

Restating the Fundamental Assumptions of Citation Analysis

Number 39, September 26, 1977

It is unfortunate that many people who use citation data in sociological studies fail to explicitly state their assumptions. Most of these assumptions are by now quite obvious to research workers in science policy studies. But even for those familiar with the methodology of citation analysis it is still too early in the field's development to take certain assumptions for granted. Authors who do not state their assumptions risk later criticism for their omission.

The two reprints which follow, "On the Use of Citations in Studying Scientific Achievements and Communication" by Belver C. Griffith, M. Carl Drott, and Henry G. Small,¹ and "Why I Am Not a Co-Citationist" by David Edge,² discuss some of the fundamental assumptions underlying citation analysis and co-citation studies.

The first article is a concise statement of the assumptions made in most citation analyses. The authors also discuss three "massive" qualifications on the use of citation data. The three qualifications concern the quality of the *Science Citation Index*[®] and *Social Sciences Citation Index*[™] files, the citation behavior of authors,³ and the scale and pace of research activity in the discipline

being studied. This article is a thoughtful response to critics of citation studies.

Both papers take a critical view of citation analysis. The first paper exhibits a healthy, constructive skepticism. The second takes a more pessimistic view of the sociological value of studying citation relationships. But neither paper questions their validity for information retrieval.

Edge admits that his perspective "makes research more difficult than it would be under more simple-minded premises, but then I happen to believe that sociology *is* difficult." No one—least of all citation analysts—is saying that sociology is easy. But I see no reason to make it even more difficult by excluding a whole realm of valuable research work from consideration.

Edge also mentions the "relatively trivial behavior of adding citations to papers." Citation behavior may not be the most complex human behavior one could study, but even when it is abnormal, it certainly is not trivial.

Finally, Edge advances an argument that, if taken to its logical conclusion, would mean the elimination of all sciences which rely upon quantitative and statistical analy-

ses. He claims that "citation and co-citation analysis, in striving to *accumulate and average*, destroys the evidence we need of individual variations." There is no reason that quantitative and qualitative studies should be mutually exclusive. Subjective studies of individual variations can peacefully co-exist with objective studies of aggregates.

It is strange that neither paper discusses the one fundamental assumption that I know disturbs many people—and justifiably so. It is often assumed, because of the first-author arrangement of ISI®'s citation indexes, that *all* conclusions derived from such first-author data will be wrong. There is little doubt that citation analyses based on first-author data inadequately recognize some authors who now publish mainly as "secondary" authors. This was not true for authors who published before the fifties, when single authored papers were the rule. The best way to eliminate this inadequacy is to compile data based on all co-authorships. In fact, ISI is now completing an intensive study whose aim is to eliminate the first-author artifact. When this study is completed some of the results will be published here. Meanwhile, we

will soon publish a list of the 250 most-cited first authors.⁴ For this kind of study the problem of "noise" caused by errors is of minor consequence. These 250 authors are so often cited that such errors rarely have an observable effect.

Some people are also disturbed by the use of citation data to indicate areas of activity and merit in science. However, criticism which does not offer constructive alternatives has little value. Those who question the use of citation data as "indicators"⁵ of scientific activity and accomplishment should offer alternative cost-effective means for dealing with the problem. Finding significant information about the immense world-wide scientific enterprise is not easy. There is considerable confusion in science about where the action is and where it ought to be, and citation analysis is a valuable tool for reducing the confusion.

It is in the interest of balance and fairness that these two papers are reprinted. Together they demonstrate that in citation studies—as well as in the rest of science—it is wise to take little or nothing for granted.

REFERENCES

1. Griffith B C, Drott M C & Small H G. On the use of citations in studying scientific achievements and communication. *Society for Social Studies of Science Newsletter* 2:9-13, Summer 1977.
2. Edge D. Why I am not a co-citationist. *Society for Social Studies of Science Newsletter* 2:13-9, Summer 1977.
3. Garfield E. To cite or not to cite: a note of annoyance. *Current Contents*® No. 34, 22 August 1977, p. 5-9.
4. ———. And who shall occupy the 250th chair among the citation immortals? *Current Contents* No. 22, 31 May 1976, p. 5-6. (Reprinted in: *Essays of an Information Scientist*. Philadelphia: ISI Press, 1977, Vol. 2, p. 496-7.)
5. Garfield E, Malin M V & Small H. Citation data as science indicators. *Toward a Metric of Science*. (Elkana Y. et al., eds.) New York: Wiley, 1977. (in press).