Evaluating web-sources: Internet literacy and L2 academic writing

Paul Stapleton

Using the World Wide Web (hereinafter the 'Web') has become an increasingly viable way to source information in academic writing. However, the Web is a fundamentally different type of resource from conventional sources such as books and journals. This paper argues that there is a need among learners for a heightened critical awareness of web-source nuances. Specifically, although the Web contains a much greater amount of information than conventional research sources, a great proportion of its sites harbour ideological agendas. It is suggested that L2 learners, in particular, may require both consciousness raising and practice in recognizing the biases that exist in websites. The paper concludes by describing a pilot study which produced some practical steps that teachers can take to introduce and enhance Internet literacy.

Introduction

In a recent Internet key word search using *Google,* I entered a word that turned up 4 million results. The word I had keyed in was 'abortion'. This led me to try other words that are at the heart of some touchstone issues in society, as well as those often discussed in academic writing texts (see Ramanathan and Kaplan 1996: 248). 'Equal rights' turned up 8,880,000, 'gun control', 4,070,000, 'death penalty', 3,360,000, 'euthanasia', 590,000, and 'marijuana legalization' measured in at 89,500.

These numbers highlight a larger issue. Slightly over a decade after user-friendly browsers first appeared, the Web has become both a widespread and legitimate source of information retrieval. The fact that APA and other referencing systems now have a formalized way to cite Internet sources suggests that the Web is viewed as a major source of information in academic writing as well. However, these numbers also illustrate that strategies are needed to deal with the scope and quality of information available. Despite this growth in size and validity, significant differences remain between the new electronic sources of information, and conventional paper and ink ones. Such a statement brings to light the much-discussed recent notion asserting that the meaning of 'literate' is constantly changing (Valmont 2003). Valmont argues that because of the recent shift to telecommunications technology, a continuous change in instructional methods is also required. In effect, the move from paper to the screen as the dominant form of communication is predicted to have far-reaching effects (Kress 2003: 9) and mark a new era which requires

an 'electronic literacy' (Shetzer and Warschauer 2000). In this paper, I will argue that one facet of this new literacy, i.e. effectively using the Web as a research resource, can be accomplished, first via a recognition of the differences from conventional sources, then by way of teaching Internet literacy and practicing Web sourcing.

The library as a filter

Although we generally don't think about it, when we pick up a book or journal in a library, we have an unconscious understanding that it has undergone a screening process before it has reached our hands. The printed words that reach our eyes were once a manuscript that first had to have been submitted to editors who, with the assistance of reviewers, decided upon its worthiness. A final decision to publish was made based on the manuscript's quality (and marketability), along with the author's degree of cooperation in making requested changes. If the resulting book or article happened to come from a reputable publisher, and librarians or professionals in the relevant field were convinced of its significance, the newly published work will have found its way to the library. In this sense, the paper and ink sources found in libraries have gone through a rigorous filtering process before reaching the eyes of the reader.

On the other hand, although there are some web pages which have undergone a similarly rigorous screening process, most do not (Kirk 2002). This is because anyone with Internet access and knowledge of web page construction can 'publish' their own page, and become one of the 4.3 billion pages (as of September, 2004) scanned by *Google*. In comparison, the world's largest library, the U.S. Library of Congress, has 18 million books. This fundamental difference in quality between conventional and Web sources for information retrieval has implications for academic writing because it suggests there is a need to develop a heightened level of vigilance and critical evaluation when using the Web as an information source (Shetzer and Warschauer 2000). This need may be further underlined in the fields of English for Specific Purposes and L2 writing, where the few existing studies in this area have pointed to students' uncertainty about the status of web-sources in academic writing (Slaouti 2002: 110), and how quality varies between different types of sources, e.g. government, interest group, and news site (Stapleton 2003).

Agendas of persuasion

Librarians quickly recognized this need soon after the Web became a viable source of research information in the mid-1990s. Now, dozens of websites detailing web page evaluation are available, mostly written by librarians (see Auer 2002 for a partial list). Many of these reach the same conclusion, and are summarized below as key questions which should be asked by those retrieving information from the Web:

1 Who is the author?
2 What authority does the site have?
3 How current is the information?
4 What is the intended audience?
5 What agenda (if any) does the author have?
6 Is the content biased?

Naturally, these same questions are assumed by good critical readers, whether they are reading electronic or conventional texts. However, while

books and journals follow formats with title pages which include the author's name and date of publication, among other standard features, such as tables of contents and indexes, web pages have fewer conventionalized standards. Indeed, a good proportion of web pages fail to include the authors' names or the date of publication—something which is rare in paper and ink materials. In addition, the non-linear textual nature of a web page, which generally includes greater use of photos, graphics, colors, sounds, and video, plus hyperlinks to other sites on the web at the click of a mouse, represents a fundamentally different form of media.

Perhaps a more important difference between the Web and conventional sources, however, concerns questions 4–6 above. These items hint at the web's various agendas, which are more often motivated by commercial and ideological rationales than the generally less subjective, information-dispersal agendas found in conventional sources. Certainly such a statement is not suggesting that ideological agendas are confined to the Web. Indeed, much writing in the conventional sources found in libraries is motivated by ideology. For example, environmentalists publish books with conservationist messages, or postmodernists espouse the existence of multiple discourses. However, the Web represents a fundamentally different form of media which, unlike the publishing industry, has numerous motives and causes. As Tillman (2000) claims that '[t]he Internet has enabled a vast new group to enter the world of publishing—those who didn't learn the culture of the print publishing trade.' For example, a domain name such as '.com' is suggestive of an enterprise, while the '.org' domain name could be attached to NGOs (non-government organizations), or groups with political, religious, or any number of other ideological associations. Compared to conventional research sources found in libraries, the Web's wider raison d'être means that an awareness of these agendas of persuasion among students learning to write for academic purposes are essential (Stapleton 2003: 242). This awareness begins with the use of a search engine.

Search engines When using the Internet for research purposes, search engines serve a similar function to a library catalogue; however, there is a key difference. *Google*, which is now used for about eighty percent of all Internet searches on sites that license its technology, uses a system of one hundred or so closely guarded algorithms to generate its search results (McHugh 2003). These algorithms 'allocat[e] relevancy to a page according to the number and importance of pages linked to it, the number and importance of pages linked to each of those pages, and so on' (McHugh 2003: IV¶4). Because the user does not know exactly how search results are generated, there remains a concern as to how the order of the results is produced. No one has the time to probe through thousands of screenfuls of information; therefore, if the first few pages of results have a bias towards a given ideological position, it is possible that a certain agenda could persuade the reader in a direction that is disproportionally represented. This is especially the case for L2 students who have less of a grasp of English, or those who may have been raised in a culture where there is said to be less of a critical mindset (Ramanathan and Atkinson 1999: 61).

In order to illustrate how a simple key word search can turn up skewed results, I entered the words 'whale hunting' into *Google*'s search engine as a test case. The hunting of whales for their meat or for research purposes is a prominent political and social issue with strong arguments both for and against, which I often use in academic writing courses that I teach at a university in Japan. Among the first thirty resulting links, thirteen were against the hunting of whales, eight were neutral, five did not address the particular issue of hunting whales for meat or research, three were not in English, and two were supportive of whale hunting. Even though the word 'hunting' is a neutral term ('killing' would skew the results even more towards the anti-hunting position), these results show a decided bias towards an environmentalist ideology. And even though *Google* is generally acknowledged as being ethically responsible, the algorithms that they use may not be able to consistently generate balanced results which are neither skewed toward or against a particular agenda among the first few screenfuls of links.

While it is certainly true that a search in the library could turn up books that are equally motivated by ideology, such a search is unlikely to produce either a similar bias or sources in such numbers. In the results above, interest groups and personal websites, i.e. 'authors' that are less likely to have their works turn up in libraries, were over-represented. Although this one sample search cannot be generalized to other searches, it demonstrates how search results can be skewed towards certain agendas, and highlights the need for an additional set of skills for researching on the Web. Again, this need appears to be magnified in the case of those L2 learners who may have less familiarity with some of the issues that are typically discussed in academic writing classes in English language programs (see Stapleton 2001).

Critical evaluation of web-sources

Once one has arrived at a website, a determination of the type and overall quality, including the potential biases of the author (based on questions 4–6 above) is also essential. Here, the term 'type' refers to the kind of organization that has created the website: company, government, interest group, educational institute, news site, personal website, etc. For example, primary sources of information, such as scholarly journals, are generally regarded as being more desirable references than secondary sources. Because a large number of the 4 billion-plus sites accessible on the Web have been uploaded without any editing or vetting, students need to be made aware of what constitutes a primary and secondary source.

Determining biases in ideology requires a close examination of the information and arguments set forth. Even if one is simply searching for numerical data, there is potential for subtle biases. For example, in a search that I made of opposing sites, this time on the topic of gun control, such biases were clearly evident. Both sides (pro- and anti-gun control) argued about assault weapons, using statistics that strongly supported their views. The pro-gun control site claimed that assault weapons comprised only 1% of privately owned guns, yet they accounted for 8.4% of all guns traced to crime. On the other hand, the anti-gun control website claimed that only 2% of confiscated guns were assault weapons,

and just over 2% of criminals that used guns used assault weapons. Clearly, via careful wording, either side can use statistics to support their own viewpoint. However, statistics can be used in even more insidious ways. For example, the anti-gun control site argued that an act of self-defence with a gun occurs every thirteen seconds. While such a statistic (assuming it is accurate) appears to strongly support the anti-gun control agenda, it fails to acknowledge how many of those acts of self-defence were among gang members or criminals.

Pedagogical approach

This concern about the suitability of using web-sources in academic writing leads to the question of how to engender an ability in students to evaluate websites and recognize the partiality and weak reasoning associated with ideological agendas. Accordingly, the following description explains how I attempted to lead seven students (five Japanese, one Russian, and one Bulgarian) in a media and communications graduate school, taking a course called 'Critical Thinking and Writing', towards an ability to critically evaluate web-sources using a four-step approach. Their English writing level ranged from high-intermediate to fluent. The participants in this small class included four females and three males with an average age of twenty-eight years. All students had their own laptop computers, Internet connection, and email addresses indicating some familiarity with electronic media.

Step 1
Search engine practice

This step required some instruction and practice in searching techniques including performing searches using multiple key words to narrow down results. To begin, a list of questions offered an orientation towards looking for factual information. These included items such as: 'When did Gutenberg invent the printing press?' and 'Who won the gold medal in the men's 100 meter race in the 1996 Olympics?' In order to practice searching for the type of analytically oriented information required in academic writing, a second series of exercises in which students had to support their opinions on given issues helped to refine students' searching skills. Here, beyond key word searches using generic search engines, students were encouraged to 'power search' (Tuman 2002), that is, find specialist databases which use the search engines supplied by specialist online databases. One example of this type of exercise asked students to complete sentences similar to the following: 'I believe whale hunting should continue/be banned because . . .' Students were required to supply reasons and evidence, plus supporting citations to websites.

Step 2
Assessing websites

In the second step, students were introduced to the six criteria (questions listed above) by which a website can be evaluated. The first three questions, which aim to identify the author, sponsor, and date, required exposure to several websites, and practice in locating these elements, because unlike books and journals, web-sources have not yet regularized these features. For the fourth question, (What is the intended audience?), the significance of domain names (see above) was discussed. It should be noted here that main domains, such as.com are not always informative; geocities.com, for example, is mostly for personal web pages, not companies. The fifth question (What agenda does the site have?) can

sometimes, but not always, be determined by looking at a site's 'Mission Statement', or 'About Us' statement. Here again, the students were exposed to several examples of these statements at various sites. The final question (Is the site biased?), quite arguably the most important question, required considerable training and is explained below.

Step 3
Describing the various forms of weak reasoning associated with bias

In order to familiarize the students with various types of bias, I introduced several forms of weak reasoning associated with bias that I had isolated from the two whaling websites above, plus other websites. The terms used below for some of the common forms of fallacious reasoning were taken from Ramage and Bean (1999: 239) and may be over-represented in websites harboring ideological agendas. Beside the misleading statistics described above, the following types of weak reasoning associated with bias were described to the seven students.

Fallacies:

red herring—diverting the readers' attention to an irrelevant matter

straw person arguments—the deliberate oversimplification of an opponent's arguments

hasty generalization—taking one example, and assuming it represents the norm

false analogy—making comparisons that stretch believability

slippery slope—the notion that any step in a certain direction will result in extreme consequences

appeal to emotional premises—using emotional rather than logical arguments

Selectivity

the omission of arguments that were unfavorable to the position held by the organization; the inclusion of links that are only favorable to the position taken; the selection of non-textual data, such as photos which cast the chosen position in a favorable light/opposing position in an unfavorable light

Loaded terms:

the use of emotional terms, such as 'outrageous' or 'heart-breaking'

Vague terms:

the use of unquantifiable expressions, e.g. 'the majority believe . . .'

Misinformation and disinformation:

supplying wrong or out-of-date information, either inadvertently or deliberately

Sarcasm:

stating the opposite of what is intended in order to ridicule or be amusing

While these errors of logic or conscious efforts to persuade via unethical means are sometimes obvious, more often they are not. This is why academic writing textbooks often include sections on fallacies and other common difficulties with arguments. In this crucial step, students were exposed to multiple examples of these forms of bias so that they could more easily recognize parallel examples in their own web-searching. Ideally, in keeping with task-based learning, teachers need to determine the interests of students, and then find sites on the Web which display the weak reasoning so often associated with bias. (For example, the issue of whale hunting was determined to be of high interest, given the Japanese context of the seven students).

It is important to note here that because the vast majority of websites are written in English,[1] L2 learners are linguistically disadvantaged. This is especially true when loaded or vague terms are used. Sarcasm is especially difficult for non-native speakers to recognize. L2 learners may also be culturally disadvantaged with regard to recognizing the quality and bias of web-based sources. Large NGOs and think tanks reflecting a spectrum of political and religious ideologies often have sophisticated websites cloaked in a veneer of objectivity. While native English speaking students are often aware of the agendas of such organizations as the NRA (National Rifle Association) in the United States, or Greenpeace, simply by virtue of being brought up in the West, L2 learners may be much less familiar with these types of organizations. This unfamiliarity, coupled with a less than fluent grasp of English, may leave them more open to being persuaded by ideas that lack balance.

| Step 4 Assignment (and results) | In order to put the critical evaluation of web-based sources into practice, the final step followed a similar approach to that used above, where two opposing sites were identified (this time by students themselves) and then assessed. This procedure required students to choose topics of their own personal interest; however, these topics had to concern an issue of controversy in the public domain. Because websites had to be written in English, there was some compunction upon the students to take up broader social, environmental, or political issues, as opposed to local issues in which English language sites may have been lacking (see Stapleton 2005). Certain website types, e.g. news sites and personal websites, were discouraged because bias is more difficult to detect at news sites, while it is often too obvious on personal web pages. |

The assignment itself asked students to identify at least five instances of weak reasoning reflecting bias, and label them using the terms discussed in Step 3. They also had to explain how these biases dovetailed with the agendas that each web-source was determined to have based on the estimation of its perceived audience and Mission Statement. As an additional step, when an area of biased information was found, students had to seek out more objective information at alternative websites.

Upon receipt, the Step 4 assignment revealed that participant students had self-selected web-topics in the following areas: legalization of marijuana, smoking, abortion, euthanasia, human cloning, gun control, and the question of the return to Greece from the British Museum of

the Parthenon Marbles. It bears repeating here that although most of the above topics have been used for years in academic writing classes (see Ramanathan and Kaplan 1996), the Web provides a whole new medium for the collection of views on these topics from groups and individuals that were (and still are) excluded from conventional sources of information, such as libraries.

The indications from this pilot exercise were encouraging with the seven students identifying a total of 75 distinct instances of weak reasoning and bias covering all of those areas outlined above. Students had also labeled most of them accurately (except sarcasm). In addition, the students were able to draw parallels between the biases and their associated agendas. Most importantly, as a result of the training, most succeeded in finding alternative web-sources that appeared to provide a balanced view, and could be considered appropriate for citation in an academic paper.

Conclusion

This paper has focused on reasons why web-based sources need to be treated more critically than conventional sources in academic writing. It has also discussed how sources found on the Web can be evaluated, and how these evaluation skills can be passed on to learners of English, who may be at a particular disadvantage in noticing agendas of persuasion, which tend to be more pervasive than in paper and ink sources. This heightened awareness of bias in research will become increasingly important as information sourcing in electronic form becomes even more widespread.

Final revised version received March 2004

Notes

1 *Global Reach* claims that 68.4% of websites are in English. Japanese comes second at 5.9%. http://global-reach.biz/globstats/refs. php3#overlap (last accessed Feb. 11, 2004).
2 Search engines such as scholar.google.com now provide more focused searches of academic sites.
3 *Global Reach* claims that 68.4% of websites are in English. Japanese comes second at 5.9%. http://global-reach.biz/globstats/refs. php3#overlap (last accessed Feb. 11, 2004)

References

Auer, N. J. 2002. 'Bibliography on evaluating web information.' Available at: http://www.lib.vt.edu/research/evaluate/evalbiblio.html (last accessed Feb. 11, 2004).
Kirk, E. E. 2002. 'Evaluating information found on the Internet.' Available at: http://www.library.jhu.edu/elp/useit/evaluate/ (last accessed Feb. 11, 2004).
Kress, G. 2003. *Literacy in the New Media Age.* London: Routledge.
McHugh, J. 2003. 'Google vs. evil.' Available at: http://www.wired.com/wired/archive/11.01/google_pr.html (last accessed Feb. 11, 2004).
Ramage, J. and **J. C. Bean.** 1999. *Writing arguments.* Allyn and Bacon, Boston.
Ramanathan, V. and **D. Atkinson.** 1999. 'Individualism, academic writing, and ESL writers'. *Journal of Second Language Writing* 8/1, 45–75.
Ramanathan, V. and **R. Kaplan.** 1996. 'Some problematic "channels" in the teaching of critical thinking in current L1 composition textbooks: Implications for L2 student writers'. *Issues in Applied Linguistics* 7/2, 225–249.
Shetzer, H. and **M. Warschauer.** 2000. 'An electronic literacy approach to network-based language learning'. In Warschauer M. and R. Kern (eds). *Network-Based Language Teaching: Concepts and Practice.* Cambridge University Press, Cambridge.
Slaouti, D. 2002. 'The world wide web for academic purposes: Old study skills for new?'. *English for Specific Purposes* 21/2,105–124.

Stapleton, P. 2001. 'Critical thinking in Japanese L2 writing: Implications about content familiarity and assumptions'. *Written Communication* 18/4, 506–548.

Stapleton, P. 2003. 'Assessing the quality and bias of web-based sources: Implications for academic writing'. *Journal of English for Academic Purposes* 2/3,227–243.

Stapleton, P. 2005. 'Using the Web as a research source: Implications for L2 academic writing'. *Modern Language Journal*,85.

Tillman, H. N. 2000. Evaluating quality on the net. Available at: http://www.hopetillman.com/findqual.html (last accessed Feb. 11, 2004).

Tuman, M. 2002. *Criticalthinking.com: A guide to deep thinking in a shallow age.* Xlibris Corporation,.

Valmont, W. J. 2003. *Technology for Literacy Teaching and Learning.* Houghton Mifflin Company, Boston MA.

The author

Paul Stapleton is a professor at Hokkaido University in Japan. His current interest is in academic writing.
Email: paulstapleton@gmail.com