

Opinion Paper

On the Objectives of Citation Analysis: Problems of Theory and Method

B. C. Peritz

Graduate School of Library and Archive Studies, Hebrew University, Jerusalem, P.O. Box 503, 91904, Israel

Citation analysis can go beyond its present preoccupations and explore new areas if it follows several rules of research design—some of them known and generally accepted, others specific to this discipline: (a) a rigorous definition of the research objectives is essential in deciding on its design, including the selection and definition of variables and the measurement of their effects; (b) the ascertainment of content-related variables enhances the theoretical interest and practical usefulness of citation analysis, although it may entail the use of smaller samples; (c) the set of papers to be compared with respect to citation frequency should be stratified in order to make them as similar as possible to each other; (d) the dependent variable, citation frequency, may comprise more than one *kind* of citation; (e) the elementary methods of analysis based on stratification (or matching) and adjustment should be complemented by model-based methods which could accommodate larger numbers of variables and would take into account the skewness of citation count distributions. © 1992 John Wiley & Sons, Inc.

Introduction

Citation analysis as it is practiced today is, historically, a byproduct of the creation of citation indexes. These search tools, initiated in the 1960s, were often found to be more informative and efficient than conventional subject indexes (see, e.g., Garfield, 1979). However, beyond their practical value, it became clear from the outset that citations also carry some social relevance which may be of intrinsic interest (Kaplan, 1965). Following Merton's approach (see, e.g., his foreword to Garfield, 1979), it is generally accepted that the publication of research papers is part of the reward system in science and hence that citations are, in some sense,

tokens of recognition. This approach has had a decisive influence on citation analysis as a discipline.

Citations, the "raw material" of citation analysis, have been studied by many authors. Cronin (1984), in particular, has submitted the citation process to a detailed theoretical scrutiny which included a review of the role and content of citations. However, in spite of the critical reservations of Cronin and other authors, it is mostly accepted that, in general, a paper is being cited in order to make a point that is relevant to the subject at hand (Smith, 1981). It is true that one occasionally encounters citations that are casual, perfunctory or immaterial but these are not frequent; there are also instances when a study is cited in a negational mode, but these too are rare, hard to define unambiguously, and sometimes do not detract from the value of the cited publication (Moravcsik and Murugesan, 1973; Chubin and Moitra, 1975; Peritz, 1983a). In some sense, therefore, the number of times a paper is cited can be used as a rough-and-ready indicator of its merit—granting, of course, variations in the citation's importance (Brooks, 1985; Cano, 1989) and the inevitable amount of error and noise. Consequently, citations can be viewed as legitimate objects of research, and, in fact, citation analysis has often proved itself to be a meaningful research tool [an extensive and balanced account of these issues can be found in part 3 of Egghe and Rousseau (1990)].

Although legitimate and meaningful, citation analysis is not devoid of conceptual and methodological difficulties. The nature of these difficulties, which tend to vary with the specific objectives of the study, is the subject of the present remarks.

In a slight paraphrase on Zunde (1971), one may distinguish three application areas of citation analysis: the evaluation of scientists, publications, and scientific institutions; the investigation of hypotheses concerning the history and sociology of science and technology; and the study of the performance characteristics of information search and retrieval procedures. The last area, which is

Received August 8, 1991; revised December 11, 1991; accepted February 7, 1992.

© 1992 by John Wiley & Sons, Inc.

in a category somewhat apart from the others, will not be considered here. The investigation of the clustering of papers by means of some citation-based link (Kessler, 1963; Small, 1973), which is part of the second of these areas of application, will only be mentioned briefly. Thus, the main body of this article deals with the *determinants of citation frequency*, whether the study is performed with an evaluative purpose or in the pursuit of some question concerning the nature of scientific research.

Reference Sets and Comparisons

To what extent was a given published innovation recognized, taken up, elaborated upon, or discussed in the scientific community? In order to answer this type of question it is clearly not enough to compile the absolute number of citations to the relevant publications and to compute the average number of citations per paper. Such averages need to be compared to the averages obtained for *other* papers dealing with “similar” subjects and published in “similar” journals at about the same time. The selection of a suitable set of such *control* or *comparison* papers (sometimes called reference standard selection) has been discussed by Narin (1976), Garfield (1979), and particularly by Braun and his team (e.g., Braun, Glanzel, and Schubert, 1985, 1987; Schubert, Glanzel, and Braun, 1988). A simple but somewhat crude solution to this problem exists in selecting controls that are *matched* to the *propositus* papers by journal, date of publication, and, in the case of multidisciplinary journals, specific discipline. A more refined approach, based on bibliographic coupling, has also been suggested by the above-mentioned authors.

Although adequate in most cases, the above solutions are not really satisfactory in a number of important instances. The underlying reason for this stems from the well-known fact that, in many fields, the community of *producers* of research results is different from the community of their *users*. As a consequence of this disparity, journals in which results from a given research area are published may be different from the journals in which they are cited. Thus, new methods of research in some discipline may be published in one kind of journal while their applications appear in a different set of periodicals. Moreover, citation of a study because of its connection with the subject matter of the citing paper may be qualitatively different from a citation indicating its *use* or *application*. In other words, citation of a paper because of some feature of its contents is indicative of a different *kind* of value than a citation which indicates the use of its results. The weights to be allocated to each one of these two kinds of value will depend on the specific objectives of the citation analysis.

The following example may help clarify the issue. Some years ago it was shown (Peritz, 1983b) that, in sociology, methodological papers are cited more often than theoretical or empirical ones. Now, suppose that,

within the framework for some evaluation process, one is asked to compare two sociology departments to each other. The two departments are of equal size and exhibit similar numbers of citations but most of department A's citations are of methodological papers while department B is cited mostly for its empirical work.

Which one of the two departments will be rated higher in this evaluation process? If one believes that the development of strong methodological tools is essential to the healthy growth of the discipline it will have to be department A. If, on the other hand, one feels that the test of a department's worth should be its contribution to the subject matter of sociology, it will be department B.

The dilemma faced by the evaluator in this example is of a philosophical nature. Once this dilemma is settled, that is, once the objectives of the evaluation are formulated in a sufficiently precise manner, the conclusions of the citation analysis will be straightforward. The presentation of its results, however, will have to deal separately with the citations of the two kinds of source papers.

It is seen, therefore, that the dependent variable “number of citations” will often have to be subdivided into two parts: citations referring to the use or application of the source paper's method, and all other citations. One may achieve this in one of three manners: (a) through a classification of the *roles* of the citations themselves in the citing papers; (b) through a classification of the source papers as in the example above; and (c) according to a classification of the citing journals. Procedure (a) is evidently the most accurate but also the most laborious, while procedure (c) is both the easiest and the least accurate. These procedures have one feature in common: they require a content-related classification of the material citations, papers, or journals, respectively.

Comparison and Evaluation

The picture of the “typical” citation analysis as it emerges from the previous section is more or less as follows: it investigates two sets of papers, “propositus” and control, made similar with respect to the publishing journal (or group of journals) and their dates of publication. There are either one or two *dependent* variables to be studied for each paper: *either* the *overall* number of citations made in a specified time period *or* the number of citations indicating use or application on the one hand and all other citations on the other.

The independent variable in this schematic description is, of course, the dichotomy: *propositus* versus control papers. Another variable is less conspicuous in this description: it is the stratifying variable which causes us to compare only “like with like” in our analyses, i.e., the specific journal or the homogeneous group of journals in which both the *propositus* and the control paper were published. The technique whereby we adjust for

the stratifying variable in the estimation of the effect under study—the effect to which the difference between propositus and control is ascribed—is beyond the scope of this article, but some remarks on stratification and modeling will be made in the next section.

There is, of course, no need to assume that the main independent variable must necessarily be dichotomous. What we have called the dichotomy between propositus and control may well be replaced by a discrete variable of three or more categories and even by a continuous quantitative variable (although the assumptions made in the latter are again outside the purview of these remarks). If the objective of the study is evaluative, one will only rarely have to consider more than one or two independent variables with two or three categories. In recent years, however, there is an increasing interest in citation analyses carried out with other objectives in mind: objectives concerned with some aspect of the nature of scientific research or with the development of citation analysis itself as a discipline. In these instances, it will become necessary to investigate simultaneously several independent variables.

The following example may serve to illustrate this point.

In a recent analysis of clinical trials (Peritz, 1990), it was found that trials that were well-designed, i.e., randomized, carried out by a double-blind procedure, and with a relatively large number of subjects, tended to be *less* cited than those that lack these characteristics. The tentative interpretation offered was that a study suggesting new ideas and approaches, even if it is only preliminary, is more likely to be cited than a *definitive* treatment of the question. The *heuristic value* of the former type of study makes it more citable while there is little reason to cite the latter study, regardless of its intrinsic value. Science simply regards the problem as solved and moves on to other matters.

In order to reach the above conclusion—however tentative—it was necessary to “control” not only for the journal (there were two: the *New England Journal of Medicine* and *Lancet*) and the date of publication but also for the three dichotomized variables mentioned above; i.e., randomized or not, double-blind or not, large number of cases or not. Thus, one had to look (jointly as far as possible) at the effects of *three* independent variables while adjusting for two others which pertain to the choice of the reference standard: the journal and the year of publication.

The three independent variables mentioned in the preceding paragraph are all *content-related*. The inclusion of such variables in a citation analysis raises major issues of ascertainment. It assumes that the source papers were scanned with regard to variables pertaining to their respective categories. It is realized that one thereby restricts oneself to citation studies with relatively small numbers of cases, but this may be the price one has to pay in order to widen the range of objectives of citation analysis.

A Remark on Methods of Analysis

The methods of assessing the effect of a so-called independent variable on one or two dependent variables while controlling for other variables—the statistical gist of citation analysis—are beyond our present scope. Briefly, the remarks of previous sections about stratification and adjustment are readily made rigorous with only a minimum of mathematical notation; it is, however, important to realize that there is at least one other strategy for achieving the above purpose: the use of a statistical model.

A statistical model is a formalized set of assumptions on the behavior of certain (dependent) random variables in the presence of other (independent) variables. One important instance is the so-called “linear model,” which assumes a linear relationship between dependent and independent variables and a normal distribution of chance error in the former. The linear model is easy to use since most software packages deal with it exhaustively. Unfortunately, this model is not applicable to citation counts, mainly because of the skewness of these variables. Logistic regression, another model which assumes that the dependent variable is dichotomous, is equally unsuitable. Other models are applicable but are difficult to implement. On the whole, it would seem that in order to apply statistical models in citation analysis correctly and conveniently, one would need a substantial methodological input from statisticians and informetricians.

In the meantime, stratification, adjustment and the nonparametric methods associated with these techniques are available and easy to use. They may sometimes be clumsy and they may not allow for the simultaneous inclusion of many independent variables, but they will suffice until more suitable parametric models become available.

Conclusions

One basic point which emerges again from the above remarks is the well-known fact that a rigorous definition of a study's objectives is crucial in deciding on its design. Basic questions—like the number of dependent variables (the kinds of citation counts) to include in the analysis, whether a weighted average of these variables should be produced, what stratifying and other independent variables should be included, and how their respective effects should be measured—all these are readily answered once an unambiguous definition of the objectives is given.

The second point is that most truly interesting research objectives require the ascertainment of content-related variables, generally on the source papers, but occasionally on the citations as well. Such ascertainment tends to be time-consuming and is sometimes conceptually difficult; it also limits the studies to small numbers of cases. However, on the other hand, it adds flavor and

interest to the practice of citation analysis and widens the range of hypotheses it can encompass.

Third, the methodology of citation analysis is still burdened by unsolved questions. The elementary methods of stratification (or matching) and adjustment are adequate for many studies, but there is still an urgent need for a statistical model that should accommodate a fairly large number of independent variables in a manner that could be easy to implement and that would take into account the inherent skewness of citation counts.

Citation analysis is now ready to break new ground. In some areas, like cocitations, the traditional methods are adequate, although, perhaps, some input from cluster analysis may be useful. In other areas, a slight change of orientation along the lines sketched in the above remarks may come in handy.

References

- Braun, T., Glanzel, W., & Schubert, A. (1985). *Scientometric indicators. A 32-country comparative evaluation of publishing performance and citation impact*. Singapore: World Scientific.
- Braun, T., Glanzel, W., & Schubert, A. (1987). One more version of the facts and figures on publication output and relative citation impact of 107 countries, 1978–1980. *Scientometrics*, 11, 9–15.
- Brooks, T. A. (1985). Private acts and public objects: An investigation of citer motivations. *Journal of the American Society for Information Science*, 36, 323–329.
- Cano, V. (1989). Citation behavior: Classification utility and location. *Journal of the American Society for Information Science*, 40, 284–290.
- Chubin, D. E., & Moitra, S. D. (1975). Content analysis of references: Adjunct or alternative to citation counting? *Social Studies of Science*, 5, 423–441.
- Cronin, B. (1984). *The Citation process*. London: Taylor Graham.
- Egghe, L., & Rousseau, R. (1990). *Introduction to informetrics*. Amsterdam: Elsevier.
- Garfield, E. (1979). *Citation indexing — its theory and application in science, technology and humanities*. New York: Wiley.
- Kaplan, N. (1965). The norms of citation behavior—prolegomena to the footnote. *American Documentation*, 16, 179–184.
- Kessler, M. M. (1963). Bibliographic coupling between scientific papers. *American Documentation*, 14, 10–25.
- Merton, R. K. (1979). Foreword. In: E. Garfield *op. cit.*
- Moravcsik, M. J., & Murugesan, P. (1975). Some results on the functions and quality of citations. *Social Studies of Science*, 5, 86–92.
- Narin, F. (1976). *Evaluative bibliometrics: The use of publication and citation analysis in the evaluation of scientific activity*. Cherry Hill, NJ: Computer Horizons.
- Peritz, B. C. (1983a). A classification of citation roles for the social sciences and related fields. *Scientometrics*, 5, 303–312.
- Peritz, B. C. (1983b). Are methodological papers more cited than theoretical or empirical ones? The case of sociology. *Scientometrics*, 5, 211–218.
- Peritz, B. C. (1990). A citation analysis of clinical trials; are definitive studies less cited than others? *Informetrics*, 89/90, 243–250.
- Small, H. (1973). Co-citation in the scientific literature: a new measure of the relationship between two documents. *Journal of the American Society for Information Science*, 24, 265–269.
- Smith, L. (1981). Citation analysis. *Library Trends*, 30, 83–106.
- Schubert, A., Glanzel, W., & Braun, T. (1988). Against absolute methods; relative scientometric indicators and relational charts as evaluative tools. In: A. F. J. Van Raan (Ed.), *Handbook of quantitative studies of science and technology*. Amsterdam: North Holland, pp. 137–176.
- Zunde, P. (1971). Structural models of complex information sources. *Information Storage and Retrieval*, 7, 1–18.