

Citation Indexing -- Its Theory and Application in Science, Technology, and Humanities

Eugene Garfield

Institute for Scientific Information

Foreword by Robert K. Merton

Foreword

I like to read some general significance into Eugene Garfield's having offered me the privilege of introducing the reader to this book, which sets out the history, method, and implications of his distinctive contribution to the advancement of science: the first multidisciplinary citation index to the scientific literature. My being a close friend cannot account for his having done so. Gene has many good friends more knowing than I about information science and the methodology of citation analysis. I suspect that the significance of his choice lies precisely there: in his having reached out for a sociologist of science rather than an information scientist. For one of the remarkable things about Eugene Garfield is that along with the imagination, pragmatic judgment, and immense energy required to invent, produce, and develop a useful tool for a seemingly routine but fundamental task in science -- searching the literature -- he has a deep intuitive sense of the social, cultural, and cognitive structures latent in the practice of science. This found expression in his seeing, from the very start, that his invention of the *Science Citation Index* would develop into a powerful tool for the historical and sociological study of science. I imagine that he wanted to have the subject of the book set squarely within this broader framework rather than the primary but narrower one of information science alone. I accept the inferred assignment with pleasure -- and understandable misgivings.

This book is long overdue. Its most immediate didactic value resides in helping the reader to understand the concept and workings of the citation index as a device for making the search of the literature much more comprehensive and far less arduous than was previously possible. The book thus fills an obvious need for a systematic account of the concept and uses of a major bibliographical tool designed to facilitate the practice of science. And it does so with the clarity and informed assurance that one would expect of the man who invented the tool.

Important as it is to fill that need for practical information about citation indexes as a bibliographical tool, for me the chief interest of the book lies elsewhere. That judgment is formed of course, -- some would say, biased -- by the perspective of a sociologist primarily interested in trying to understand how interacting cognitive and social structures of science affect the thought and behavior of scientists. Indeed, from that standpoint, the citation index itself provides a case study in the process of invention and discovery. Even in bare outline, the case holds its own sociological interest.

We can begin with one aspect of the latent social and cultural structure of science presupposed by the historically evolving systematic use of references and citations in the scientific paper and book. That aspect is the seemingly paradoxical character of property in the scientific enterprise: the circumstance that the more widely scientists make their intellectual property freely available to others, the more securely it becomes identified as their property. For science is public not private knowledge. Only by publishing their work can scientists make their *contribution* (as the telling word has it) and only when it thus becomes part of the public domain of science can they truly lay claim to it as theirs. For that claim resides only in the recognition of the source of the contribution by peers. The greatest ambition of a productive scientist is to do the kind of work that will be much used and much esteemed by fellow scientists best qualified to assess its worth- And, in general, scientific work is esteemed in the measure that others can draw upon it to advance their own future inquiry.

All this is reinforced by the reward system of science. Since recognition by qualified peers is the basic form of extrinsic reward (all other extrinsic rewards deriving from it) and since that reward can be accorded only when the work is made known, this historically evolving reward system provides institutionalized incentive for open publication without direct financial reward. Such peer recognition is usually accorded the first published contribution of its kind, later ones presumably being redundant. But since the cognitive structure of science makes for independent multiple discoveries -- functionally equivalent if not identical discoveries -- this social and cognitive complex evokes a concern among scientists to get there first and to establish, through prompt publication, their self-validating claims to priority of discovery.

The anomalous character of intellectual property in science becoming fully established only by being openly given away (i.e., published) links up with the correlative moral as well as cognitive requirement for scientists to acknowledge their having made use of it. Citations and references thus operate within a jointly cognitive and moral framework. In their cognitive aspect, they are designed to provide the historical lineage of knowledge and to guide readers of new work to sources they may want to check or draw upon for themselves. In their moral aspect, they are designed to repay intellectual debts in the only form in which this can be done: through open acknowledgment of them. Such repayment is no minor normative requirement. That is plain from the moral and sometimes legal sanctions visited upon those judged to have violated the norm through the kinds of grand and petty intellectual larceny which we know as plagiarism. (Karl Marx testifies to the possible depth of commitment to the norm: for him, plagiarism was the one altogether detestable crime against private property, as witness his preface to the first edition of *Capital* and his further thunderings on the subject throughout that revolutionary book.)

It was of course unnecessary for Eugene Garfield to identify this composite communications-intellectual-property-and-reward system in order to arrive at his concept of the citation index. He only needed the sense that the system provided the ingredients for systematically identifying, through citation indexing, links between the work of scientists that could be put to use both for searching the literature and for

exploring cognitive and social relationships in science. It is symptomatic of the direction and pace of development that about half the book is given over to modes of citation analysis designed for the latter kind of development.

Eugene Garfield's account of the birth, adolescence, and early maturity of citation indexing reports other patterns typically found in the evolution of an invention. Detailed procedural research, sometimes in the form of mini-experiments, is designed to test the feasibility of successively expanded and differentiated uses. Self-generated criticism and external criticism give rise to error-detecting and error-correcting research. The exercise of such organized skepticism exerts pressure for a deeper sociological understanding of the interaction between the communication-and-reward system in science and its cognitive development. As the sociologist Norman Kaplan (1) was perhaps the first to note a dozen years ago and as varied sociological applications of citation analysis (2) since then have made abundantly plain, we need to know more than is yet known about what references and citations do and do not represent if citation analysis is to provide further understanding of how science is socially and cognitively organized and practiced.

Certain patterns of referencing behavior would seem to set limits on the use of citation counts for tracing the long-term genealogy of ideas. One of these patterns has been described as "obliteration by incorporation": the obliteration of the source of ideas, methods, or findings by their incorporation in currently accepted knowledge (3). In the course of this hypothesized process, the number of explicit references to the original work declines in the papers and books making use of it. Users and consequently transmitters of that knowledge are so thoroughly familiar with its origins that they assume this to be true of their readers as well. Preferring not to insult their readers' knowledgeability, they no longer refer to the original source. And since many of us tend to attribute a significant idea or formulation to the author who introduced us to it, the altogether innocent transmitter sometimes becomes identified as the originator. In the successive transmission of ideas, repeated use may erase all but the immediately antecedent versions, thus producing an historical palimpsest in which the source of those ideas is obliterated.

To the extent that such obliteration does occur -- itself an empirical question that is only beginning to be examined -- explicit citations may not adequately reflect the lineage of scientific work. As intellectual influence becomes deeper, it becomes less readily visible. That influence may operate through acceptance of a theoretical framework, with its basic assumptions, or through standardized procedures of inquiry. In short, it may be canonical knowledge that is most subject to obliteration of source. As Joshua Lederberg noted in the foreword to another of Eugene Garfield's books: "The work that *everybody* knows. . . is hardly cited at all!" (4). Specialized historical and sociological studies can supplement explicit citations with tacit ones, the kind that can be reconstructed from textual evidence such as eponymous allusions, terminology bearing the stamp of the source of an idea, and the like. But it remains to be seen whether some modes of significant cognitive influence find no expression in references and citations, explicit or tacit.

In the evolution of citation analysis, questions such as these have begun to receive the systematic attention they require. Cole, Cole, and Dietrich, for example, and Garfield, Malin, and Small (5) have addressed themselves to the problem, the latter suggesting that even when obliteration has occurred, "scientific ideas that have been regarded as important or influential can be associated with one or more scientific works that are *at some time* highly cited." It is thus argued that citation counts do not favor mediocrity, since obliteration of this kind takes place only after substantial visibility through citations has occurred, a requirement which a large fraction of published work fails to satisfy. All this raises interesting question about the time frames within which citation analysis can effectively trace the genealogy of scientific knowledge.

The reader will find in this book a plentiful variety of other uses to which the data base of the *Science Citation Index* has been put. A good deal is said about one of the most disputed of these: the use -- some of us would say, the abuse -- of citation counts as the principal or determining basis for assessing the research performance and further potentialities of individual scientists. This was not unforeseen. As early as 1963, just as the first *Science Citation Index* was being published, a cautionary note was sounded about the possible "promiscuous and careless use of quantitative citation data for . . . evaluation, including personnel and fellowship selection." That forewarning came, properly enough, from Gene Garfield (6). He went on to say, "It is preposterous to conclude blindly that the most cited author deserves a Nobel prize. "

The closing chapter of this book reexamines the subject in light of the intervening 15 years of research. That chapter can be read less as a newly developed defense of the use of citation analysis for assessing individual scientific performance than as a methodological manual for those who venture into those dangerous waters. A recurring theme in the chapter is the strong reminder that citation counts cannot be responsibly taken as the controlling basis for appraisals of individual performance. At best, they are ancillary to detailed judgments by informed peers. The forensic use of citation counts to compare the impact of scientific contributions by individuals only provides an extreme type of occasion for subjecting such practices to the organized skepticism that is one of the fundamental characteristics of science.

No one reading this book can fail to note its pervading sense of exuberant promise. Citation indexing has been a standard of scientific bibliography for more than a decade but its sociological and historical research potentials presumably have not yet been fully realized. One of its notable contributions to this time has been the emerging specialty of bibliometrics, which builds on the early work of such scholars as Bradford and Gross to define and quantify parts of the multiple structures of science. The growing numbers at work in this field have in their turn begun to ask fundamental questions about citation behavior.

I can report that most of these questions are critically examined in this book. The scientific curiosity coupled with skepticism that led Eugene Garfield first to conceptualize and then to develop a new way of indexing the scientific literature are still at work. Applied to his invention of the *Science Citation Index*, this attitude of mind has led Eugene Garfield and his colleagues -- both his colleagues-at-a-distance and his colleagues in the Institute for Scientific Information which he founded -- to collaborate on what amounts to an evolving grammar of citation analysis.

ROBERT K. MERTON

REFERENCES:

Forewords ordinarily do not have references appended to them. In the case of this book, however, one simply dares not substitute allusions for explicit references.

1. **Kaplan, N.** "The Norms of Citation Behavior: Prolegomena to the footnote." *American Documentation* 16: 179-184, 1965.
2. **Cole, J.R. and Cole, S.** "Measuring the Quality of Sociological Research: Problems in the Use of the *Science Citation Index*. *American Sociologist* 6: 23-29, 1973; **Moravcsik, M.J. and Murugesan, P.** "Some Results on the Function and Quality of Citations." *Social Studies of Science* 5: 86-92, 1975; **Chubin, D. E. and Moitra, S. D.** "Content Analysis of References: Adjunct or Alternative to Citation Counting?" *Social Studies of Science* 5: 423-441, 1975.
3. **Merton, R. K.** *Social Theory and Social Structure* (New York: The Free Press, 1968) pp. 27-29, 35-38; **Garfield, E.** "The Obliteration Phenomenon." *Current Contents* No. 51/52: 5-7, (22 Dec. 1975); **Messeri, P.** "Obliteration by Incorporation: Toward a Problematics, Theory and Metric of the Use of Scientific Literature." Columbia University, unpublished ms., 1978.
4. **Lederberg, J.** Foreword to Garfield, E. *Essays of an Information Scientist* (Philadelphia: ISI Press, 1977), p. xiii
5. **Cole, S., Cole, J. R. and Dietrich, L.** "Measuring the Cognitive State of Scientific Disciplines." Pp. 221-229 and **Garfield, E., Malin, M.V., and Small, H.** "Citation Data as Science Indicators." Pp. 181-182 in **Elkana, Y., Lederberg, J., Thackray, A., Merton, R.K., and Zuckerman, H.** (eds.) *Toward a Metric of Science: The Advent of Science Indicators* (New York: John Wiley & Sons, 1978).
6. **Garfield, E.** "Citation Indexes in Sociological and Historical Research." *American Documentation* 14:289-291, 1963.

Added to WebSite :November 1996

[home]	[next]
------------------------	------------------------

It is part of the sociology of science and has applications to science policy-making. It involves quantitative studies of scientific activities, others, publications, and overlaps on bibliometrics to some extent [3,4]. Scientific literature is a reflection of scientific activity and productivity [5]. Velmurugan [6] analyzed to identify the publications, trend on fossil fuels in Indian perspective with 943 and its h-index was 73 and average citations per item was 28.63, total sum of times cited was 26,997, without self citations was 26,159, citing articles was 21,694, without self citations was... 5. Garfield E, Merton RK (1979) Citation indexing: Its theory and application in science, technology, and humanities. Wiley, New York, USA.

The Science Citation Index (SCI) is a citation index originally produced by the Institute for Scientific Information (ISI) and created by Eugene Garfield. It was officially launched in 1964. It is now owned by Clarivate Analytics (previously the Intellectual Property and Science business of Thomson Reuters).^{[1][2][3][4]} The larger version (Science Citation Index Expanded) covers more than 8,500 notable and significant journals, across 150 disciplines, from 1900 to the present.Â Citation Indexing: Its Theory and Application in Science, Technology, and Humanities. Information Sciences Series. New York: Wiley-Interscience.