Citation Analysis: A New Tool for Science Administrators

A new way to assess scientific productivity is about to come into its own. Citation analysis, hitherto an arcane tool of historians and sociologists of science, has now been refined to the point where it offers increasingly interesting possibilities to the science administrator. Proponents of the technique believe that within a few years it will find major uses in decisions at the level of national science policy, as an adjunct of the peer review process, and in evaluating the performance of individual scientists.

The starting point of all citation analysis studies is to count the number of times an article or author is cited in the scientific literature. On the general assumption that the number of citations reflects an article's influence, and therefore quality, this measure can be made to serve as the fundamental yardstick for quantifying many aspects of the cognitive and social structure of science. The basic information about who has cited whom is readily obtained from a quarterly publication, the *Science Citation Index*, which lists all the citations in some 2600 of the world's most often cited journals. Under each scientist's name appear the titles of his cited articles and of the articles citing them during the period covered. (In fact, only half of the scientific articles published are ever referred to in the scientific literature, and the average cited paper is only cited 1.7 times a year.) The technique's range of application can be seen in some of the uses already in practice or contemplated.

- Citation analysis data are being offered in a court case to prove that a woman denied tenure is as good or better than two men promoted over her.
- Citation analysis is being relied on at certain universities, such as the State University of New York, as part of the evidence for deciding cases of promotion and tenure.
- The National Science Foundation (NSF) is using the technique to assess its funding of chemistry departments and as a safety net to catch chemists who write bad grant proposals but are heavily cited.
- Citation analysis has been used in a National Academy of Sciences study showing that postdoctoral fellows supported by the National Institute of General Medical Sciences are more often cited than comparable scientists without postdoctoral training.
- The suspicion that government science policy toward universities is dominated by a coterie of "Eastern Establishment" scientists has been tested by sociologist Warren Hagstrom of the University of Wisconsin and found generally wanting; slightly more such scientists sit on government committees than would be expected by merit alone (as judged by citation analysis), but the difference is not statistically significant.
- Citation analysis could have a "very important military intelligence significance," according to Morton V. Malin, vice president of the Institute for Scientific Information, which publishes the *Science Citation Index*. Malin believes the technique could be used to identify the gaps in a country's published research and to infer the kinds of technology being pursued in these classified areas.

Counting footnotes may seem a simpleminded way of estimating anything, let alone characteristics as subtle as scientific quality and intellectual influence. But there are several reasons why citation analysis should not be taken lightly. For one thing, it clearly describes something real about the scientific world, as can be seen by a glance at the list of the 50 most cited authors (Fig. 1), 12 of whom have now won Nobel Prizes, or at the map of biomedicine in 1973 (Fig. 2).

For another, granting agencies are under increasing pressure from Congress to do more evaluation of their programs. The NSF and National Institutes of Health (NIH) are keenly interested in the technique because it is demonstrably objective (peer review has the unavoidable appearance of a buddy system); moreover it seems likely at present showing to corroborate the decisions made by the embattled peer review system. The NSF has indicated receptivity to a scheme for setting up a citation analysis data bank at a cost of some $250,000 a year.

Articles, of course, are often cited for reasons other than their importance, influence, originality, or merit, and every author can think of a dozen such reasons why citation analysis should not work. He may cite a paper because it was written by his professor, or because he disagrees with it, or because it describes some useful method. But to citation analysts, such references are simply the "noise" in the system, which they believe they can for many purposes filter out or at least reduce to insignificance. In most, but not all, attempts at validation, citation counts are found to correlate highly with almost every conventional measure of scientific quality.

Though the significance of citation counts is usually the first objection raised against the method, citation analysts consider this issue to be largely resolved. The technique has now progressed to other problems, such as the limits of its applicability. Many studies to date, for example, have been confined to physics; it is not known how confidently the results can be extrapolated to other disciplines.

Another issue that sharply divides the 30 or so people who are active in the field is the extent to which the

NIH, HEW Nominees

As expected, President Ford has nominated Theodore Cooper as assistant secretary for health of the Department of Health, Education, and Welfare and Donald S. Fredrickson as director of the National Institutes of Health. Cooper, a surgeon and physiologist with long experience at NIH, has been acting assistant secretary; Fredrickson, also an NIH alumnus, is currently president of the Institute of Medicine.

The two nominations had been pending for weeks, held up, apparently, by conservative Republican objections to Cooper, a Democrat (*Science*, 18 April). Quick Senate confirmation of both nominees is expected.—R.G.

technique can fairly be used to evaluate individuals. Some sociologists of science, such as Jonathan Cole of Columbia, believe firmly that the technique is valid when applied to large aggregates but is not so fine a measure that it can be the yardstick of an individual's performance. Others, such as Derek de Solla Price of Yale, argue that citations are an accurate measure of individual quality. A practical test of the issue may be furnished by an imminent court case in which a biochemist denied tenure at an eastern university alleges she is a victim of sex discrimination. Her case has been championed by Robert E. Davies, a physiologist at the University of Pennsylvania. Together with two colleagues in the operations research department, Nancy L. Geller and John S. De Cani, he has devised a way of estimating the "lifetime citation rate" of a scientific article, based on its citations to date. The articles published by the untenured biochemist turn out to have an expected lifetime citation rate of 53.5 times per article. The chairman of her department is a 51.4 per paper man. The two men who received tenure at the same time she was denied it have lifetime citation rates of 21.8 and 50.9 per paper, respectively.

From this and other citation measures the Pennsylvania analysts conclude that, on the whole, the woman researcher's work "is significantly better in quality than that of the two men who were promoted to associate professorships with tenure and fully comparable to the full professors in her department." But these results drew a storm of criticism when presented at a recent conference of citation analysts, attracting such descriptions as "dangerous dribble," "frightening," and "premature." Davies' position is that citations are an excellent indicator of quality because they "represent the integrated peer review of everyone in the field." Like other indicators of quality they can be misused but, with proper precautions, the lifetime citation rate he and his colleagues have developed "ought to be useful in promotion and tenure considerations."

The Pennsylvania team's work has the support of historian Derek de Solla Price, who has been an influential proponent of bringing quantitative measures to the study of science ever since his *Little Science, Big Science*, published in 1963. Price, a respected figure in the field, goes further in his claims than most citation analysts. He considers that within a few years citation analysis will be used as an adjunct to peer review. As clinching evidence of the method's accuracy, he cites a not yet published study which shows that two assessments of scientific articles by peer review correlate more closely with an assessment by citation analysis than they do with each other.*

The scientific community, Price predicts, will at first resist the outside evaluation represented by citation analysis but will accept the technique when they find it corroborates the judgment of peer review and their own beliefs about who is good and who is bad.

RANK	AUTHOR	TOTAL TIMES CITED
1	LOWRY OH	2921
2	CHANCE B	1374
3	*LANDAU LD	1174
4	BROWN HC	1150
5	*PAULING L	1063
6	†GELL-MANN M	942
7	COTTON FA	940
8	POPLE JA	933
9	BELLAMY LJ	906
10	SNEDECOR GW	904
11	BOYER PD	893
12	BAKER BR	876
13	KOLTHOFF IM	853
14	†HERZBERG G	842
15	FISCHER F	826
16	SEITZ F	822
17	DJERASSI C	801
18	BERGMEYER HU	754
19	WEBER G	750
20	REYNOLDS ES	748
21	MOTT NF	741
22	*ECCLES JC	737
23	FEIGL F	729
24	FREUD S	727
25	PEARSE AGE	726
26	ELIEL EL	721
27	STREITWIESER A	717
28	*MULLIKEN RS	712
29	*JACOB F	711
30	*BORN M	710
31	BRACHET J	706
32	WINSTEIN S	702
33	ALBERT A	687
34	LUFT JH	674
35	†DEDUVE C	673
36	†VON EULER US	668
37	FIESER LF	666
38	HUISGEN R	661
39	NOVIKOFF AB	655
40	GOODWIN TW	643
41	†BARTON DHR	632
42	FISHER RA	631
43	BATES DR	627
44	†FLORY PJ	626
45	STAHL E	626
46	DEWAR MJS	619
47	GILMAN H	618
48	FOLCH J	618
49	DISCHE Z	614
50	GLICK D	609

Fig. 1. Nobel Prizes by citation analysis. The list above, compiled by Eugene Garfield of the Institute for Scientific Information, shows the 50 most cited authors for 1967. Six authors (*) had already received the Nobel Prize when the list was compiled, another six (†) have received it subsequently.

On the level of national science policy, the United States, with its steadily declining share of the world's scientific output, could, in Price's view, take advantage of citation analysis to allocate its research funds more efficiently.

Price's confidence in the technique's future is not universally shared and, indeed, is alarming to some. Remarks an NSF official, "I am not looking forward to the day when Senator Proxmire's assistant can get a printout from the *Science Citation Index* and says, 'I see there has been a breakthrough in this field—I can't quite pronounce it —and why haven't you put half your money into it?'" Wayne Gruner, assistant to the NSF's director of research, says that the foundation "would be willing to spend a good deal of money to find out if this is a practical tool. But it would be a grave mistake for people to assume it has already been demonstrated that you can measure scientific productivity by this method."

Even if it were justifiable to use citation analysis in allocating funds, Gruner says, the attempt might be self-defeating because people would find ways of exploiting the system to their own advantage, for example by mutual citing arrangements, which would ruin its use even as a research tool. Like NSF, NIH also has let several small contracts designed to evaluate the technique's possibilities. Helen Gee, chief of the NIH program analysis and evaluation branch, hopes to use citation analysis to "demonstrate to the Office of Management and Budget that decisions made by peer review groups are not made on the buddy system but are valid assessments of merit." Another application might be to study the interdependence of NIH's various institutes. If, for example, the National Cancer Institute were found by citation analysis to be drawing heavily on the work of other institutes, that would be an argument for ceasing to increase the NCI's budget at the others' expense.

A new development with interesting possibilities for science policy lies in the definition and mapping of the relationship between scientific specialties. Developed by Henry Small of the Institute for Scientific Information, Philadelphia, and Belver Griffith of Drexel University, the technique identifies

* "A statistical procedure for evaluating the importance of scientific papers" (unpublished summary of a Ph.D. dissertation by Julie A. Virgo, Graduate Library School, University of Chicago).

"clusters of highly interactive documents in science." The basic measure of interaction is simply the frequency with which two papers are cocited (included in the same list of references) within the period of interest. Applying this test for association to the *Science Citation Index*, Small and Griffith find that articles can be grouped by cocitation links into clusters which, they believe, "represent the scientific specialties which currently exhibit high levels of activity."

One proof of this contention is that titles of articles assigned by the computer program to a given cluster exhibit an evident relationship to each other, enabling the cluster to be designated "plate tectonics," or "messenger RNA," as the case may be. Another is that clusters do not remain constant but dissolve and regroup from year to year in a pattern quite suggestive of the forward march of science. The cluster map of biomedicine in 1972 was dominated by separate specialties having to do with reverse transcriptase and chromosomes. By 1973 (see Fig. 2) these had coalesced into a single supercluster on viral genetics. The 1972 cluster on microtubule protein emerged in 1973 into the important new specialty labeled "Muscle: myosin and cytochalasin-B."

It is in a way surprising that so simple a measure of association can generate so evocative a map. Small and Griffith have not yet fully tested the correspondence of their cluster maps to reality, but their hypothesis is that the clusters, rather than conventional disciplinary boundaries, represent the true intellectual and social infrastructure of science. "Whatever their physical reality, maps of science are certainly useful as heuristic tools," say Small, Eugene Garfield, and Morton Malin in a recent paper. Such tools, they suggest, can serve to test historians' and philosophers' ideas about the development of science. More practically, the maps should help in identifying active research fronts and in distinguishing areas of science that are either neglected or overfunded in relation to societal goals.

Garfield is the president of the Institute for Scientific Information (ISI), the commercial organization which publishes *Current Contents* as well as the *Science Citation Index*. The index was originally developed as a purely bibliographic tool for finding how a particular article or idea has been followed up in the subsequent literature. At a cost of $2,850 a set for the 1975 edition, ISI has an obvious interest in promoting other uses of the index, but equally the commercial viability of Garfield's ideas indicates that they are not without foundation. ISI has already developed a fully computerized system for producing annual cluster maps from the *Science Citation Index*. The institute is now planning to make the system available to teletype and remote terminal access. A software package will enable a researcher to have specialty maps displayed to him at the desired scale, ranging from superclusters to individual papers. Garfield also envisages citation data being used "to establish the relative standing of individuals or institutions," though for purposes not of grading them but of assessing their impact in the field. At a conference held by ISI last month at Elkridge, Maryland, a senior NSF official indicated that the foundation was prepared to consider proposals from the citation analysis community to have the *Science Citation Index* put "on-line" by ISI and made available to academic users.

Despite the general expectation that citation analysis will tend to back up peer review, it is unlikely that all possible applications of the technique will fall within the interest—as narrowly construed—of the scientific community. For one thing, it would render the community open to varying degrees of evaluation by outsiders. For another, some citation analysis studies might point toward painfully radical redistributions of scientific resources. An egregious example is the article by Jonathan R. Cole and Stephen Cole entitled "The Ortega hypothesis" [*Science* **178**, 368 (1972)].

The Ortega hypothesis, which the Coles seek to disprove, holds that science is somewhat akin to a bee colony in that every laborer in the vineyard, however modest his powers, contributes his little bit toward progress. The Coles believe that half the laborers may be drones, not workers. From a study of citation patterns among American physicists, they find that the best scientists (as judged by citation rates) tend almost exclusively to cite each other's work, while the work of the average researcher "is rarely the work that is influential in the production of high impact scientific research." Even in a high quality journal such as *Physical Review*, 80 percent of the articles were being cited seldom or not at all

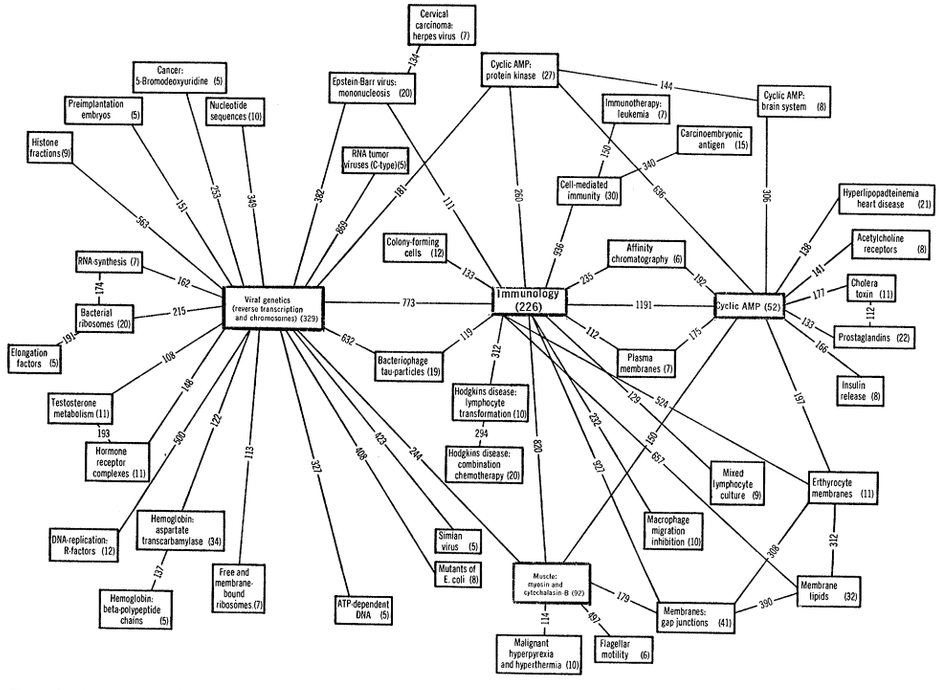

Fig. 2. Map of biomedicine constructed by Henry Small and Belver Griffith from the 10,000 documents cited 15 times or more in the 1973 *Science Citation Index*. Application of the clustering program produced several groups of articles, including a 1000-strong group with obvious relation to biomedicine. The map shows this group displayed in its interlinked clusters according to the threshold criterion that, to be a member of a cluster, an article must be cocited with one or more other members at least 11 times. The numbers on the lines show the sum of cocitations between documents in connected clusters. [Map taken from article by Steve Aaronson, Institute for Scientific Information, in *Mosaic*, March 1975]

within 3 years after publication, and were presumably making little impact on the development of science. "Thus the basic question emerges," conclude the Coles, as to "whether the same rate of advance in physics could be maintained if the number of active research physicists were to be sharply reduced. . . . We suggest that it may not be necessary to have 80 percent of the scientific community producing 15 or 20 percent of the work that is used in significant scientific discoveries, if perhaps only half their number could produce the same work."

The impact of citation analysis on the scientific community cannot yet be assessed because all that has really been demonstrated so far is promise, not practicability. The NSF and NIH have political reasons for pressing ahead with the technique, although their interest, it should be noted, predated the present bout of congressional hostility. On the other hand, both the agencies and most sociologists of science are afraid that the technique, having taken so long to reach the threshold of acceptability, could be set back disastrously by a few premature uses. Citation analysis is therefore likely to be introduced rather cautiously. Nevertheless, for those who wish to influence the way the technique is applied to the practice of science, the time is probably now.—NICHOLAS WADE

Stanford Campus Debates Fate of Student Program

The Stanford University campus has been aroused by protests lately, not over the familiar issues of war in Southeast Asia or military research sponsorship, but over the proposed merger of a popular independent study program with the regular academic departments.

On 14 January, James L. Gibbs, Jr., dean of undergraduate studies, announced he would recommend to the administration that SWOPSI (Stanford Workshops on Social and Political Issues), merge with the rest of the university after June 1976 as a cost-saving measure. Since then, students have demonstrated with a mock funeral mourning the "death of underground education," inveighed against the merger in the columns of the Stanford *Daily*, the campus newspaper, and participated in a raft of faculty-student committees that are reviewing Gibbs' recommendation.

"The students have come out for this thing in the way they haven't come out for anything in years," comments one observer.

Gibbs' recommendation and the controversy it has aroused are part of a national trend in universities to cut back on para-academic activities in the interest of saving money. Stanford, like many other universities, agreed to the establishment of such programs in the late 1960's as constructive channels for then-rampant student political activities. But now, faced with growing budget deficits, universities must decide which of these "relevance" activities are most peripheral to their primary purposes, and hence which can be eliminated or consolidated to reduce costs.

Like other prominent universities (*Science*, 14 March 1975), Stanford is engaged in a major belt-tightening effort to trim $10 million from its $70 million yearly central operating budget over the next 3 years.

The SWOPSI program, which will enroll 700 undergraduates for course credit in the current academic year at a cost of $26,231 in university funds, is perhaps the best known of such "relevance" programs. It received national publicity for a hard-hitting 1971 study of Pentagon-sponsored research on the Stanford campus (*Science*, 25 February 1972 and 22 November 1974); and it is known locally for a series of workshop reports that investigated a local air pollution control board, the impact of a proposed rapid transit plan, and destructive logging practices in the San Francisco Bay region.

The program operates with a full-time staff director. Students suggest workshop ideas; a faculty-student policy board approves them; then the student, with SWOPSI's aid, finds both a regular faculty member who will sponsor the workshop as well as a workshop leader—usually not on the faculty—who will contribute his or her time. The goal of each workshop is a report that is published independently of the university. Workshop leaders and students, after completing the studies, have sometimes been invited to serve on decision-making boards or to testify on subjects which they have researched. Some of the reports have become popular reading; one, on transportation, titled *Ride On!*, still sells in area bookstores.

Gibbs maintains that the SWOPSI workshops can be picked up individually by the academic departments without losing their character because the program, labeled an experiment, has been so successful. He stresses that he has never recommended the termination of SWOPSI, although he has recommended that two other experimental programs under his purview be ended to save money.

Defenders of an independent SWOPSI claim that if the ideas for workshops and the non-faculty workshop leaders had to pass muster according to departmental standards, the workshops would lose their current character and appeal. Charles Drekmeier, a political science professor, says, "The idea of SWOPSI is that there are a lot of experts running around who do not have Ph.D.'s and students should have the benefit of their expertise." Dan A. Lewis, a former director of SWOPSI, believes that the program will be transformed beyond recognition if Gibbs' plan goes through.

The current SWOPSI director, Andrew Parnes, also objects to the recommendation. In a letter to the Stanford *Daily* he estimated that instead of saving $25,000 the move to the departments could cost Stanford $300,000 yearly.

Some SWOPSI studies have offended members of the Stanford faculty and one claim is that Gibbs' recommendation is a form of political retaliation. (One SWOPSI report on faculty-student relations, for example, critized the performance of Gibbs' office.

"It's not political retaliation," counters Gibbs. "It's simply a matter that the university is in a difficult financial situation."—DEBORAH SHAPLEY