas or create new or alternative ways of explaining a subject.

When such students come to the United States, they find that the teacher is asking them to read several sources and to provide their own synthesis of what they mean. They are not used to this, and often go looking for a "perfect answer" in the literature to provide to the teacher.

U.S. students may be asked to work together on a project, and all students in the group may be given one grade for the communal work. Asian students are unfamiliar with this practice, and often assume that this permits copying and submitting the same work.

WHAT IS "SELF PLAGIARISM"?

In the last several years, Western science journal editors have begun using plagiarism software to check new science articles submitted by researchers, to catch the very few cases when a scientist fails to credit a source. To their surprise, the most common form of plagiarism was using the exact same words, sentences, and paragraphs the scientist had used in a previously published journal. Since science journals usually require that authors submit only new research, the use of already-published research or work based only on extensive quotations from previous work, called "self plagiarism," is a violation of the journal rules and also a waste of money (publishing the same study twice). This practice has been found to be far more common than suspected. The university-level pressure to publish as many articles as possible is well understood but is not considered an excuse. Self-plagiarism is found across all cultures. This practice occurs in the classroom when a teacher requires a paper, and does not clearly state that it must be a new composition and any work written for a previous class may not be resubmitted. While penalties for extensive plagiarism by university professors in the U.S. are severe, often involving loss of job, self-plagiarism is usually considered a minor infringement and when discovered by editors, the paper is generally rejected with no other consequence for the writer.

THE DANGER OF INDIRECT AND HEDGED CITATIONS

A student or researcher may locate an idea expressed in a book that in turn has been cited from another book. In writing a new paper, the student may lift the words or ideas from the book, repeating the citation, without going back to the original cited source to check the original wording and context. This is a common practice, and does not violate any rules or code of conduct. But such a line of serial quotation going back through many authors, each of whom have used the quote in their own context, has the danger of distorting or drifting away from the meaning intended by the original source. At the least, this is a laziness on the part of the current writer.

At the worst, it may be an unknowing misuse that is now completely inappropriate, and seem to provide support for a conclusion when the support is weak or unjustified. Simkin and Roychowdhury (2002) have used a mathematical method, based on the rate at which errors in citations are repeated, and estimate that "...only about 20% of citers read the original."

It is also necessary to be very careful to cite another scientist's conclusion with attention to the limitations, or "hedged," the scientist puts on the conclusions. For instance, research may show that a drug is forty percent effective in reducing certain types of lung tumors. A student reporting on this may quote or paraphrase it as "being effective against lung tumors," failing to note the limitation to "forty percent" or to just "certain types." As one researcher describes this problem (Horn, 2001): "When scientists cite and paraphrase the conclusions of past research, they often change the hedges that describe the uncertainty of the conclusions, which in turn can change the uncertainty of past results."

CLEAR DESIGNATION OF CREDIT

Modern research more-and-more relies on teams of researchers. Therefore today, a science article is more likely to contain many co-authors. However, isolated cases of science fraud have revealed that some co-authors have limited or no knowledge of the published research where their name appears as co-author. It is common for the authorship to list administrators who secured funding, math specialists who helped in the data analysis, and others whose hands may not have participated in the actual labwork. This problem has been explained in an early edition of "Honor in Science" published by Sigma Xi in 1984 which recommended the following:

- All co-authors should have read the manuscript prior to submission for publication.
- All co-authors should be ready to take responsibility for the article's content, that is, attest that it is real data and accurate.
- Where different members of the team contributed different skills, the "Acknowledgments" section should clarify each author's role as in the example: "Smith took the data, Jones analyzed it, and Brown fed the animals" (Sigma Xi, 2000).

While this last recommendation has been around for over 20 years, very few scientists and journals are using this method of clarifying the nature of co-author's contributions. When data are found to be questionable or fraudulent, co-authors continue to be unable to defend the published paper. Description of the role of each co-author, as demonstrated in the next paragraph below, is now being revived as an important requirement that would help detect faulty science, as in the recent case of fraudulent cell cloning by the Korean scientist Hwang Woo-Suk.

Acknowledgments

Schrock drafted the original text of this article from lectures given at Guangxi Normal University, Shaanxi Normal University, and Beijing Normal University in 2006. Yang redesigned the Wilmott-Harrison worksheet for Chinese examples and translated this full text in Chinese.

References Cited:

- Horn, Kelly. 2001. The Consequences of Citing Hedged Statements in Scientific Research Articles. *BioScience* 51(2): 1086-1093.
- Qu, Sanqiang. 2002. *Copyright in China*. Foreign Languages Press; Beijing. 436 pages.
- Sigma Xi. 2000. *Honor in Science*. Research Triangle Park, NC. 41 pages.
- Simkin, M. V. and V. P. Roychowdhury. 2002. Read Before You Cite! *Condensed Matter* website: <http://www.arxiv.org/abs/cond-mat/0212043>
- Wilmott, Chris J.R. and Tim H. Harrison. 2003. "An Exercise to Teach Bioscience Students About Plagiarism." *Journal of Biological Education* 37(3): 139-140.