Essays of an Information Scientist, Vol:8, p.403-409, 1985 Current Contents, #43, p.3-9, October 28, 1985

Current Comments®

EUGENE GARFIELD
INSTITUTE FOR SCIENTIFIC INFORMATION®
3501 MARKET ST., PHILADELPHIA, PA 19104

Uses and Misuses of Citation Frequency

Number 43 October 28, 1985

About two years ago, I discussed in *Current Contents®* (*CC®*) when and how it might be relevant to use citation analysis to evaluate faculty research.[1] On August 30, 1985, a symposium on the use and abuse of citation analysis was held at the annual meeting of the British Association for the Advancement of Science at the University of Strathclyde in Glasglow, Scotland. On that occasion, ISI®'s director of research, Henry Small, and I reviewed the many reasons for evaluation methods as well as their limitations. Much of that discussion is incorporated below.

However, in January of this year my *CC* essay took on special significance. An editorial published in the *New York Times* used citation-frequency data from a table in the faculty-evaluation essay to question the need for increased funding of biomedical research.[2] The table is reprinted later in this essay (Table 1).

The table is derived from the references cited in the *Science Citation Index®* (*SCI®*) from 1975 to 1979. It demonstrates that only a small percentage of papers are frequently cited. When selecting *Citation Classics®*, we vary the citation threshold according to the field under investigation. There is a typical hyperbolic distribution when citation frequency is plotted against the number of papers cited at that frequency. This phenomenon was reaffirmed recently when we published a similar analysis for papers cited over 450 times between 1961 and 1982[3] (Figure 1). The complete data for all papers in the 1961-1982 *SCI* are included in Table 2.

It is unfortunate that the *New York Times* editorialist misinterpreted the data in Table 1, which was based on just five years of data. He stated that "only 36 percent of published scientific articles are cited two or more times in subsequent research reports. The rest—two-thirds of researchers' published output—may contribute only negligibly to the march of science."[2] It is even more unfortunate, however, that he used these data to justify proposals for cuts in biomedical research. My response to the editorial was not published by the *New York Times*, but I am including it here partly because many readers have urged me to respond. Apart from this specific inappropriate use of the data, a more general tendency to misuse such data warrants comment. By basing funding or even scholarly tenure and hiring decisions on quantitative bibliometric data, there is always the potential for making two serious mistakes: one, in believing that mere publication or citation *counting* is equivalent to citation *analysis*; and two, in believing that citation analysis, even when carefully performed by experts, is sufficient by itself to ensure objectivity.

The citation investigation of a researcher or research project can be a complex procedure. It should involve more than a glance at the *SCI*, *Social Sciences Citation Index®* (*SSCI®*), or *Arts & Humanities Citation Index*™ (*A&HCI*™) in print or online. For example, citations tell us nothing about a researcher's teaching ability, administrative talent, or other nonscholarly contributions. And they do not necessarily reflect the usefulness of research for curing disease, finding new drugs, and so on.

Apparently, some naive persons infer from our studies of highly cited papers that we are condemning less-cited papers and the work they represent as mediocre or even useless. This is doubly unfortunate because it is patently untrue and because we have never suggested such a spurious conclusion. Perhaps it is time to reiterate some oft-stated general observations. But first, my letter on the *Times* editorial.

April 22, 1985
The Editor
The New York Times
229 W. 43rd Street
New York, NY 10036

To the Editor:

In the *New York Times* January 28, 1985, editorial on biomedical research, your editorialist misused data I published about a year and a half ago. In an article about faculty evaluation, published in *Current Contents*® (*CC*®), November 7, 1983, I provided a table giving frequency distributions from the *Science Citation Index*® (*SCI*®). It is true that of the millions of items cited *in a five-year period*, 36 percent were cited two or more times. But this does not justify the conclusion that "only 36 percent of published scientific articles are cited two or more times."

The purpose of this table was to demonstrate the ranked-frequency distribution of citations to papers cited in a given period for all branches of science. We use such data to estimate the number of published *Citation Classics*® that we can obtain at each frequency level. (By definition, a *Citation Classic* is a paper or book that has been highly cited in its field.) However, your editorial distorted the meaning of the statistical data in order to justify a specious policy recommendation to curtail support of *biomedical* research.

The *SCI* indexes all high-impact scientific journals, including biomedical journals. Indeed, my article also included a table that showed the five-year impact of the articles in biomedical publications like the *Journal of Experimental Medicine*, the *New England Journal of Medicine*, and so on—those journals that typically publish articles by NIH-sponsored researchers. The five-year citation impact of articles published in such journals ranges from 5 to 70. As reported by Helen Gee and Charles U. Lowe (*New York Times*, p. A22, February 20, 1985), the seven-year impact of NIH-sponsored articles published in 1973 was about 15. The essential point is that the lifetime citation expectancy of peer-reviewed biomedical-journal articles published in the past decade would be almost double the range cited above. These figures would be quite different from those for the millions of items of all kinds published over the entire history of science. Clearly, a large fraction of the citations to those papers occurred in earlier periods. Their average frequencies would depend upon the size of the literature in those years.

All of the above factual information notwithstanding, what is its relevance to the issue of funding more or less research? There will always be a wide range in the citation-frequency distribution for a large population of papers. This alone cannot tell us anything about the average quality of research. No matter how small or large the budgets for research, there will always be a group of "well-cited" papers. By definition the vast majority will be cited less frequently. An unknown variable proportion will never be cited at all for a variety of reasons. The stress to publish or perish guarantees the publication of many preliminary communications that will rarely be cited. Most of these are superseded by reviews or other archival papers that will continue to be cited until they too are superseded by other more relevant papers.

By comparison with older papers, the average quality of modern scientific papers is quite high. They are often packed with information generated with the help of computerized instrumentation. The output of modern science is prodigious, and we do need better methods of evaluating it. Statistical data can help in such evaluations. However, there is always the danger of spurious conclusions.

Kenneth Warren (*New York Times*, February 5, 1985) indicates that the number of excellent proposals that must be turned down is very high. It is ironic that you would use evidence of the very success of modern medical research to limit even more of that success. The number of unconquered diseases is large, not to mention the related problems of population control, famine, and so on, and solutions rest upon increasing our fund of knowledge. If biotechnology, and biomedicine in general, is on the verge of great breakthroughs, this is precisely the time to increase our support of research and the educational institutions that nurture it.

Since most *CC* readers have some grasp of statistical methods and probability theory—something not to be expected of journalists—Figure 1 is not surprising to them. This is the kind of curve that we have learned to expect for such a large population of papers. Whether it is the total output of worldwide research or the papers published in a particular field, there will usually be a hyperbolic distribution. This was observed in the very first experimental citation index we created in the 1950s. And Derek J. de Solla Price formalized this notion in a classic paper on networks of citations[4] and later in his model of what Robert K. Merton earlier identified as the process of the cumulative advantage, or the "success-breeds-success" process.[5,6] In any large population of scientists, most produce a few papers while a decreasingly smaller number produce increasingly larger numbers of papers. Similarly, most papers are cited a few times, if at all, while an increasingly smaller number are cited at ever-higher frequencies.

Whenever we have published lists of most-cited authors and papers, I have always been concerned about "who shall occupy the 41st chair."[7]

Every year, more scientists are eligible for Nobel prizes than can win them. This means that there has always been an accumulation of "uncrowned" laureates who are the peers of prizewinners in every sense except that of having the award. These scientists, like the "immortals" who happened not to have been included among the cohorts of forty in the French Academy, may be said to occupy the "41st chair" in science.[6,8]

Space restrictions in *CC* require that we select some arbitrary cutoff points. This inevitably means that some highly cited papers or authors of equivalent quality to the ones chosen will not be listed. Does this mean that we should avoid mentioning any highly cited authors and papers at all? I think not. In a democratic society we aspire to select our best representatives. In the academies of sciences we hope that the best are selected sooner or later. The spectrum of talent is quite broad. A fixed threshold or cutoff point means that not all of the "best" can be admitted.

Selections of interesting papers through citation analysis or by peer-review committees will never be perfect. The use of citation frequency as one *indicator* of impact is legitimate. It provides a measure of the past or present influence of earlier work. It does not necessarily guarantee that the impact, in the future, will be the same. In citation analysis and peer review, past performance is only a guide to, not a guarantee of, the future. Furthermore, one must find out not only how often, but *why* someone is being cited.[9-11] One must also try to determine the citation-distribution rates for the fields involved. The number of citations received by different types of

Table 1: Citations received by items cited one or more times in the 1975-1979 cumulated *SCI®*. The table includes an unspecified number of duplicates cited in variant forms. A = total citations, 1975-1979. B = cumulative number of items. C = cumulative percent of items.

A	B	C
≥ 1	10,641,000	100.00%
≥ 2	3,874,000	36.00
≥ 5	1,531,000	14.00
≥ 10	670,000	6.30
≥ 17	313,000	3.00
≥ 25	155,000	1.50
≥ 51	44,000	.40
≥ 101	10,500	.10

papers and the length of time for these papers to reach their peak vary enormously from field to field. However, correctly applied, citation analysis can clarify the selection procedure and increase its objectivity.

In contrast to the evaluation of individuals by peer review or publication and citation frequency, the use of citation data becomes increasingly useful and valid as the size of sample populations increases. Publication and citation data from the *SCI* have been used regularly by the National Science Founda-

Figure 1: Citation distribution for all items cited at least 450 times in the *SCI®*, 1961-1982. Items were tabulated for ranges of citations. The midpoint of each citation range is plotted. The horizontal bars through the points indicate the actual ranges. The table of values shows the actual ranges of citation values (X) and of number of items (Y).

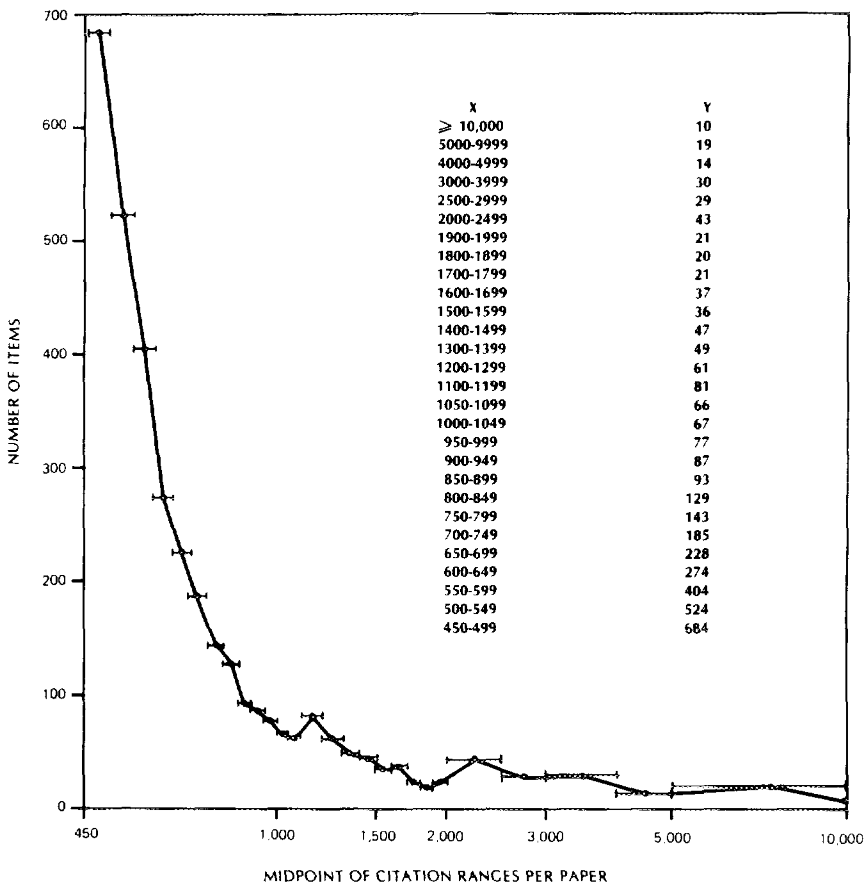

Table 2: Citations received by items cited one or more times in the *SCI*®, 1961-1982. The table includes an unspecified number of duplicates cited in variant forms. A = total citations, 1961-1982. B = cumulative number of items. C = cumulative percent of items.

A	B	C
⩾ 1	23,139,000	100.00%
⩾ 2	9,731,000	42.05
⩾ 5	4,310,000	18.63
⩾ 10	2,286,000	9.88
⩾ 17	1,283,000	5.55
⩾ 25	736,000	3.40
⩾ 51	287,000	1.15
⩾ 101	77,000	.33

tion in its biennial *Science Indicators* reports. Six of these have been published so far—the most recent one in 1983.[12] Aggregated data on countries, regions, specialties, or even departments are of intrinsic interest to a wide number of readers. And *SCI* data have been used to confirm questionnaire-based rankings of academic departments published by the National Academy of Sciences. This series includes four volumes: biological sciences, engineering, mathematical and physical sciences, and social and behavioral sciences.[13]

Recently, we published a list of *Citation Classics*[3] that included one of the most-discussed papers of all time—the 1953 Watson-Crick paper on the double-helix structure of DNA.[14] Yet this paper has been explicitly cited in the *SCI* "only" about 850 times. Many classic papers receive thousands of explicit citations in their lifetime. There is, I believe, an important message here for those who jump to conclusions without fully studying or understanding the data.

The Watson-Crick case is an important reminder that citation behavior is governed by many complex factors. Any citation analysis should take into account the publishing and citing conventions of the field, the reasons that the papers are being cited, and, of course, adjustments for self-citations. Determining the reasons for citation requires examination of the citing papers in two ways: *context*[15] and *content*[16] analyses. Context analysis shows which aspect of the cited work was actually mentioned by other publishing scientists; content analysis tells whether these scientists were criticizing or supporting that aspect of the cited paper.

In any case, most papers have a limited lifetime citation expectancy. *Sooner or later* they are subject to "obliteration by incorporation":[17-19] they are no longer cited because their substance has been absorbed by current knowledge. This point needs to be repeated regularly. One reason citations to certain papers are not obliterated lies in their potential for controversial discussion. Clearly, Darwin's work falls in that category. Less controversial work, such as Watson and Crick's, may be more susceptible to obliteration. Although difficult to prove categorically, this hypothesis warrants careful investigation. Undoubtedly, future computer-aided full-text analyses will illuminate the ways in which both natural-language and cited references symbolize evolving concepts in science. The advent of universal optical character recognition machines makes this prospect realistic in the next decade.

In this same connection, however, we should not forget about what are properly called uncited or hardly cited classics. We should take a much closer look at this phenomenon. It has been frequently encountered in the over 2,000 published commentaries by the authors of identified *Citation Classics*. Those commentaries often cite other papers, by the same authors as well as others, that are best described as uncited classics. That is, either by delayed recognition, incorporation, or other factors, these papers received relatively few explicit citations. Whether they are victims of cryptomnesia, an unconscious plagiarism in which creative ideas expressed as new are actually unrecalled memories of another's idea,[20-22] or were superseded for other reasons remains to be seen. But one can usually find examples of bibliographic amnesia in almost every field. The very existence of *Citation Classics* calls for finding ways of systematically identifying uncited or hardly cited classics. This is an important facet of the more general phenomenon of uncitedness.

Some uncited authors are upset that papers by others often supersede their own earlier papers. However, there may be complex social or other factors that, at a particular time, affect the adoption of a particular paper as the symbol of its field. When ISI identifies a particular review or other well-cited paper in a field as a *Citation Classic*, we may sometimes inadvertently do an injustice to another more nearly primordial paper. However, our basic goal is to recognize and give greater visibility to the discovery, method, or field identified. Sometimes this means a review paper will be called out, while a less-cited primordial "classic" paper is not. Delayed recognition of papers reporting significant research is not uncommon. Some remain virtually uncited for years due to the phenomenon of "premature discovery" or because they can't yet be related to other current research. Or they may be awaiting methodological breakthroughs that will allow them to be fully exploited. Others simply remain obscurely expressed. Talented reviewers may give such discoveries the added clarity and significance to call them to the attention of a previously unaware community of readers. Lack of explicit recognition is a risk all scientists face in the era of big science. It is worth noting that many authors of *Citation Classics* have never received any other formal recognition of their impact.

It is arguable whether the Watson-Crick elucidation of the structure of DNA was more or less "significant" than numerous other discoveries before and since. Perhaps the fact that it is only one in a thousand papers that have been cited as much tells us something important about the way scientific knowledge cumulates. It is precisely because it is difficult to assign numeric values to this or that discovery or breakthrough that we should not confuse intrinsic value with the "intellectual influence" reflected in citation counts.

The esoteric nature of much modern research seems to be the antithesis of any analysis that, at first, appears to be a purely quantitative, algorithmic, and uncritical technique. But proper citation analysis requires dedicated commitment. It is one important methodology for identifying and evaluating creativity and excellence. Citation analysis is not a substitute or shortcut for critical thinking; it is, instead, a point of departure for those willing to explore the avenues to thorough evaluation. Although peer-review and citation analyses are highly correlated, there is enough variance to warrant using both procedures in tandem.

If the *New York Times* wants to recommend modifications in the expenditure of public funds, it has every right to do so. But it has an obligation to use published data responsibly. It is hard to believe that the editorial in question, like most others involving science and technology, could not have benefited from peer review.

Every year at hundreds of universities and colleges and other research centers throughout the world, academic administrators, corporate managers, and research directors go through the painful process of evaluating researchers and research programs for employment, promotion, funding, or awards. In most cases, decisions are based on peer review—usually a committee's appraisal of each researcher's qualifications, demonstrated ability, research performance to date (including the number of published papers), and involvement in the scholarly community. The published output of researchers may sometimes be used as a measure of the effectiveness of the program in which they are involved.

Quantitative measures in science, like qualitative judgments, can be used or abused. Although measures of the publication output of countries, institutions, and even individuals can be useful, especially when employed comparatively, their uncritical overuse in the evaluation of an individual has led, long before systematic citation analysis became possible, to the phenomenon known as "publish or perish." The use of citations was originally intended to provide a quantitative tool to help differentiate the prodigious output of teams of scientists. But the abuse of citation "impact" is also possible. While there is no discernible trend in the misuse of citations, some adminstrators, like some journalists, will grab at straws under the pressure of externally established deadlines. They, like editors, need to be vigilant, so that the final result is not a set of misleading conclusions but rather better science.

* * * * *

My thanks to Abigail Grissom for her help in the preparation of this essay.

REFERENCES

1. **Garfield E.** How to use citation analysis for faculty evaluations, and when is it relevant? Pts. 1&2. *Essays of an information scientist.* Philadelphia: ISI Press, 1984. Vol. 6. p. 354-72. (Reprinted from: *Current Contents* (44):3-11, 31 October 1983; (45):3-12, 7 November 1983.)
2. How much research is enough? *NY Times* 28 January 1985. p. A14.
3. **Garfield E.** The articles most cited in the *SCI* from 1961 to 1982. 7. Another 100 *Citation Classics*: the Watson-Crick double helix has its turn. *Current Contents* (20):3-12, 20 May 1985.
4. **Price D J D.** Networks of scientific papers. *Science* 149:510-5, 1965.
5. --------------. A general theory of bibliometric and other cumulative advantage processes. *J. Amer. Soc. Inform. Sci.* 27:292-306, 1976.
6. **Merton R K.** The Matthew effect in science. *Science* 159:56-63, 1968.
7. **Zuckerman H.** *Scientific elite.* New York: Free Press, 1977. p. 42.
8. **Houssaye A.** Histoire du 41me fauteuil de l'Académie Française (History of the 41st chair of the French Academy). Paris: L. Hachette, 1856. 368 p.
9. **Cole J R & Cole S.** Measuring the quality of sociological research: problems in the use of the *Science Citation Index. Amer. Sociol.* 6:23-9, 1973.
10. **Moravcsik M J & Murugesan P.** Some results on the function and quality of citations. *Soc. Stud. Sci.* 5:86-92, 1975.
11. **Chubin D E & Moitra S D.** Content analysis of references: adjunct or alternative to citation counting? *Soc. Stud. Sci.* 5:423-41, 1975.
12. **National Science Board.** *Science indicators 1982: report of the National Science Board.* Washington, DC: National Science Foundation, 1983. 344 p.
13. **Jones L V, Lindzey G & Coggeshall P E,** eds. *An assessment of research-doctorate programs in the United States.* Washington, DC: National Academy Press, 1982. 4 vols.
14. **Watson J D & Crick F H C.** Molecular structure of nucleic acids. *Nature* 171:737-8, 1953.
15. **Small H G.** Citation context analysis. (Dervin B & Voigt M J, eds.) *Progress in communication sciences.* Norwood, NJ: ABLEX, 1982. Vol. III. p. 287-310.
16. **Cole S.** The growth of scientific knowledge. (Coser L A, ed.) *The idea of social structure: papers in honor of Robert K. Merton.* New York: Harcourt Brace Jovanovich, 1975. p. 175-220.
17. **Merton R K.** Foreword. (Garfield E.) *Citation indexing—its theory and application in science, technology, and humanities.* New York: Wiley, 1979. p. vii-xi.
18. --------------. *Social theory and social structure.* New York: Free Press, 1968. p. 27-9; 35-8.
19. **Garfield E.** The "obliteration phenomenon" in science—and the advantage of being obliterated! *Essays of an information scientist.* Philadelphia: ISI Press, 1977. Vol. 2. p. 396-8.
20. **Merton R K.** *On the shoulders of giants.* New York: Harcourt Brace Jovanovich, 1985. 300 p.
21. --------------. *The sociology of science.* Chicago: University of Chicago Press, 1973. p. 402-8.
22. **Garfield E.** Pageless documentation; or, what a difference a page makes. *Current Contents* (17):3-6, 29 April 1985.