Evaluating Credibility of Information on the Internet

Copyright 2003 Ronald B. Standler

As school pupils and undergraduate students surf the Internet to find information, these pupils and students also need to evaluate the credibility of the information that they find. Because my own professional experience is in scientific research, electrical engineering, and law, the discussion in this essay focuses on those three areas. I urge that readers make their own evaluation of credibility of sources of information, instead of blindly relying on peer review, or instead of ignoring non-peer-reviewed sources.

First, I review the three traditional methods for evaluation of credibility: (1) peer review, (2) credentials of the author, and (3) writing style. These three traditional methods are easy for the reader, but are *not* sophisticated and may give misleading results. Second, beginning at page 4, I present methods that readers can use to make their own evaluation of credibility of information. Finally, beginning at page 6, I suggest that the Internet may be more useful for delivering information than archival journals in a library.

I. Traditional Methods of Evaluation

1. peer review

The traditional standard for credibility was either:

- 1. publication of a book by a reputable publishing company (e.g., Mc-Graw-Hill, Wiley, Springer Verlag, Cambridge University Press, etc.)
- 2. publication in a peer-reviewed, scholarly journal, such as those published by recognized professional societies.

Such criteria for credibility only shifts the burden for determining credibility from the reader to the reviewers chosen by the publisher. I want to emphasize that the credibility of information should depend on the information itself, and *not* the opinion of some expert who endorsed the information. Readers can, and should, make their own evaluation of credibility, even if the work has been endorsed by peer review.

Some physicians and scientists have submitted fraudulent research results to respected journals, which published those fictitious results after approval by peer-review. While one hopes that such episodes are rare, it is clear that peer-review offers no protection against fraud by authors.¹

Alan Sokal, a professor of physics at New York University, deliberately wrote an article full of rubbish, as a parody of "postmodernist literary theory", then submitted his article to *Social Text*, which published it after approval by peer-review.² This incident is sometimes cited as showing that the peer-review process is imperfect,³ although I think this incident more importantly shows the lack of intellectual rigor in some parts of the humanities, which makes whole fields of knowledge and entire journals, not just an individual article, not credible.

It is inherently *unscientific* to trust an expert, instead of making a personal evaluation of the reasons for the expert's conclusion. Peer review is useful to increase the quality of scholarly journals, but peer review is *not* a substitute for the reader taking the responsibility to evaluate the credibility of information in an article.

2. credentials of the author

A commonly used standard is prestige. Has the author earned a doctoral degree from an accredited university? This is a superficial standard for evaluating credibility, but it has its place. For example, I would not consider a book or article about medicine to be credible unless its author was either a professor of medicine or a practicing physician, which includes having an earned M.D. degree. Similarly, I would not consider a book or article about law to be credible unless its author was either a professor of law, judge, or a practicing attorney, which includes having an earned degree in law. Having said this, let me also say that (1) appropriate credentials of the author is *no* guarantee that his/her article will be either credible or significant and (2) there are rare examples of people who are self-taught in an area of medicine or law, so that they can produce

¹ Arnold S. Relman, "Lessons From the Darsee Affair," *New England Journal of Medicine*, Vol. 308, p. 1415-17, 9 June 1983. ("What kind of protection against fraud does peer review offer? The Darsee affair gives a clear answer: Little or none.") Also see, Eric J. Lerner, "Fraud Shows Peer-Review Flaws," *The Industrial Physicist*, Vol. 8, Dec 2002, http://www.aip.org/tip/INPHFA/vol-8/iss-6/p12.pdf. (Discussing fraudulent data in sixteen papers published in prestigious, peer-reviewed journals by Hendrick Schön.)

Alan Sokal, "Transgressing the Boundaries: Towards a Transformative Hermeneutics of Quantum Gravity," *Social Text*, Nrs. 46/47, pp. 217-252, Spring/Summer 1996. Prof. Sokal has a very informative collection of webpages about his hoax at http://www.physics.nyu.edu/faculty/sokal/.

³ See, e.g., Colleen Bell, *Critical Evaluation of Information Sources*, 20 Jan 2003, http://libweb.uoregon.edu/guides/findarticles/credibility.html.

credible and significant work, despite their lack of formal credentials.

If one evaluates credibility by examining the credentials of an author, then an anonymous work has *un*certain credibility. Further, one must be suspicious of anonymous works, because the author is apparently *un*willing to take personal responsibility for his/her work, which suggests that the author knew the work is either poor quality or bogus.

When I find information on the Internet that is posted at a personal website, I expect to find, somewhere in his/her website, both his/her e-mail address and his/her credentials (e.g., a resume or c.v.). The absences of an address or credentials makes me distrust the author.

Any author who has written at least several papers that were published in respected, peer-reviewed scholarly journals is obviously capable of continuing to produce similar quality work, even if that work is posted on the Internet, instead of published in a peer-reviewed journal.

Sometimes one finds a policy statement issued by a professional society, which is allegedly a consensus of members of that profession and sometimes published anonymously. I tend to ignore such pronouncements. Truth does *not* depend on the number of experts who endorse it. There are many examples in history of science where new ideas were opposed by established experts, who perhaps felt personally threatened either by new and *un*familiar ideas or by technology that might make them obsolete. Endorsement by authorities is more characteristic of religious dogma than scientific Truth.

3. writing style

One can easily distinguish scholarly writing from less credible forms of writing by the writing style. Scholarly writing is distinguished by:

- citations, either in text or in footnotes.
- no misspelled words.
- few, if any, grammar errors.
- appropriate vocabulary, including the proper technical terms and no slang.
- precise writing neither vague statements nor sweeping generalizations; careful use of absolute words like "always", "never", or "ideal".
- statements are internally consistent.
- alternative interpretations or opposing viewpoints are considered, unless the purpose of the article is clearly stated advocacy of one viewpoint.
- no sarcasm.
- no *ad hominem* attacks on other people.
- includes a date of last revision or date of submission for publication.4

⁴ A date is important because information or opinions are sometimes made obsolete by subsequent events or discoveries.

• use of generic names instead of brand names or trademarks. In summary, scholarly writing displays the erudition of the writer and the careful revisions that the writer undertook to produce work that he/she felt was of acceptable quality to display publicly.

It is commonly asserted that the hallmark of scholarly writing is large numbers of citations to previously published material in professional society journals, books, government publications, and other trustworthy sources. Citations make it easy for the reader to verify information; clearly acknowledge earlier discoveries, insights or contributions to the subject matter (thus the author avoids misrepresenting work by others as the author's original contribution); and also show the author's awareness of past work (i.e., the author's erudition).

However, publications in theoretical or experimental science tend to have few citations and a short bibliography, because these authors concentrate on presenting their own calculations or measurements, with citations to previous publications typically to provide a motive for the new work and to interpret the new results.⁵ In contrast, review articles in science cite many publications, to explain the historical evolution of ideas in a subject area and to critically review these publications to spot discrepancies and suggest questions that need to be answered by future scientific research.

On the other hand, a credible paper in history or law would contain a large number of citations to authority, either as citations in the text or as footnotes.

II. Evaluation of Content

A more sophisticated standard for evaluating credibility of information is to read the material and then decide if it is credible. This is a complex process that involves many factors. One of the most important factors is the documentation of reasons to believe the conclusion:

- 1. In theoretical science, there should be a derivation of a result, starting from accepted laws of nature and with clearly stated assumptions and approximations. There will be enough detail in the publication to allow a skilled reader to duplicate the derivation.
- 2. In experimental science, there should be a description of the methods of the experiment with enough details to allow a reader (who has adequate resources) to duplicate the experiment. The reader needs to consider whether the methods are adequate, and whether the methods have disadvantages or problems that were *un*recognized by the author.
- 3. In law, there should be quotations from authorities: constitutions, statutes, judicial opinions

An extreme example is Albert Einstein, "Zur Elektrodynamik bewegter Körper," *Annalen der Physik und Chemie (Leipzig)*, Fourth Series, Vol. 17, pp. 891-921, 26 Sep 1905, which contains *zero citations* to previously published work. http://www.wiley-vch.de/berlin/journals/adp/890_921.pdf
This paper gave birth to the theory of special relativity, an extraordinarily novel achievement, but nevertheless citations to earlier relevant work could have been made, and in a review paper two years later Einstein did refer to earlier work (e.g., by Lorentz).

See http://www.garfield.library.upenn.edu/essays/v5p091v1981-82.pdf.

from the U.S. Supreme Court, judicial opinions from state supreme courts, ... as well as citations to the *Restatements of Law*, venerable treatises, and other respected sources. While law follows authority much more than science, even in law there are often ways to get at something more fundamental than appeal to authority. For example, there may be mention of economic principles, philosophy, public policy, etc.

The key principle here is that a conclusion is only as good as the reasons for believing that conclusion. A conclusion stated without reasons, or with *in*adequate reasons, is worthless. It is the reader's task to decide if the reasons are adequate and convincing.

plausibility of information

Another important factor in evaluating the credibility of information is the plausibility of that information. For example, there are situations in science where accurate data can be extrapolated according to accepted laws of nature, to produce an *implausible* (and *wrong*) result. Recognizing that the result is wrong inspires the reader to go back and carefully examine the assumptions and limitations of those "laws of nature" and find the mistake. Running this process in the opposite direction, scientists and engineers who are *both* experienced *and* competent can often "guess" (i.e., estimate) the correct result, without doing a formal calculation or measurement.

Strange or unexpected conclusions need more supporting evidence than ordinary conclusions.

Learning to do this kind of evaluation of credibility of information is an essential part of a genuine education.⁶ Elsewhere, I have listed some specific characteristics of hoaxes about computer viruses, to teach people to be more critical in their acceptance of information.⁷

credibility may be controversial

Finally, it is important to recognize that there may be no unique right answer for credibility of some information at the cutting-edge of new knowledge, where experts who are familiar with the detailed subject matter may disagree about either the credibility, plausibility, or significance of new work. After tens of years, a consensus will emerge about the credibility of past work. But current work at the frontiers of knowledge may be controversial and disputed.

⁶ Ronald B. Standler, Why Attend College?, 28 Apr 2003, http://www.rbs0.com/edu.htm.

⁷ Ronald B. Standler, Computer Virus Hoaxes, 15 June 2002, http://www.rbs2.com/choax.htm.

III. Internet v. Traditional Paper

Many instructors, librarians, and academic administrators continue to believe that publishing in traditional paper media (e.g., books and articles in peer-reviewed, scholarly journals) is somehow "better" or "more credible" than posting a webpage on the Internet. This belief puts too much emphasis on peer review and ignores both the content of the work and its impact.

If an author writes a technical book in science or engineering, the publisher of that book might consider it a success if it sells 500 copies in a few years. Scholarly journals are commonly available in university libraries and almost nowhere else. Every time I visit a university library that specializes in either science, engineering, or law, I am dismayed at the small number of people (often I am the only one) there who are either reading or photocopying scholarly journals. Most of the people in a university library seem to be students who use the library as a quiet place to read a textbook, or to do homework. I am dismayed at this lack of use of scholarly journals and books in university libraries, because these libraries contain valuable resources that are *not* found on the Internet, but it does seem to be a fact that printed journals are not frequently used.

The fact that many practicing engineers and many practicing attorneys ignore articles in archival journals is deplorable, but it *is* a fact. If one wants to reach practicing engineers and practicing attorneys, then one will need to post information on the Internet, instead of publish in archival, printed journals. Additionally, because law libraries are segregated from other libraries, scholarly journals in law appear to be read *only* by law professors and some law students, even though legal issues affect everyone. Posting legal essays on the Internet makes information about the law easily available to everyone who is interested.

In contrast to this limited distribution of books and scholarly journals, articles that are posted on the Internet are easily available worldwide, any time of the day or night. Further, inclusion of webpages in the major search engines (e.g., Google, Alta Vista, All the Web, Open Directory Project, etc.) makes information much easier to find than in a traditional library. The Internet is inherently egalitarian: the Internet makes information available to *everyone*, not just the few who work in a university community.

To search for articles printed in scholarly journals before the mid-1960s, one needs to tediously search through the annual index of printed abstracts (e.g., *Physics Abstracts, Chemical Abstracts*, etc.) year by year, a process that is so tedious that it is rarely done. Since the mid-1960s, most abstracts are available in proprietary databases for convenient online searching, but such searches are expensive. For example, in April 2001 I did a search of *Electrical Engineering Abstracts*, in the Dialog collection of databases, for articles on fire or explosion hazards of surge arresters or varistors: I paid US\$ 412 to do this search and print the bibliographic data and abstract of each hit. The combination of the tediousness of manual searches of printed

abstracts and the expense of online searches of abstracts means that articles in the archival, scholarly literature are ignored more than they should be. In contrast, search engines on the Internet are free and easy-to-use.

Studies show that many articles published in science and engineering journals are never cited in any subsequent article,⁸ which means that many peer-reviewed articles in archival, scholarly journals have no impact on the development of subsequent new ideas. I believe that the lack of citations in subsequent articles shows that many peer-reviewed articles have little, or no, importance to the evolution of ideas and progress, perhaps because:

- the article was either insignificant, poorly written, or not credible (i.e., peer review was not adequate to distinguish significant articles from dross).
- people who needed the information in the article were unaware of the article's existence.
- it was too inconvenient to obtain a photocopy of the article.

Posting articles on the Internet, instead of publishing in printed journals, removes the latter two impediments to readers finding an article, thus increasing the impact of an article.

Frankly, it makes no sense for me to spend hundreds of hours of my time doing scholarly research and then writing and revising a paper to publish in an archival printed journal, when only a trivially small fraction of people who need my information will read such printed publications. Since July 1998, I have chosen to write material for my websites, instead of publish in printed journals. The large number of hits on my professional website (more than 7000 hits/week in March 2003) has convinced me that the Internet is the best way to distribute my ideas and opinions, thus giving my writing a larger impact that if it were hidden in journals in university libraries. I have no doubt that my essays on law, posted at my professional website, are read by many more people (and more practicing attorneys) than if they were published in a peer-reviewed, printed legal journal.

⁸ David P. Hamilton, "Publishing by – and for? – the Numbers," *Science*, Vol. 250, pp. 1331-32, 1990. http://www.the-scientist.com/yr1998/july/hamilton1.html and http://www.garfield.library.upenn.edu/papers/hamilton1.html;

David P. Hamilton, "Research Papers: Who's Uncited Now?," *Science*, Vol. 251, p. 25, 4 Jan 1991. http://www.the-scientist.com/yr1998/july/hamilton2.html and http://garfield.library.upenn.edu/papers/hamilton2.html;

David A. Pendlebury, Letter to Editor, "Science, Citation, and Funding," *Science*, Vol. 251, p. 1410-11, 22 March 1991 http://www.the-scientist.com/yr1998/july/pendlebury.html and http://garfield.library.upenn.edu/papers/pendlebury.html;

Eugene Garfield, "Commentary: I Had a Dream ... about Uncitedness," *The Scientist*, Vol. 12, p. 10, 6 July 1998. http://www.garfield.library.upenn.edu/commentaries/tsv12(14)p10y19980706.pdf and http://www.the-scientist.com/yr1998/july/comm 980706.html

peer review on the Internet

Acceptance of work by colleagues is an important part of professional life, which is perhaps one reason that many academic administrators continue to insist that the only appropriate place for a professor to publish is in a peer-reviewed archival journal or book. While it is true that one can post anything on the Internet, it is also true that there are analogs of peer review in cyberspace, for example:

- links to a webpage by a professor or recognized professional.
- inclusion of a webpage in required, or recommended, reading for a college class.
- links to a webpage at a website maintained by a recognized professional society.

The fact that there are websites operated by either crazy or poorly educated people is no reason to avoid the Internet, just as the existence of comic books and pulp fiction is no reason to avoid scholarly books.

Internet can be fresher than printed archival media

Books and scholarly journals are frozen in time: once printed, the material can never be changed. In contrast, materials posted on the Internet can be revised at any time, to include new information, to indicate a change in an author's opinion, or to explain something more clearly.

link rot

The most significant disadvantage of the Internet is the lack of a stable URL for a document. It is common for webmasters or authors to remove documents and to give no hint for the new location of the document. Link rot is a particular problem when the author of a article cites a document that is available only on the Internet. When that document vanishes, so does the support for the author's article.

Two professors at the University of Nebraska at Lincoln developed three online biochemistry classes, with 515 links. When they found that they were spending about four hours/month checking and revising links on their pages, they did some research and found that the half-life of their links was only 55 months. The current version of their report on link rot is posted at: http://www-class.unl.edu/biochem/url/broken_links.html.

There are several ways in which link rot occurs:

- a webmaster reorganizes the directory structure on a webserver, and moves existing webpages to different subdirectories.
- someone who posts at their employer's website changes jobs, and their former manager removes all of their old webpages.
- an author decides to remove his/her work. (Why? Hard disk space is cheap. Old webpages should remain on the webserver for continuity of links to those URLs. If there is an error in

the old webpages, the error should be fixed, instead of removing the entire webpage. If the author is no longer interested in maintaining his/her webpage, perhaps a note might be added to the top of the webpage stating that the information is no longer current, but that the webpage remains on the Internet to prevent link rot.)

These causes of link rot are *not* inherent in the Internet, but are the result of bad choices by people. People interested in preventing link rot can purchase their own domain name, so they can have a permanent address in cyberspace, even when they move to a different geographical location or change employer.

Perhaps the library analog of link rot is a citation to an out-of-print book that is not found in any local library, or a citation to an obscure journal that is not found in any local library. Such difficult-to-find items can sometimes be retrieved through interlibrary loan service of a university library. Alternatively, a commercial literature retrieval service can make a photocopy of an article in an obscure journal, typically in exchange for a fee of approximately US\$ 20.

citing material posted on the Internet

The Modern Language Association ⁹ (MLA) and Columbia University Press¹⁰ have promulgated citation styles for material on the Internet, which shows that such material *can* be cited in scholarly works. The MLA encloses the URL in angular brackets, e.g., http://www.rbs0.com/tw.htm. The MLA's angular brackets seem superfluous to me, as everyone knows that a phrase that begins with "http://www." is a URL.

When citing material on the Internet, I suggest simply following the same bibliographic style as for a book(let), omitting the name of the publisher and the city of publication, and appending the URL at the end of the citation. The date in bibliographic data is normally the date of last revision of the webpage. For citations or bibliographies that will be posted on the Internet, the URL should be formatted as a hypertext link, to permit the reader to easily access the reference. I put a space between the last character in the URL and the period at the end of the citation, to clearly distinguish the URL from the period in the citation. As an example, this webpage might be cited in a footnote as:

⁹ The official version is posted at http://www.mla.org/, then choose "MLA Style". then choose "Frequent Asked Questions about MLA Style", then choose "How do I document sources from the World Wide Web"

¹⁰ Janice R. Walker and Todd Taylor, *Columbia Guide to Online Style*, 18 Nov 2002, http://www.columbia.edu/cu/cup/cgos/idx_basic.html .

¹¹ If the date of last revision can not be determined, then use the (most recent) year in the copyright notice. If neither date can be found, then one could include "accessed [date]", which the MLA format always requires.

Ronald B. Standler, *Evaluating Credibility of Information on the Internet*, 6 May 2003, http://www.rbs0.com/credible.pdf .

Some bibliographic styles list the total number of pages in a book(let). Such information is inappropriate for a document formatted in HyperText Markup Language (e.g., a URL that ends in either .htm or .html), because the length of the printed document in HTML depends on the type size, font, and paper size that is chosen in the reader's webbrowser. However, a document formatted in Adobe Portable Document Format (e.g., a URL that ends in .pdf) does have a fixed number of pages that could be cited in a bibliography.

Citations in text or footnotes to published material commonly include a specific page number. Again, mention of a page number is meaningless for documents that are in HTML. While page numbers are meaningful in PDF, one might choose to ignore them, because they may change if the author revises the PDF document in the future.

Of course, pupils and students should always follow the citation style that their teacher or professor suggests, instead of relying on my personal preferences that are expressed here.

Why not post peer-reviewed journals on the Internet?

Currently authors need to choose between (1) publishing in a prestigious, peer-reviewed journal published on paper that few people read and (2) posting information on the Internet, where the information is freely available to everyone. However, it is certainly possible to put peer-reviewed, scholarly journals on the Internet, where they could be freely available to everyone.

There are several impediments to posting scholarly journals on the Internet:

- Some traditional journals are published by for-profit corporations, who make a profit selling subscriptions to their journals. For example, in April 2003, a subscription to *Nature* costs a library US\$ 920/year, and costs an individual person US\$ 159/year. There is no way that Macmillan Publishers Ltd., the commercial company who publishes *Nature*, can afford to give free access to *Nature*.
- 2. Nonprofit professional societies who publish scholarly journals derive significant income from the sales of subscriptions to libraries, and to their members. There is no way that these professional societies can afford to give free access to their journals, without either increasing membership dues or decreasing other services.
- 3. If scholarly journals on paper become obsolete, librarians will become curators of museums of pre-Internet information. Librarians have a bias toward continuing to manage the only local source of information. It is *not* a criticism of librarians to note that *no one* likes to be made less important.

However, there are some significant benefits to online publishing. If authors were required to submit material in Adobe Portable Document Format (PDF) with embedded fonts, then the online journal could avoid the following substantial expenses:

- professional typesetting, with its occasional typographical errors
- printing (e.g., costs of paper, making printing plates, binding)
- postage

Journals operated by professional societies traditionally relied on *un*paid editors and *un*paid peer reviewers, so there are already no expenses for those critical services. Reduction of publishing expenses should reduce costs to users of information.

Not only can online publishing eliminate publishing expenses, but also libraries would no longer be burdened by the continually increasing floor space to store these heavy printed journals. Searching online sources for occurrence of certain words or phrases is much more efficient for users than searching printed abstracts. Moreover, paper journals are more easily damaged by fire or vandalism than journals stored on a hard disk drive, for which an exact duplicate copy can be easily made.

IV. Links to Other Resources

The following webpages offer different perspectives on the problem of evaluating the credibility of sources of information. The following webpages are *not* a bibliography for my essay, but do offer alternative points of view. The dates are largely meaningless, since the following webpages are frequently revised.

- Jan Alexander & Marsha Ann Tate, *Evaluating Web Resources: Bibliography*, 23 Apr 2003, http://www2.widener.edu/Wolfgram-Memorial-Library/webevaluation/webstrbib.htm.
- Nicole J. Auer, *Bibliography on Evaluating Web Information*, 25 Apr 2003, http://www.lib.vt.edu/research/evaluate/evalbiblio.html . (Also contains links to some "example websites", so readers can test their evaluation skills.)
- Joe Barker, Evaluating Web Pages: Techniques to Apply & Questions to Ask, 6 Jan 2003, http://www.lib.berkeley.edu/TeachingLib/Guides/Internet/Evaluate.html.
- Colleen Bell, Critical Evaluation of Information Sources, 4 May 2003, http://libweb.uoregon.edu/guides/findarticles/credibility.html.
- Don E. Descy, Evaluating Internet Based Information, 1999, http://www.lme.mankato.msus.edu/class/629/Cred.html (Dr. Descy also has a collection of links to hoax webpages, so readers can test their evaluation skills.)

- Robert Harris, *Evaluating Internet Research Sources*, 17 Nov 1997, http://www.virtualsalt.com/evalu8it.htm.
- Iowa State Univ. Library, *Evaluating Information on the Web*, 29 Oct 2002, http://www.lib.iastate.edu/commons/resources/evaluation/.
- Elizabeth E. Kirk, *Evaluating Information Found on the Internet*, 5 June 2002, http://www.library.jhu.edu/elp/useit/evaluate/.
- M. Phillips, Critical Evaluation of Resources, March 2001, http://www.lib.berkeley.edu/TeachingLib/Guides/Evaluation.html.
- Ray Schroeder, Evaluating Online Resources Notebook, 1 Apr 2003, http://www.uis.edu/~schroede/valid.htm.
- Alastair Smith, Evaluation of Information Sources, 5 May 2003, http://www2.vuw.ac.nz/staff/alastair_smith/evaln/evaln.htm. (A very large collection of links to essays on evaluating sources of information.)

V. Conclusion

Readers can, and should, make their own determination of credibility of information, instead of blindly relying on peer review. And, if an article is good, it is good because of the author's skill, insight, diligence, and creativity, *not* because some expert or authority endorsed it. A conclusion is only as good as the reasons for believing that conclusion. A conclusion stated without reasons, or with *in*adequate reasons, is worthless. It is the reader's task to decide if the reasons are adequate and convincing.

This document is at http://www.rbs0.com/credible.pdf revised 6 May 2003

return to my homepage at http://www.rbs0.com/