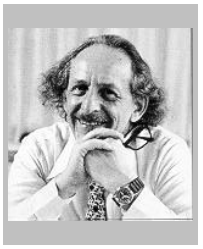
Essays of an Information Scientist: Science Reviews, Journalism Inventiveness and Other Essays, Vol:14, p.255, 1991



How to Avoid Spreading Error
Scientists Must Search for Corrections

Reprinted from *THE SCIENTIST* ® 1(19): 9, 10 August 1987.

In 1954 Conway Zirkle reviewed the fascinating history of the patterns and context of citations to the falsified scientific experiments published by Viennese zoologist Paul Kammerer. Using two types of salamanders and the male of the midwife toad, Kammerer claimed in the 1920s to have shown that acquired characteristics were inherited. But, as Zirkle recounts, "the acquired characters ... turned out to be india ink." (*Science* Vol. 120, 1954. p. 189).

The truth about Kammerer's work was known and revealed in print as early as 1926. Nonetheless, later researchers, many of whom were unaware of the nature of Kammerer's evidence, repeatedly cited his publications and took the conclusions he presented at face value. Zirkle warned readers that the consequences of fraud in scientific publication may be far-reaching. "A single knowing misrepresentation may start a chain reaction," he wrote (p. 190).

Commenting on Zirkle's paper shortly thereafter, P. Thomasson and J.C. Stanley agreed that "the uncritical citation of disputed data by a writer, whether it be deliberate or not, is a serious matter. Of course, knowingly propagandizing unsubstantiated claims is particularly abhorrent, but just as many naive students may be swayed by unfounded assertions presented by a writer who is unaware of the criticisms. Buried in scholarly journals, critical notes are increasingly likely to be overlooked with the passage of time, while the studies to which they pertain, having been reported more widely, are apt to be rediscovered" (*Science* Vol. 121, 1955. p. 610).

That last sentence is as relevant today as it was over three decades ago. And in the context of the ongoing debate about the extent and consequences of falsified data appearing in the journal literature, there seems to be a heightened awareness of the significant role that correction or retraction notices play in scientific communication. Such notices (when not trumpeted in the press as evidence of fraud, most of which are no such thing) are important devices to ensure that science progresses on firm ground.

Moreover, they are doubly important these days as verification of published results through the duplication of experiments becomes a rare activity. Who has time or

resources to repeat the experiments of colleagues? With the specter of information overload looming over the bench of almost every scientist, correction and retraction notices are apt to be missed even more often than they were in 1954.

So I welcomed Donald A.B. Lindberg's recent announcement that the U.S. National Library of Medicine's *MEDLINE* system has since 1984 annotated some 37 items in its files to notify users that the findings reported in these publications were later retracted (*The Scientist*, June 15, 1987. p. 13). I commend this innovation and heartily agree with Lindberg's statement elsewhere that "the general reader of the published scientific literature must be able to learn that an article he or she has read has subsequently been retracted" (*Science* Vol. 235, 1987. p. 1308).

I wonder, however, whether scientists and librarians realize how many correction notices have been indexed in ISI's *Science Citation Index*? Of the 10 million journal items indexed in the *SCI* since its inception, over 50,000 were coded as explicit corrections. In 1986 alone, we indexed more than 4,500 such items. These vary from corrections of simple typographical errors to retractions of and outright apologies for "bad" data or data that cannot be verified.

Correction or retraction notices invariably reference an original publication. Such a citation links the unfounded or erroneous data and conclusions in the original publication to the correction or retraction of those data and conclusions in a subsequent publication. Forward cross-referencing is the heart and soul of citation indexing. A citation index has the great advantage over other types of reference tools of permitting both retrospective and prospective searching. By looking under the entries for a particular publication in the *SCI*, you will learn what subsequent publications have cited the earlier work and among these entries you may find a correction notice.

But there is an even larger point I wish to make: a citation index, by reflecting in aggregate terms the entire scientific literature, is the ultimate "correction index."

It is often said that scientific knowledge is cumulative and scientific research is self-correcting. In the former instance, new knowledge rests upon previous knowledge. In the latter, untenable conclusions that provide the impetus for extending work along certain lines will in the long-term be uncovered as further experiments fail to provide expected results. So, aside from explicit corrections, there is a constant and inexorable reassessing and refining of earlier work to improve upon current understanding. All of this is reflected in the scientific literature and formalized in citations. While it is useful to learn about overt corrections and retractions, it is equally important to be aware of more subtle and possibly more significant "corrections," which are the very progress of science. That some earlier paper was wrong or has been superseded is the significant message.

It is not only prudent to teach students to look for "corrections" both large and small every time they publish. It is also ethically important and essential to the claim of originality. As countless published apologies have demonstrated, duplication occurs because authors and referees fail to determine whether or not the work cited in a paper has been corrected or superseded. In the courtroom, failure to check references (precedents) can prove disastrous to an attorney. The consequences in scientific research are obvious but no less damaging.

The unwillingness of abstracting services to cover all errata is understandable. An index entry to one of these may be longer than the erratum itself! Indexing errata is expensive and only occasionally does one receive any gratitude for helping to clean up the mistakes of others. Most readers pay little attention to errata and librarians find them costly if not impossible to insert at the appropriate page of a journal. For a while it may become fashionable to publish "retractions" of various kinds, but the incentive to continue doing so will be missing unless their potential impact is communicated by editors, referees and teachers.

Scientists ought to develop the habit of looking for corrections or retractions of works they cite in their publications. It is a habit they ought to instill in their students, too. Error in important work will eventually come to light, but the sooner uncovered the better for all. If researchers learn to search the journal literature for corrections routinely, there would be greater efficiency in the use of the time and resources of scientists and publishers alike.

In this issue of *The Scientist* (pp. 10-11), Stanford University president and biologist Donald Kennedy emphasizes the traditions of confirmation and peer review as the chief means by which the scientific community keeps the journal literature as free as it can be of "noise" such as inadvertent error. To these honorable traditions I would like to see added habitual and circumspect searching of the scientific literature before scientists go to press. ■