The Web of Knowledge

A Festschrift in Honor of Eugene Garfield

Edited by

Blaise Cronin
and
Helen Barsky Atkins

ASIS Monograph Series

Information Today, Inc.
Medford, New Jersey

First printing, September 2000

The Web of Knowledge: A Festschrift in Honor of Eugene Garfield

 Published by Information Today, Inc., 143 Old Marlton Pike, Medford, New Jersey 08055.

♾ This book is printed on acid-free, archival-quality paper.
Printed and bound in Canada.

Library of Congress Cataloging-in-Publication Data

The web of knowledge : a festschrift in honor of Eugene Garfield / edited by Blaise Cronin and Helen Barsky Atkins.
p. cm. -- (ASIS monograph series)
Includes bibliographical references and index.
ISBN 1-57387-099-4
1. Indexing. 2. Science -- Abstracting and indexing. 3. Garfield, Eugene.
I. Garfield,
Eugene. II. Cronin, Blaise. III. Atkins, Helen Barsky. IV. Series.

Z695.9 :W37 2000
025.4'95—dc21

00-057554

Publisher: Thomas H. Hogan, Sr.
Editor-in-Chief: John B. Bryans
Managing Editor: Janet M. Spavlik
Production Manager: M. Heide Dengler
Copy Editor: Pat Hadley-Miller
Designer: Jeremy M. Pellegrin
Indexer: Sharon Hughes

CHAPTER 14

A Short History of the Use of Citations as a Measure of the Impact of Scientific and Scholarly Work

Jonathan R. Cole

Today, citation analysis is used extensively to measure the impact and quality of scientific work as well as the intellectual influence of scientists and scholars. It is also used as a tool to describe social and intellectual networks, and as a measure of individual achievement. The range of uses is quite extraordinary. Consider only a few. Social network analysts map the structure of academic fields through increasingly sophisticated structural models of the interrelationship of patterns of citation. Citations are used to measure the cross-fertilization of scholarly ideas from one field to another. They are used as a measure of impact of an individual's work on his or her peers. They are used to compare the relative impact of scholars who have received various forms of honorific recognition such as the Nobel Prize, membership in the National Academies, with scientists or scholars not similarly honored. Tracing intellectual influence from one generation of scientists or scholars to another is often done by examination of "intergenerational" citation patterns. Career trajectories of different types of scientists and assessments of the impact of a scientist's work at various stages of his or her career are constructed by examining the trend of citations to scientists' work. More sophisticated citation studies examine how recognition through citations changes the probability of success from one career stage to another.[i]

Given the recent American obsession to rank everything, it should not surprise us that citations have been used to rank-order almost every aspect of universities and colleges. The forces that contribute to the desire among faculty members of departments, schools, hospitals, colleges, to say nothing of academic administrators and state legislators, to be ranked number one, or among the top five, or ranked as "distinguished" or among "the best" has led professionals as well as amateurs to use citations in these rating games—without any real discussion of their correlates. To cite several examples, professional schools are rank-ordered in terms of the average number of citations received by members of the faculty. The quality or reputations of American Ph. D. programs, which are described in periodic reports from the National Research Council, use citations as one indicator of the quality of the faculty at the evaluated departments, thus producing their own rank-ordering.[ii] The "impact" of scholarly journals has been measured and ranked in terms of the concentration of citations to articles that appear in them.

The distinguished and prolific United States seventh circuit judge Richard Posner (1990), who teaches law at the University of Chicago when not working at the bench, has recently used citations to map the influence of decisions made by Justice Benjamin Cardozo. I have seen citation comparisons become part of depositions in litigation over whether a candidate has been denied tenure on grounds other than the quality of teaching or published research. The lawyers who use them are rarely informed about the appropriate and inappropriate uses of these measures. Citations are even used as a bibliographic tool to identify relevant literature that is of interest to a scientist or scholar—the original intent of the citation indexes produced by Eugene Garfield and the Institute for Scientific Information.

In short, citation analysis has become a small cottage industry today. The legitimacy of its use for a variety of purposes has been established. In fact, that legitimacy has extended to the point where many suspect uses of citations are accepted without significant skepticism. Consider an example drawn from personal experience. Probably one-third of the more than 400 cases for tenure at Columbia that I have sat in on as Provost have made a case for the candidate based, in part, upon the impact of scholarship measured through citation counts. Rarely, if ever, are these counts, which compare one individual with another, accompanied by a set of caveats about the limits of such individual comparisons or the bases on which the comparisons are being made.

When I listen to the presentations of citation analysis as an argument for a tenure case, I often wonder whether these advocates have any idea of the origins of the uses of citations as indicators of impact or quality or how the measure was developed. Few ask where this practice or set of measures came from or about the sociological and historical origins of the uses of citations as measures of impact, quality, influence, or academic reputation. When I have, on occasion, asked, there is surprise that much of the origins can be found in Columbia history.

This brief historical excursion attempts to locate in time and place the origin of one aspect of citation analysis and in doing so to properly acknowledge Eugene Garfield, without whom the development of these indicators would surely have been much delayed. In this retrospective, I am not, of course, tracing the historical origins of the "footnote," which has been done with great skill and wit recently by Anthony Grafton (1997). My project is far more limited. I will discuss the development of the use of citations as a means of developing analytic methods in the sociology of science. The technology of the *Science Citation Index (SCI)* made these new forms of analysis possible.

The first and major section of the paper focuses on a description of the origins of the uses of citations and the indexes that went beyond their original intent as a bibliographic tool. The story represents a case of how the development of measures was coupled with conscious efforts to construct an academic specialty. The precise origins of specific uses may represent an instance of Robert Merton's concept of "independent multiple discoveries." The rest of the paper will discuss briefly the early skepticism and resistance to the use of citations as a measure of impact—as an illustration of appropriate scientific skepticism about new measurement techniques and in part as resistance to the social studies of science. The paper will also casually address the lack of clarity in the measurement tools, the concern over the improper use of citations in the reward system of science, and the diffusion of the idea and techniques from sociology to other disciplines.

The Historical Development of the Use of Citations as a Measure of Impact, Quality, and Reputation

My story begins in the early 1960s and extends for about a decade. The putative father of the sociology of science is Robert K. Merton. If you were searching for the origin of the sociology of science, you might reach back further than Merton's Harvard doctoral dissertation and subsequent book, "Science, technology and society in seventeenth-century England" (1938), but you would be indulging in adumbrationism in doing so. Merton began the sociological study of science in that work, but curiously wrote only intermittently about the subject for the next twenty years.[iii] In the meantime, he established himself as one of the leading American social scientists of the twentieth century. He did write critically important papers about the normative structure of science during the period where he concentrated on other subjects, and one can find in all of his work a number of interrelated themes that link science with the analysis of other elements in social systems.[iv] Although Merton published occasionally on science, he made no effort during this period to create a new specialty.

In August 1957, Merton strategically chose to focus on science in his presidential address to the American Sociological Association. His paper, "Priorities in scientific discovery," (1957) returned to historical examples. But, more importantly, it set the stage for the sociological study of science in a systematic way. Merton had demonstrated several things in his paper and in the cumulative work on science since his dissertation: first, that the social study of science, beyond the Marxian treatment of it, could yield important results and new questions for empirical analysis; second, that science could be an object of useful and informative analysis carried on by non-scientists (a proposition that met with substantial resistance during these early years)[v]; and, third, that sociologists could treat the organization of science—its social structure, its reward systems, its norms and values—as an object of study that would yield results of both theoretical and practical value—and that science was just as legitimate an object of sociological study as was the family, the economy, deviant behavior, religion, and other major social institutions.

Merton's interest in science as an exemplar for sociological theory and research "in the middle range," led to his formation of a Columbia seminar and then research program in the sociology of science. This began in the fall of 1965. Merton recruited three students to "the first round" of the seminar who would make the study of science a principal part of their academic careers: Harriet Zuckerman, who had already begun her now renowned work on Nobel laureates,[vi] Stephen Cole, who had previously worked on problems of deviant behavior and historical sociological themes and who would go on to become one of the major international figures in the sociology of science, and myself, a second year graduate student. The program lasted for more than a decade and led to the recruitment of many graduate students. It was supported throughout most of the period with substantial grants from the National Science Foundation.[vii]

Much of the early work in the seminar dealt with aspects of the social organization of science, its normative structure, and its reward system as it played out in the careers of physical scientists. We did not focus on the content of science *per se.* Robert Merton is known of course as a social theorist, but in many ways he was more insistent on developing empirical tests of ideas, often stemming from his own work, than on further development of theory without focused tests of those theoretical ideas.[viii] So the seminar's fundamental objective was to raise good questions for which we had no answers, and to do so within a theoretical framework. Consequently, early on, the following issues became problems for inquiry: Was science a truly meritocratic or fair system that allocated rewards on the basis of the quality of work? Or, did functionally irrelevant factors, such as sheer quantity of research output, or gender, or "old boy" connections determine rewards? Was the communications system of science really an open one? What sociological variables were causally related to recognition in science? What effects did social, or academic context, have on the distribution of rewards and scientific productivity? What sociological factors, independently of innate ability or individual attributes, determine visibility and "eminence?" These were some of the puzzles that occupied us in those early days.

In the salad days of that enterprise, when we all felt on a mission to build a specialty of knowledge that had not existed before, we were also working to connect our work with knowledge generated in allied disciplines, particularly the history of science. Stephen Cole and I were interested initially in two questions, among others: What determined scientific eminence? Was there truth to the widespread belief that in science you either publish or perish—which we translated into whether the quantity of scientific production was more important than its quality in determining scientific recognition? There were three major Columbia influences on our early work. One was Merton's own theoretical work on the normative structure of science, of course. A second was the methodological approach associated with causal analysis developed by Paul F. Lazarsfeld. A third was Bernard Barber, also a member of the Columbia faculty at the time, who was pursuing both theoretical and empirical studies in science and whose students were also students in the Merton seminar.[ix] In a variety of studies carried out by Barber's students, including Daniel Sullivan, citations were used as measures of impact and influence.[x]

A fourth, and critically important influence, could be found in the published work by the peripatetic and enormously creative Yale historian of science, Derek J. de Solla Price. In his imaginative representations of the historical growth of scientific knowledge, Price had charted new territory in "Little science, big science" (1963), where he presented a variety of indicators to measure scientific productivity and the rate of growth of knowledge in various scientific disciplines. Price was interested principally in estimating accurately the growth curves of science—which he hypothesized was logistic. He provided us with both mathematical and empirical indicators of the functional form of the curve. Similarly, he was interested in the distribution of productivity across the population of scientists in a field, and he proposed that a small proportion of scientists accounted for a very large proportion of the published scientific literature. Price's conjectures proved accurate. This productivity inequality would enable many of us to use indicators such as citations for populations and large samples that would prove adequate because of the functional form of productivity in science. Price had a passion for outlining the paths of growth—he was concerned with these "dependent variables."

We had a different objective. In studying sociological problems related to scientific careers, scientific productivity, and recognition, we were interested in explaining variations in scientific productivity and whether rewards were principally determined by the quality of performance. In short, we wanted to go beyond the accurate representation of patterns of productivity to explaining what determined them and what were the consequences of these different patterns. We were interested in the "independent variables" that were correlated with Price's patterns of growth and with other patterns of differential recognition, and ultimately with causal forces operating in these social systems. This interest led directly to the question: What stable indicators could we find and use that would help us explain productivity patterns?

We know that Eugene Garfield's idea for a citation index was formulated in the mid-1950s. Some of these origins and the initial resistance are made plain in correspondence between Nobel Prize biologist and polymath Joshua Lederberg and Gene Garfield (see Lederberg's contribution in this volume for a fuller account). Lederberg encouraged Gene to pursue the idea since he could immediately see the value of such a bibliographic tool in the field of genetics. Lederberg was convinced that the idea for a citation index was a good one, as he made clear to Garfield in a letter of May 9, 1959:

> "Since you first published your scheme for a 'citation index' in *Science* about 4 years ago, I have been thinking very sincerely about it, and must admit I am completely sold. In the nature of my work I have to spend a fair amount of effort in reading the literature of collateral fields and it is infuriating how often I have been stumped in trying to update a topic, where your scheme would have been just the solution! I am sure your critics have simply not grasped the idea, and especially the point that the author must have to cooperate by his own choice of citations and then <u>he</u> does the critical work.
>
> "Have you tried to set this out in an adequate experiment? Would you look for support from the NSF? Of course you have to count on opposition from the established outfits which have already succeeded in blocking any progressive centralization of the Augean tasks." [Source: http://garfield.library.upenn.edu/lederberg/050959.html]

Garfield responded in a May 21, 1959 letter:

> "I hope you won't be embarrassed by a show of emotion, but your memo almost brought tears to my eyes. It then seemed that over six years of trying to sell the idea of citation indexes had not been completely in vain ... You ask whether I have tried to set this out in an adequate experiment. In the beginning (way back at Johns Hopkins when I first started on the idea in 1953) I did do a little experimenting. I had my own doubts at first. Actually it took quite some time and effort just to get together the material on the Seleye article ... For some time after [1955] I tried to convince certain illiterates at the National Science Foundation to give me a small grant to conduct research on citation indexes. In those days Sputnik was an unknown word and the NSF Office of Scientific Information was an equally unknown entity. However, some congressman, at an NSF budget hearing, asked the naïve question 'why don't you Shepardize the literature—we lawyers don't seem to have the troubles you do.' As a result of this simple questioning the NSF made a statement in one of its reports that

> it was planning to support research on citation indexes..."
> [Source: *ibid.*, 052159.html]

The existence of the *Science Citation Index (SCI)* as a product of the Institute for Scientific Information came to Steve's and my attention during the first year of the Merton Seminar. The first two years of the index already existed—1961 and 1964. I do not recall the precise date or source of information that led us to the index, but it started us on a very long journey of using it to develop measures and indicators of key concepts in the sociology of science. In fact, I believe our first use (and perhaps the very first use of the *SCI* as a tool to measure the impact or quality of scientific work) came in data that we collected in 1965–1966 on a sample of 120 physicists.[xi] Steve and I produced a brief "pre-data collection outline," which really was a short paper with hypothetical tables that was the basis for a presentation to the seminar. It was entitled "The measurement of scientific eminence" (May 4, 1966). It is apparent from the "paper" that we were struggling to define the relationships between concepts and indicators. In that distant 1966 memo, we referred to citations as measuring both "impact" and "quality." We began with simple ambitions:

> "This paper is a methodological one. Its principal aim is to suggest a way to measure scientific eminence and to prove that the suggested technique is more valid than other techniques which have been used in the past. As is well known among the small but growing band of sociologists of science [we did have a feeling of being part of a tribe], the major impediment to rapid advance is the difficulty in measuring scientific output. If we are to discover the social conditions most favorable for the advance of science, we must be able to measure scientific output. This paper is the first in a series that will deal with measurement problems." [Cole and Cole private seminar paper. It was presented to the seminar on May 4, 1966. A copy of that unpublished outline and narrative was located in Merton's famous files.]

The narrative of this early outline goes on to describe an empirical project of studying the "current population of American university physicists." We sent questionnaires to all physicists who worked in academic departments that had granted at least two Ph. D.s per year between 1959 and 1963—a total of 2,079 working in 86 departments. We gathered biographical information on these physicists from "American Men of Science" *(sic)*, publication counts for the physicists from "Science abstracts," and citation counts from the *Science Citation Index (SCI)*. The questionnaire asked the physicists to tell us whether they were familiar with the work or even the names of a random sample (stratified by age, university affiliation, productivity and number of awards) of 120 physicists.

We were attempting to explain the variations in recognition of the 120 physicists in terms of the prestige of their affiliations, levels of scientific productivity, quality of their work, honorific recognition, age, and other variables hypothesized to be causal determinants of scientific recognition. It was in that first empirical study that Steve and I staked out a claim for using citations to measure "quality" of scientific work:

> "It has often been maintained that using the sheer number of papers as a measure of scientific output is misleading because it ignores the crucial variable of 'quality.'... 'Quality' like the word 'creativity'... has caused much conceptual confusion among sociologists of science. Without here getting involved in a conceptual discussion of 'quality' let us take as a measure of quality the impact of a man's work on his colleagues as measured by the number of citations to his work." (We are currently planning a paper that will demonstrate how citation indexes may be used to measure "quality" of scientific work.) [*ibid.*, p. 4]

The working paper went on to discuss ways in which we were anticipating using citations as a measure of quality and impact and to weigh citations against other factors in determining the visibility of scientists and other forms of their recognition. But for my purposes here, the important point is that this seminar paper and its antecedent work represented one of the first times, to my knowledge, that citations were being used systematically as a measure of scientific quality or impact with relatively large samples of scientists.[xii] The results of that study were first presented to the annual meetings of the American Sociological Association in August 1966 and the paper was published in the *American Sociological Review*, as "Scientific output and recognition: A study in the operation of the reward system of science" (Vol. 32, June 1967, 391- 403).

It remains somewhat unclear whether our use of citations as an independent variable in studying social aspects of science was in fact the first such use. Perhaps the first published work to discuss correlates of citation measures was an Autumn 1966 paper by Alan E. Bayer and John Folger, "Some correlates of a citation measure of productivity in science" (1966) [Volume 39, Issue 4 (Autumn, 1966), 381-390)], published in Notes and Reports of the *Sociology of Education*. Bayer and Folger's abstract of the paper begins:

> "The *Science Citation Index* provides an easy way to derive criterion measures of scientific accomplishment. Measures derived from citation counts, the principal criterion, have high face validity ... "

The paper opens with the acknowledgement that the measurement of the factors that affect scientific accomplishment have attracted the attention of sociologists and psychologists but that adequate measurement of productivity has frustrated researchers.

> "In theory, evidence of achievement can be identified and measured, but in practice, research workers have found the collection of criterion measures difficult and often very expensive" (p. 381).

They then draw our attention to a promising new way of measuring productivity:

> "This paper describes the use of a new index of citations which has the potential to simplify research by providing a reasonably adequate criterion of productivity in science for a number of science fields. The use of the index is illustrated by a pilot study of productivity in bio-chemistry.... Within the past few years a new indexing system has been introduced that greatly simplifies the problem of obtaining citation counts. Under the sponsorship of the National Science Foundation and the National Institutes of Health, the Institute for Scientific Information in Philadelphia has compiled and published the *Science Citation Index* (SCI). The index lists citations by cited senior author and citing authors" (pp. 381-383).[xiii]

In fact, in these early days of the *SCI*, Eugene Garfield and Irving Sher were interested in demonstrating the multiple uses of the Index. As early as 1963 Garfield was discussing the uses of citation counts as an indicator of impact.[xiv] At the point of its earliest uses, there were important debates about whether citations actually measured impact of work, rather than quality. Despite his early cautions about the use of citations as a measure of quality, Garfield and Sher were among the first to juxtapose the average number of citations to work produced by scientists who had been "certified" as producing high quality work with those scientists who had not achieved such lofty recognition. They showed, for example, that 1962 and 1963 Nobel laureates in physics, chemistry and medicine received far more citations to the work they published prior to receipt of the prize than the average cited author in the *SCI* file for 1961.[xv] Garfield was interested in the widespread use of his indexes and he wrote extensively about their various uses as found in the papers of social scientists.

It remains unclear that we can identify the "discoverer" of citation use as a measure of quality. The construction of the index made work with it possible and several individuals began to use it as an indicator of more abstract sociological concepts. It is fair to say, however, that the published papers and books that came out of the Merton seminar and program, predominantly by Stephen Cole and me, but also from Harriet Zuckerman and Merton, as well as several other students, represented the first systematic use of citations as a measure of impact, quality, and intellectual influence. Others may have published a paper that refers to or even used citations at roughly the same time, but they did not focus on demonstrating the value of this indicator and its shortcomings, nor did it play a central role in their research program over time. It is possible that the "discovery" of the use of citations as a measure of quality, impact, and influence is yet another demonstration of the abundant scientific phenomenon of "multiple discoveries," but there can be little doubt, I believe, that the systematic development of concepts and indicators involving citations in the sociology of science came from the Columbia group, and most specifically from Steve's and my published work (Merton, 1961).[xvi]

Regardless of the precise origin of the use of citations as indicators of important concepts in the sociology of science, the papers that used the *SCI* began to emerge with increasing frequency in the social science literature. None of this would have happened without the technological innovation of Eugene Garfield. The creation of the *SCI* represents a good case study of how technological innovations very frequently create the necessary conditions for significant advance in scientific fields. That story has been told often, of course, and does not bear repeating here. We can say that the Columbia program in the sociology of science was among the "early adopters" of this new technology.

An irony existed in the early development and use of this new technology by those of us at Columbia. We were studying, among other things, the normative code of science, among which was the prescription that there be open communication of new theories and methods. Our work with citations involved increased dependence on a technology that was created with a proprietary interest in mind. We encouraged the development of the *SCI* because of its potential use by scholars. But the use of the indexes was restricted to those who would have access to them and those who would pay for the information. This new intellectual property was owned by the ISI, which was created as a for-profit business. It represents an early example of what has become commonplace today: a new structural relationship between industry and research universities. As much as we could see the value of the citation indexes, we were frustrated with some of our early experiences with the ISI.

We did not fully appreciate the goals of a start-up company that was interested in advancing the growth of knowledge but was also born as a profit-making enterprise. We were prone to view the ISI as an appendage to research universities and their libraries and therefore to expect that it would be governed by norms of exchange that governed work inside the academy. We came to learn early on that we were wrong. Thus, if we needed citation counts from ISI, it was hard for us (perhaps naïve of us) to understand why they refused to provide us with the data we needed for our research at minimal costs. Gene was on a mission but it was one that had a very heavy business side to it.

To his great credit, Garfield was always interested in having the *Citation Index* used for scholarly purposes. It was also not lost on him that we were among his best advertisers of ISI indexes within the academy. But he was creating these indexes to make money as well as to develop knowledge. Today, these mutual goals do not seem to us as particularly incompatible. We are now well acquainted with the consequences of the Bayh-Dole Act of 1982. Probably no other recent act of Congress has had a more positive impact on moving scientific, particularly biomedical, discoveries to the market in a timely way. Indeed, an extraordinary number of biomedical researchers have formed their own start-up companies. But, in the early 1960s, we were interested in being able to use the *Citation Index* at an affordable price; Garfield was willing to provide "educational discounts," but he was never willing to give the contents of the files away for research purposes. Today that seems self-evident, but at the time I thought it quite unreasonable that indexes that were constructed for "other purposes" would be guarded carefully against free access.

Early Resistance to the Use of Citations as a Measure of Quality or Impact

One of the defining norms in the code of science is that of skepticism. Doing science involves a skeptical stance toward assertions of fact without documented and unimpeachable evidence. Adherence to this norm is as important in the social sciences as elsewhere. One by-product of conformity to this norm is the refinement of hypotheses, the scrubbing of data used to support an argument, and the improvement of the relationship between concept and indicator.

When we first proposed the use of citations as a measure of quality and impact of work, the idea was met with a great deal of skepticism. The sources of criticism were important. You first had to deal with internal criticism within the program at Columbia. Merton, himself, was a continual and exceptional critic. As a member of the seminar, you had the feeling that if your ideas could pass muster with Merton, you were apt to be able to defend them against any other authority. This proved to be true in the sociological community, but was not the case in the scientific community. There was great initial skepticism

among scientists about the use of citations as a measure of impact or quality. Scientists, including very eminent ones, were skeptical about the entire enterprise of the social study of science. Many did not believe that social scientists could have anything useful to say about a scientific field without detailed expert knowledge of its cognitive content. When it came to the use of citations, the scientists were doubly skeptical. They could point to cases of highly cited papers or individuals who received the lion's share of their citations for new methods or useful experimental tools. The quintessential example always raised was the fact that O. H. Lowry had the most citations in the entire index and that his work represented a tool that everyone in an area would reference, regardless of whether or not it had a significant impact on his or her work.[xvii] Others would point out the occasional (rare) Nobel laureate who had relatively few citations compared with people who had clearly not done nearly as important work. They would point to the number of self-citations as a confounding element in citation counts; they would, if they knew enough about the *SCI*, point to the problem of homophily in the way the index was constructed. They would point out that the index only recorded citations to first authors and in many fields the most "important" author in a collaborative work often was located at the end of an author sequence. Even those who were predisposed to accept the measure, or at least suspend their disbelief, often questioned whether citation counts really measured quality. These people were unconvinced that there were close linkages between the general concept and the indicator used to measure it.

Many of these skeptics were not out of place. Their questions deserved answers. We wrote papers to answer them and to make clear how citations could be used legitimately and how their use could be abused.[xviii] Our own studies relied on careful collection of data that did not count self-citations; that presented correlations between the counts as they would emerge from the *SCI* and a full count of citations to all authored papers regardless of an individual's position in the author set. We vehemently objected to the use of citations as a way of drawing inferences about quality in individual comparisons. The counts were to be used for statistical aggregates; they were not useful for making individual comparisons across fields and specialties; it was absurd to draw strong inferences from fine-grained differences in citation totals of different individuals, yet that was and still is often done. Given the various problems with citation counts it is inappropriate, for example, to conclude that a scientist with fifteen citations has had a greater impact on his field than one whose work has received ten citations.

Over the next decade, there appeared an increasing number of scholarly papers and books that used citations as an independent variable. Of course, there was, and still is, much mindless use of citations, but there was increasingly a body of literature forming that used the *SCI* in appropriate ways—studies that enabled the empirical domain of the sociology of science to grow.

The uses began to be diffused and reached across national borders and across disciplines. Other social scientists began to use citation data in ways alluded to above. Policy analysts began to use citations to examine the relationship between research and development and scientific output and achievement.

There continued to be debates over precisely what citation counts measured—particularly whether they were measures of impact or quality. Throughout much of the history of the use of these counts, there has been a lack of clarity about the distinction between impact and quality. Citations have been used appropriately to chart "intellectual influence," and in that sense, to measure the impact of work by individuals or groups of scientists on others. The difficulty in making the distinction between impact and quality can be seen by empirical example. There is no question that in the aggregate citation counts are strongly correlated with other independent measures of the quality of scientists' work. Many studies have demonstrated this over the years. However, there are many cases where scientists have received large numbers of "negative" citations, that is, their work is referenced, but often critically. Take for example, works like "Time on the cross" (1974), by the economic historians Robert Fogel and Stanley Engerman. This work of quantitative history about the antebellum South claimed to have tested a set of hypotheses about the nature and consequences of slavery. This two-volume work met with an angry reception, to say the least. If there were substantial errors in the work, they proved "fruitful errors." The book triggered the production of scores of volumes on slavery, almost all of which took issue with Fogel and Engerman or used the work as a point of reference or departure. Many of these new works on slavery proved to be major works. The work represented in "Time on the cross" was a catalyst, in some important sense, in re-energizing a field. How does one assess the nature of citations to it and whether the high number of citations suggest that the books simply had an impact or, despite the negative citations, represented "high quality" if disputed work? Do the citations to it represent impact, quality, or both? We have argued that papers or books that have enormous "potential for elaboration," to use Paul Allison's (1973) apt phrase, are books that have had a high impact and more often than not qualify as works of quality.

Without the creation of the *Science Citation Index,* it would have been far more difficult to develop empirical studies of the social structure of science. It took the vision of Eugene Garfield and his dogged determination to create a new bibliographic tool; it took his messianic zeal to communicate widely the value of the Index; it took business sense and his desire to "sell" it and derivative products, such as *Current Contents,* to the world, to make the technology that made much of our early empirical work possible. Few people who use the *Index* know of Gene's contribution, few people today know how the *Index* was first used for sociological purposes, but such "obliteration through incorporation" is commonplace in science.

Acknowledgment

I want to thank Maritsa Poros for her help in researching the history of citation use in the literature of the sociology of science and related fields.

References

Allison, P. (1973). Social aspects of scientific innovation: The case of parapsychology. Master's Thesis. Madison, WI: Univ. of Wisconsin.

Barber, B. (1952). Science and the social order. Glencoe, IL: Free Press.

Barber, B. & W. Hirsch (Eds.).(1962). The sociology of science. New York, NY: Free Press.

Bayer, A. E. & J. Folger (1966). Some correlates of a citation measure of productivity in science. Sociology of Education, 39(4), 381-390.

Clark, K. (1957). America's psychologists: A survey of a growing profession. Washington, DC: American Psychological Association.

Cole, J. R. (1969). The social structure of science. Ph. D. Dissertation. New York, NY: Columbia Univ.

Cole, J. R. (1970). Patterns of influence in scientific research. Sociology of Education. 43(4), 377-403.

Cole, J. R. & S. Cole (1971). Measuring the quality of sociological research: Problems in the use of the *Science Citation Index.* The American Sociologist, 6, 23-29.

Cole, J. R., & S. Cole (1972). The Ortega hypothesis. Science, 4059(178), 368-375.

Cole, J.R., & S. Cole (1973). Social stratification in science. Chicago, IL: Univ. of Chicago Press.

Cole, J. R. & S. Cole (1976). The reward system of the social sciences. In C. Frankel (Ed.), Controversies and decisions: The Social sciences and public policy (pp. 55-88). New York, NY: Russell Sage Foundation.

Cole, J. R. & S. Cole (1981). Peer review in the National Science Foundation: Phase II. Washington, DC: National Academy of Sciences.

Cole, S. (1970). Professional standing and the reception of scientific papers. American Journal of Sociology, 76, 286-306.

Cole, S. (1975). The growth of scientific knowledge: Theories of deviance as a case study. In L. Coser (Ed.), The idea of social structure: Papers in honor of Robert K. Merton (pp. 175-220). New York, NY: Harcourt Brace Jovanovich.

Cole, S. (1979). Age and scientific performance. American Journal of Sociology, 84, 958-977.

Cole, S. & J. R. Cole (1966). The measurement of scientific eminence. Sociology of science seminar. New York, NY: Columbia Univ.

Cole, S. & J. R. Cole (1967). Scientific output and recognition: A study in the operation of the reward system in science. American Sociological Review 32(3), 377-390.

Cole, S., J. R. Cole & L. Dietrich (1978). Measuring the cognitive state of scientific disciplines. In Y. Elkana, J. Lederberg, R. K. Merton, A. Thackray & H. Zuckerman (Eds.), Toward a metric of science: The advent of science indicators (pp. 209-252). New York, NY: Wiley.

Cole, S, J. R. Cole & L. Rubin (1977). Peer review in the American scientific community. Scientific American. 237(4), 34-41.

Cole, S., J. R. Cole & G. Simon (1981). Chance and consensus in peer review. Science, 214, 881-886.

Cole, S., L. Rubin & J. R. Cole (1978). Peer review in the National Science Foundation: Phase II. Washington, DC: National Academy of Sciences.

Fogel, R. W. & S. L. Engerman (1974). Time on the cross: The economics of American Negro slavery. Boston, MA: Little, Brown.

Garfield, E. (1963). Citation indexes in sociological and historical research. American Documentation, 14(4), 289-291.

Grafton, A. (1997). The footnote: A curious history. Cambridge, MA: Harvard Univ. Press.

Merton, R. K. (1938). Science, technology and society in seventeenth century England. In G. Sarton (Ed.), OSIRIS: Studies on the history and philosophy of science and on the history of learning and culture (pp. 362-632). Bruges, Belgium: St. Catherine Press.

Merton, R. K. (1957). Priorities in scientific discovery: A chapter in the sociology of science. American Sociological Review, 22(6), 635-659.

Merton, R. K. (1961, October). Singletons and multiples in scientific discovery. Proceedings of the American Philosophical Society, 105(5), 470-486.

Posner, R. (1990). Cardozo: A study in reputation. Chicago, IL: University of Chicago Press.

Price, D. J. de Solla. (1963). Little science big science, New York, NY: Columbia Univ. Press.

Quandt, R. E. (1976). Some quantitative aspects of the economics journal literature. Journal of Political Economy, 84, 4, 741-756.

Stigler, G. J. & C. Friedland (1975). The citation practices of doctorates in economics. Journal of Political Economy, 83(3), 477-508.

Stigler, G. J. & C. Friedland (1979). The pattern of citation practices in economics. History of Political Economy, 11(1), 1-20.

Stinchcombe, A. (1975). Merton's theory of social structure. In L. Coser (Ed.), The idea of social structure (pp. 3-11). Berkeley, CA: Harcourt Brace Jovanovich.

Zuckerman, H. (1967). Nobel laureates in science: Patterns of productivity, collaboration and authorship. American Sociological Review, 32, 391-403.

Zuckerman, H. (1977). Scientific elite: Nobel laureates in the United States. New York, NY: Free Press.

Endnotes

i There has been a tendency over the years to reify the meaning of citations. Citation counts are not measures of quality in and of themselves. When the measure is used as an indicator of the quality of an individual's work without examining the "meaning" of the citations to the work, particularly the type of citation and the characteristics of those who cite the work, the probability increases of drawing inappropriate inferences about the impact or quality that an individual's scholarly and scientific work has had on his specialty or field.

ii Scholars in different disciplines have compared the citations to scholars in departments of differing ranks. Some of the best earlier work can be found in economics by Stigler and Friedland (1975); Quandt (1976); Stigler and Friedland (1979), among others.

iii There were, of course, important precursors to the efforts to establish a specialty prior to the late 1950s and early 1960s. One example is the important contribution by Bernard Barber, "Science and the social order" (1952), in which Barber discusses science within the framework of a social system analysis—one which drew on the theoretical work of Talcott Parsons and Merton. Barber also published an important early reader on the subject, "The sociology of science," (1962) Bernard Barber and Walter Hirsch (Eds.).

iv See Arthur Stinchcombe's paper in a *Festschrift* honoring Robert K. Merton [Coser (1975)]. Arthur Stinchcombe brilliantly traces some of the core theoretical and methodological themes in Merton's work.

v If it were not for Merton's own extraordinary skills at crafting highly readable, even entertaining, papers that made wide use of the literatures of cognate disciplines, and thus could be enjoyed in the reading by non-sociologists, there would probably have been even more significant resistance to the growth of the specialty. In short, Merton spoke to audiences both within and outside the profession—and most particularly to historians of science and technology and scientists who were interested in the history of science.

vi Zuckerman first published her work on the laureates in 1967 in an *American Sociological Review*, 32, 391-403, paper entitled "Nobel laureates in science: Patterns of productivity, collaboration and authorship." The classic study was published as "Scientific elite: Nobel laureates in the United States" (1977)

vii The work to build the sociological study of science at Columbia was not carried out in isolation. There was an important group of individuals and in a number of cases clusters of individuals who were interested in the social organization of science and were studying it. Among the early leaders in the 1960s were Warren Hagstrom who was developing a program at the University of Wisconsin, Madison; Norman Storer, Diana Crane, Wayne Dennis, Irving Sher (from ISI), Alan E. Bayer, John Folger, and the groups of historians of science led by Price at Yale; and somewhat later by Everett Mendelson and Gerald Holton at Harvard; and Arnold Thackray at the University of Pennsylvania.

viii It has always interested Merton's former students how ironic it was that Merton the theorist was continually focusing his students on empirical tests of ideas, while his close colleague and collaborator—and the teacher of some of us—Paul F. Lazarsfeld, who was one of America's premier social science methodologists, was constantly focusing his students on the theory that would help to explain the patterns of empirical evidence that students were trying to understand.

ix The "Merton seminar" became the Merton/Zuckerman seminar after Harriet Zuckerman joined the Columbia faculty.

x The seminar, which had the atmosphere of a workshop in which ideas were exchanged, debated, and part of an on-going dialogue, involved collaborations among the various student members of the seminar. We learned much from each other and collaborations were formed among some of the early participants in the seminar that lasted for decades. Not only did I collaborate with Stephen Cole for decades on aspects of science, but Harriet Zuckerman and I began our work on women in science as an outgrowth of work in the program related to testing the norm of universalism as it affected women's careers in science.

xi In these early days, there were no electronic databases. The ISI would publish annual volumes of citations and we would have to count the citations (eliminating some such as self-citations) by hand. We had many students who worked "counting citations." This had to be one of the most boring of tasks and we were often concerned if our students would work too many hours "at the index," we might have to admit them to the psychiatric ward of Columbia Presbyterian Hospital. Today, the *SCI* is on-line and counts can be more easily done.

xii Perhaps the first study to examine correlates of citations was Kenneth E. Clark (1957, pp. 54-57). Clark studied eminent psychologists and found that measures of visibility (measures of peer esteem) had a higher correlation with citation counts to the individual (r= .67) than with other measures of productivity (bibliographic counts), with professional income, and number and quality of Ph. D. students studying with the individual. [These findings are described in Bayer and Folger (1966, p. 385)] However, Clark's work did not use the *SCI*, which was not published until 1961. Clark created his own citation counts from psychology journals, but he did not create a citation index. He describes his procedure as follows: "3. *Journal citation counts.* This index is the number of times an individual's work is cited in the published literature by other research workers. Through the use of IBM equipment it was possible without great effort to determine the number of times an individual's works were cited in articles of a research nature written in various journals." (Clark, 1957, p. 43). Clark did not include all citations in his counts. For example, he excluded self-citations and he limited counts to "those articles which presented at least some evidence of data collections and presentation ..." (*ibid.*) In short, Clark created counts similar to those that we constructed in using the *SCI.*

xiii Bayer and Folger went on to show that citation counts were correlated with the source of biochemistry doctorates. "Almost twice as many graduates of low quality departments than of the high quality ones had no citations. At the other end of the continuum, three times as many graduates of high quality biochemistry departments than of low quality ones were cited more than 15 times ... The product-moment correlation coefficient between institution quality and citation count is .21, which is significant at the .001 level" (p. 388).

xiv Eugene Garfield claimed in a 1963 paper that citation counts were measuring impact of work but not its significance or importance. See pp. 289-291.

xv Bayer and Folger, *op. cit.*, refer to the Garfield and Sher finding: "Whereas the mean number of references cited was 3.4 [for all references in the file], and the mean number of citations per author was 5.5 for the total group, for the Nobel prize winners the mean number of references cited was 58.1, and the mean number of citations per author was 169.0" [p. 385].

xvi S. Cole and J. R. Cole (1967); S. Cole (1970); J. R. Cole (1970); J. R. Cole and S. Cole (1971,1972,1973); J. R. Cole and S. Cole (1976); S. Cole, J. R. Cole, and L. Rubin (1977,1978); S. Cole, J. R. Cole, and L. Dietrich (1978); S. Cole (1975, 1979); S. Cole, J. R. Cole, and G. Simon (1981); J. R. Cole and S. Cole (1981)

xvii These scientists probably underestimated the actual impact and quality of the Lowry work (Lowry, O. H., N. J. Rosebrough, A. I. Farr & R. J. Randall (1951). Protein measurement with the folin phenol reagent. *Journal of Biological Chemistry* 193: 265-275). Citation patterns and interviews with scientists suggested that there was a fairly clear hierarchy of value associated with scientific work: theoretical contributions outranked those that were fundamentally experimental and experimental work outranked technological breakthroughs.

xviii For some of our early publications that dealt with the problems associated with citation counts, see J. R. Cole (1969, 1970); J. R. Cole and S. Cole (1971); J. R. Cole and S. Cole (1973), "Social stratification in science."