

An Investigation of the Validity of Bibliographic Citations

Robert N. Broadus

School of Library Science, University of North Carolina, Chapel Hill, NC
27514

Edward O. Wilson, in his famous work, *Sociobiology, The New Synthesis* [9], makes reference to a pair of articles by W. D. Hamilton, but misquotes the articles' title. No less than 148 later papers make reference to both Wilson's book and Hamilton's articles, by title. Thus, there is provided an opportunity to test the charge, made by some critics, that writers frequently lift their bibliographic references from other publications without consulting the original sources. Although 23% of these citing papers made the same error as did Wilson, a further perusal of the evidence raises considerable doubt as to whether fraudulent use was intended.

Introduction

The use of citation counts for various purposes in library and information science, as well as in the history and sociology of science and scholarship [1] naturally has raised fundamental questions about the nature, purposes, and validity of bibliographic references themselves. Differences in judgment are sharp indeed.

One viewpoint, stated impressively by J. M. Ziman, a British authority on science and its history, holds that citations "vouch for the authority and relevance of the statements that they are called upon to support. . . One relies on the citations to show. . . [the citing paper's] place in the whole scientific structure, just as one relies on a man's kinship affiliations to show his place in his tribe" [2, p. 318]. Hodges, in a dissertation produced at the University of California at Berkeley, summed up,

A survey of more than eighty papers, analyses of many writings about writing, and interviews with forty-two leading scholars reveal enough regularity in apparent or stated 'reasons for citing' to warrant the conclusion that there is a generally accepted 'code' governing citation practice. Those interviewed are almost unanimous in believing that the references called for by such a code are not only sound but useful. [3, p. 2]

Almost all the people interviewed by Hodges said that references provided good leads to other publications [3, p. 192] and many said that looking at a paper's list of references gave them an idea as to whether the paper itself was worth reading [3, p. 195].

On the other hand, there are strong opinions that some bibliographic references are made for purposes other than giving proper credit, while really important papers may have been overlooked by authors in their searches of literature. Kaplan, in 1965, asked a set of leading questions: "How often are the works of others cited without having been read carefully? How often are citations simply lifted from the bibliography in someone else's work without either reading or giving credit to the man who did the original search of the literature?" [4, p. 181].

Some critics have gone beyond the questioning stage and have charged that authors are fraudulent in making references to other publications. May declared flatly, "The author selects citations to serve his scientific, political, and personal goals and not to describe his intellectual ancestry." One consequence, according to May, is carelessness; another is "plagiarism of other people's citations without having actually used them" [5, p. 890]. Still another bitter attack on citation behavior was made by Davies:

Or we can tell you how extraordinarily well read we are, by quoting fifteen peripherally relevant articles in fifteen languages. Most of these articles will have been accumulated by accretion. Someone else referred to them and we added them to our list because they looked nice. This illustrates a fundamental law of reference giving—it is quite unnecessary to have read or even seen the reference yourself before quoting it. [6]

While the charge of carelessness and inaccuracy is serious enough, that is a fault that can be detected without great difficulty. There is indeed considerable evidence that, in general, the accusation is justified. Most of us have been frustrated by references that give wrong

Received May 17, 1982; accepted August 9, 1982

© 1983 by John Wiley & Sons, Inc.

volume numbers or paging. Key and Roland checked the papers accepted for publication by a medical journal over a 13-month period, and found that of the 1867 references attached to these papers, only 40 percent were free of errors (corrections were made, however, before the articles were published) [7]. Boyce and Banning examined articles actually published and found that in one journal, 13.6% of the references contained some error; for their second journal, the rate was 10.7% [8].

Errors of this kind affect citation counts in different ways, depending on the purpose of the particular study. Rather few, for instance, involved title of cited journal; therefore, the practice of ranking journals in the order of numbers of citations received would be affected only slightly. The fact that in a reference the volume number or paging was wrong would affect the results of the very few citation studies. Incorrect name of author, however, would have serious consequences for those studies using this method to assess the importance or eminence of the individual whose works are cited, and mistakes in date of publication would generally lead to wrong conclusions about ages of materials cited.

The accusation of outright deception in the use of bibliographic references is far more serious, and is more difficult to prove (as well as to defend against), for it has to do with the private intentions of those who purport to discover and announce "truth." The controversy is of utmost importance not only for citation studies, but for the whole structure of scientific and scholarly investigation.

How can we obtain evidence on the integrity of citation behavior? It is even more difficult than detecting falsification of laboratory data. Putting the question to a sizable sample of authors who make bibliographic references is not feasible, and even if it could be done, we should accept only those answers that had been sanctioned by polygraph tests. Our only avenues are indirect ones.

The present study, by use of one peculiar situation, seeks to furnish some evidence on the matter.* In 1975 Prof. Edward O. Wilson published a book that was to become famous, and to be cited by a great many other authors. In his own huge list of references (about 2500 entries), Wilson included the following:

Hamilton, W. D. 1964. The genetical theory of social behaviour, I, II. *Journal of Theoretical Biology*, 7(1):1-52 [9].

The title of the article is incorrect;† the actual words should have been, "The genetical evolution . . ." (the same error is repeated in the abridged edition of *Sociobiology*, published in 1980). Perhaps the word "theory" was substituted because of the title of the journal.

*I have included rather more detail than usual, because my conclusions are a matter of personal judgment, and these details may help readers to make their own assessments.

†Called to my attention by Jane Morley, a student in the School of Library Science, The University of North Carolina at Chapel Hill.

A check of *Ulrich's International Periodicals Directory* indicates that *The Journal of Theoretical Biology*, begun in 1961, had by 1975 a circulation of some 1600 [10, p. 196]. In view of this relatively small circulation and the great popularity of Wilson's book, it seems reasonable to suppose that some of those making reference to Hamilton's paper(s) first encountered them through Wilson.* An automated search of the databases of the *Science Citation Index* and the *Social Science Citation Index*, together with a manual search of the *Arts and Humanities Citation Index* (the topic is, of course, interesting to many fields of study), revealed that in the years 1976 through the time of the searches, December 1981-January 1982, 180 published papers made reference to both Wilson's book and to one or both of the Hamilton articles. Did any of these later writers make the same mistake as Wilson? If so, what, if any, is the evidence that they simply added the reference to their own papers without consulting the original article(s)? (There is no objection at all, presumably, to a writer's going from Wilson's reference to the original paper cited; then in turn making reference to that paper after having examined it directly. Kaplan [4, p. 181] does imply by the last part of his second question that the author who suggested a reference should be given credit, but (1) is there any obligation to credit a subject bibliography or periodical index for such help? (2) Should every reference made include a chain of credit going back to the first citer?)

Any conclusions here obviously must be tentative. It is possible that a later writer did use fraudulently the bibliographic reference in Wilson and then by good fortune had the actual wording corrected later perhaps by a zealous research assistant making a final check of the footnotes. On the other hand, even if a writer did include the reference exactly as Wilson worded it, that would not be sure proof that such critics as May and Davies are right. Surely, it is not unusual for a person to jot down a bibliographic reference while reading a book or article, then to look up the cited article in a journal but fail to note the error in the original reference. Despite these caveats, it may be said that our situation does provide some evidence.

Findings

Table 1 gives a breakdown of the 180 papers that co-referenced Wilson's book and the article(s) by Hamilton. Of the 148 that included title of article in their references to Hamilton, 114 used the correct word "evolution," and 34, or 23%, the erroneous word "theory." Papers with the defective reference originated in some of the leading institutions and appeared in some of the most prestigious journals.

*Although some may have picked it up from a reprint in Williams, G. C., Ed. *Group Behavior*. Chicago: Aldine-Atherton; 1971.

TABLE 1. Papers coreferencing Wilson's book and Hamilton's article(s), according to ISI Indexes.

	Paper Not located	No title for article(s)	"Evolution" (Correct)	"Theory" (Incorrect)	Total
All 3 Indexes	0	0	0	1	1
<u>SCI</u> and <u>SSCI</u> , but not <u>A&HCI</u>	2	2	21	5	30
<u>A&HCI</u> only	3	0	10	0	13
<u>SCI</u> only	0	17	36	11	64
<u>SSCI</u> only	4	4	47	17	72
Totals	9	23	114	34	180

These data seem to support the suspicions raised by Kaplan, and even the direct charges of May and Davies, against the assertions of Ziman and the findings of Hodges. Before conclusions are nailed down, however, consideration of other evidence may modify our views somewhat.

The index to Wilson's book reveals a total of 20 references to seven papers by Hamilton. Four of these references are to the article(s) in question. It might be supposed that one bad turn would call for another, and that an author having swiped one reference under false pretenses would face the temptation to lift more. This supposition could be checked, for those papers located through the *Social Science Citation Index*, because these printouts included the complete list of references made by each paper. Twenty-three papers uncovered through SSCI used "theory" and coreferenced Wilson and Hamilton. Only two of them also made reference to another of the Hamilton papers cited by Wilson. This fact does not imply that wholesale borrowing from Wilson has widespread; in fact it suggests, if anything, the contrary.

Still dealing with these 23 papers: Their other references were checked against Wilson's bibliography to determine the amount of duplicate citing. In all, the papers made 1281 references (not counting the two that each

made to Wilson's book and the Hamilton papers in question). Of these 1281 references, 269, or 21%, were referenced by Wilson also. In view of the large number of items in Wilson's bibliography, this amount of duplication does not seem excessive, (It is possible, of course, if the worst be contemplated, that some writers lifted references from Wilson and from other writers as well to make up their own bibliographies without consulting any of the original documents, but that seems to carry things too far.)

References made by some of the individual papers give a somewhat different picture. In 11 of these papers, the list of references shows a duplication rate with Wilson's list of 20% or more, but in four papers the actual number of duplicates is small: 4, 4, 7, and 9, respectively. One paper, however, coreferences 79 of its 248 items with Wilson; another 14 of 37. Even more suspicious is a third paper that makes 76 references, 33 of them in common with Wilson, for the highest percentage of duplication in the group, 43.4%. In addition, the latter paper lists the paging of Hamilton's articles as "12-45"—strange in that the paging of the two articles actually is 1-16 and 17-52. Questions are raised further by the fact that in Wilson's bibliography, the next item after the Hamilton reference in question is another item by Hamilton, bearing the page numbers

12-45. Nevertheless, a comparison of the texts of these three articles with the text of Wilson does not give the impression that plagiarism of the references took place.

Conclusion

Unfortunately, we still have some important but unknown factors in the way of making an estimate of how many writers, if any, lifted the reference from Wilson without consulting the original source. My own guess is that of the 23 discussed above, one writer may have done so, and that with another there is room for a measure of doubt. Even if 4-8% of authors engage in this kind of unethical practice, that presents an important problem for scholars performing certain kinds of citation studies. It should be noted, however, that where such studies are used to provide information on which to base decisions about collection development, the difficulty is a little less serious, because of the fact that research people so often look for a work in the library having located a bibliographic reference to it [11, p. 302-305, 328].* In such cases, a user may seek a given work even though the footnoting be fraudulent.

Cronin is right in saying that perhaps it is "asking for

*But note also a study in which 52% of the subjects were undergraduate students. Here, footnotes were less of a stimulus than was the group labeled "Index, abstract, or bibliography." Golden, G. A.; Golden, S. U.; Lenzini, R. T. "Patron approaches to serials: A user study." *College & Research Libraries*, 43(1):22-30; 1982.

the moon" if we expect "a full-blooded attempt to explore the phenomenology of citations" [12, p. 311], but the evidence shown here should be of some help.

References

1. Cozzens, S.; Ivory, M.; LaPorte, J.; Small, H.; Stanley, J. *Citation Analysis: An Annotated Bibliography*, 2nd ed., Philadelphia, PA: Institute for Scientific Information; 1978.
2. Ziman, J. M. "Information, communication, knowledge." *Nature*, 224(5217):318-324; 1969.
3. Hodges, T. L. "Citation indexing: Its potential for bibliographical control." Ph. D. dissertation. University of California, Berkeley, CA, 1972, 642 pp.
4. Kaplan, N. "The norms of citation behavior: Prolegomena to the footnote." *American Documentation*, 16(3):179-184; 1965.
5. May, K. O. "Abuses of citation indexing." *Science*, 156(3777): 890, 892; 1967.
6. Davies, D. "Citation idiosyncracies." *Nature*, 228(5278):1356; 1970.
7. Key, J. D.; Roland, C. G. "Reference accuracy in articles accepted for publication in the *Archives of Physical Medicine and Rehabilitation*." *Archives of Physical Medicine and Rehabilitation*, 58(3):136-137; 1977.
8. Boyce, B. R.; Banning, C. S. "Data accuracy in citation studies." *RQ*, 18(4):349-350; 1979.
9. Wilson, E. O. *Sociobiology, The New Synthesis*. Cambridge, MA: Belknap Press of Harvard University Press; 1975.
10. *Ulrich's International Periodicals Directory*, 16th ed., 1975/1976. New York: R. R. Bowker; 1975.
11. Broadus, R. N. "The applications of citation analyses to library collection building. In *Advances in Librarianship*. M. J. Voigt and M. H. Harris, Eds. New York: Academic; 1977, Vol. 7, pp. 299-335.
12. Cronin, B. "Some reflections on citation habits in psychology." *Journal of Information Science*, 2(6):309-311; 1980.