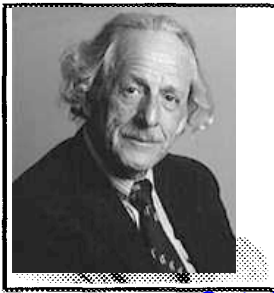
Essays of an Information Scientist: Science Reviews, Journalism, Inventiveness and other Essays, Vol:13, p.1, 1990 Current Contents, #1, p.3-10, January 7, 1991



Current Comments®

EUGENE GARFIELD
INSTITUTE FOR SCIENTIFIC INFORMATION®
3501 MARKET ST., PHILADELPHIA, PA 19104

The Impact of Citation Indexes on Biochemists and Sociologists—A Survey by L. Hargens and H. Schuman. Part 1. Methodology

Number 1 Go to Reprint January 7, 1991

Beginning with this first issue for 1991, there is a change in the editorial schedule of *Current Contents* ® (*CC* ®). *Citation Classics* ® and the *ISI* ® *Press Digest*, including *Hot Topics*, will now appear every other week. They will alternate with either an essay by me or an invited guest or a reprint with an appropriate introduction. An example of the latter appears in this issue.

During a recent trip to Eastern Europe, I observed a wrenching exercise in progress in several countries where the national science academies have employed most research scientists for the last 40 years. How, for example, does an Eastern European country determine which of its research scientists should continue to be supported? Now that party considerations have been reduced, if not eliminated, some concrete indicators are needed to evaluate these people—much like an academic tenure exercise at an American university.

As I've discussed at length,[1-3] faculty evaluation is a complex business, and, while the situation in Eastern Europe is unique, it has many relevant parallels. The introduction of international peer review systems should do much to separate the chaff from the wheat. But since so many positions need to be evaluated rather quickly, there will be a tendency to seek unobtrusive means of evaluating research performance.

Indeed, the late science historian Derek J. de Solla Price believed that the distinction between a research scientist and others with scientific training was that a research scientist is one who regularly publishes. Half the scientists holding research positions in Eastern Europe probably do not meet this criterion.

For each of those in the relatively large group that remains, one may then reasonably ask, "How often and where was this scientist's work cited?" Even in the harshest of the Communist regimes, some brilliant scientists functioned and produced work of international caliber. Others remained prominent in their own countries, perhaps because they concentrated on national interests involving, for example, earth or soil sciences.

All this is by way of reminding *CC* readers that the use of citation analysis for job performance evaluation has increased significantly. Indeed, not until I read the article by Lowell L. Hargens, University of Illinois, Champaign-Urbana, and Howard Schuman, University of Michigan, Ann Arbor, from the journal *Social Science Research*,[4] did I realize the extent to which citation indexes were used for this purpose.

In their article, Hargens and Schuman inform us that, in fact, the vast majority of biochemists and sociologists they surveyed had used the *Science Citation Index* ® (*SCI* ®) or the *Social Sciences Citation Index* ® (*SSCI* ®) primarily for information retrieval, though respondents also noted that it is used for tenure evaluations. Unfortunately, the simple questionnaire they used did not make clear how the *SCI* was used in these evaluations. As I have recommended myself, the *SCI* is invaluable for providing a quick bibliographic purview and impression of a scientist's work.

I think their study is quite revealing and should eliminate any doubt about the utility of the *SCI*/*SSCI* in a scientific library, even though I am somewhat disappointed to learn

that a large minority of biochemists never use the *SCI* for any purpose. There are, of course, scholars who never use indexing services of any kind, and, like all evangelists, I'm disappointed to learn that so many are not yet converted. I should not be surprised because hundreds of *CC* readers tell me that *CC* is about the only information service they use. They depend on their personal indexing systems for retrospective retrieval.[5] But, I believe even they will see the light when compact discs are universally available at every scholar's workstation.[6]

In closing, let me say a little about the authors. Hargens received his PhD in sociology from the University of Wisconsin, Madison, in 1971, and has held faculty positions at the University of Washington, Seattle; Indiana University, Bloomington; and the University of Illinois, where he is now a professor of sociology. His primary research area is the sociology of science, and, in recent years, he has studied the editorial review process of scholarly journals. His recent publications on this topic include "Scholarly consensus and journal rejection rates"[7] and "Variation in journal peer review systems."[8] The former article is an extension of the classic paper by Harriet A. Zuckerman and Robert K. Merton, Columbia University, New York, "Patterns of evaluation in science: institutionalization, structure and functions of the referee system."[9,10] Lowell is a founding member and has been active in the affairs of the Society for the Social Studies of Science.

Schuman is a professor of sociology and a program director of the Survey Research Center of the Institute for Social Research, University of Michigan. He received his PhD in sociology, in 1961, from Harvard University, Cambridge, Massachusetts. Prior to coming to the University of Michigan in 1964, he was a research associate for three years with the Center for International Affairs at Harvard, spending approximately half of that time as the field director of a project on development in Dacca, Bangladesh.

His main research and teaching interests are in survey research methods, social psychology (especially the study of attitudes, beliefs, and behavior), and race and ethnic relations. He is the author or coauthor of seven books and monographs, and he has written 60 journal articles and chapters. Schuman received a Guggenheim Fellowship (1980-1981) and was a Fellow at the Center for Advanced Study in the Behavioral Sciences, Palo Alto, California, from 1985 to 1986. He is presently editor of the journal *Public Opinion Quarterly* and served as editor of *Social Psychology Quarterly* from 1977 to 1979.

Hargens and Schuman began their collaboration on the use and evaluation of citation index data by academics when they realized that the data could be used to study both the sociology of science and general social-psychological questions.

REFERENCES

1. **Garfield E.** How to use citation analysis for faculty evaluations, and when is it relevant? Parts 1 & 2. *Current Contents* (44):5-13, 31 October 1983; (45):5-14, 7 November 1983. (Reprinted in: *Essays of an information scientist*. Philadelphia: ISI Press, 1984. Vol. 6. p. 354-72.)
2. ----------. Uses and misuses of citation frequency. *Current Contents* (43):3-9, 28 October 1985. (Reprinted in: *Ibid.*, 1986. Vol. 8. p. 403-9.)
3. ----------. Evaluating research: do bibliometric indicators provide the best measures? *Current Contents* (14):3-10, 3 April 1989.
4. **Hargens L L & Schuman H.** Citation counts and social comparisons: scientists' use and evaluation of citation index data. *Soc. Sci. Res.* 19(3):205-21, September 1990.
5. **Gabaldon D J.** Bibliography managers save aggravation. *The Scientist* 3(19):22, 2 October 1989.
6. **Garfield E.** Announcing the *SCI Compact Disc Edition*: CD-ROM gigabyte storage technology, novel software, and bibliographic coupling make desktop research and discovery a reality. *Current Contents* (22):3-13, 30 May 1988. (Reprinted in: *Op. cit.*, 1990. Vol. 11. p. 160-70.)
7. **Hargens L L.** Scholarly consensus and journal rejection rates. *Amer. Sociol. Rev.* 53:139-51, 1988.
8. ----------. Variation in journal peer review systems. *JAMA—J. Am. Med. Assn.* 263(10):1348-52, 9 March 1990.

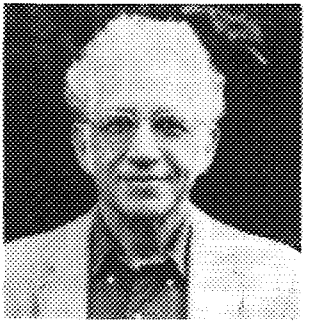
9. **Zuckerman H A & Merton R K.** Patterns of evaluation in science: institutionalization, structure and functions of the referee system. *Minerva* 9:66-100, 1971.
10. . . . . Citation Classic. Commentary on *Minerva* 9:66-100, 1971. *Current Contents/ Social & Behavioral Sciences* 22(32):20, 11 August 1986.

Reprinted from: *Social Science Research* 19:205-21, 1990.

Back to Introduction

Citation Counts and Social Comparisons: Scientists' Use and Evaluation of Citation Index Data

← Lowell L. Hargens
Department of Sociology
University of Illinois
Urbana, IL 61801
and
Howard Schuman →
Institute for Social Research
University of Michigan
Ann Arbor, MI 48109



Data from samples of biochemists and sociologists show that nearly all are familiar with citation indexes and that the two groups are equally likely to have used a citation index for bibliographic purposes. We develop three hypotheses from social comparison theory to account for variation in use and evaluation of citation counts as indicators of scientific achievement: (1) more highly cited scientists will more often use and more highly evaluate citation counts as indicators of scientific achievement than will less cited scientists, (2) these relationships will be stronger for sociologists than for biochemists, and (3) sociologists as a whole will more often use and more highly evaluate citation counts than biochemists. Finally, among sociologists, we hypothesize that those primarily interested in quantitative research areas will use and favor citation counts more than those with primarily qualitative or theoretical interests. Our data support all but one of these hypotheses. We also report unexpected differences in use and evaluation of citation counts by sex and departmental prestige. ©1990 Academic Press, Inc.

Academic scientists are ambivalent about attempts to measure scholarly contributions. They often view such attempts negatively because they fear that using quantifiable characteristics to gauge contributions leads to the distortion of research products. For example, widespread use of publication counts as a basis for promotion decisions is sometimes blamed for a deluge of trivial publications. Scientists see their research as craft work (Whitley, 1984:6-7), and many believe that using one or two easily quantifiable aspects to assess a scientist's scholarly product tends to debase that product.[1]

Yet assessing scholarly contributions quantitatively has undeniable attractions as well. Decisions about tenure, promotion, and other academic awards are necessary, and quantitative information about performance oridinarily plays a role in them (Braxton and Bayer, 1986). Furthermore, reliance on quantitative measures may protect evaluators from charges that their decisions are particularistic, or are based on candidates' ascriptive characteristics (Lewis, 1975:40-42). Finally, academic research work is a nonroutine, often ambiguous activity with infrequent formal assessments of one's performance. Individuals in such circumstances are likely to seek evidence about their relative performance (Festinger, 1954), and quantitative forms of evidence

[1]Stigler (1984) makes this point forcefully in his satire "An Academic Episode" in which an academic administrator radically changes faculty members' behavior by setting up and altering a system for measuring scholarly merit.

may be especially attractive because they appear to be "objective."

Since its initial publication in 1964 by the Institute for Scientific Information®, the *Science Citation Index*® (*SCI*®) has made it relatively easy to count how often an individual has been cited by other scholars during a given year. Although the *SCI* was developed as a bibliographic tool to help scientists trace the literature in their areas of interest (Garfield, 1979:49-61), measuring the impact of individuals' work has become the *SCI*'s most visible and controversial use (Wade, 1975; Garfield, 1979:240-252). The controversy about such use bears witness to scientists' ambivalence toward citation counts as measures of scholarly performance.

In this paper, we report results from a survey of academic scientists' use and evaluation of citation count information. In part, we sought to determine if patterns of use and evaluation are consistent with Festinger's (1954) social comparison theory, especially as integrated with attribution theory (Goethals and Darley, 1977). Festinger hypothesized that people desire to evaluate their own abilities, and that when they lack objective measures, they resort to comparing themselves to others. Goethals and Darley added that people want to find that their abilities—necessarily measured in terms of performance—compare favorably with others'. We developed three hypotheses about scientists' use and evaluation of citation counts from social comparison theory. The first and most general hypothesis is

> 1. Scientists who are highly cited will be more likely to use citation counts for gauging scholarly contributions than will infrequently cited scientists. The former will also evaluate citation counts for this purpose more highly than the latter.

We assume that most scientists feel that their own research contributions are important, but that they also seek support for these self-evaluations. Discovering that one's work is highly cited confirms positive self-evaluations and at the same time validates citation counts as a measure of scholarly contribution. This is a kind of construct validation in which *both* constructs—the merit of one's work and the value of citation counts—support each other. In contrast, infrequently cited researchers should be less likely to regard citation counts a valid measure of scholarly contributions because citation counts do not support their tendency to evaluate their own work positively.

We tested this hypothesis by drawing samples of scientists at U.S. universities in two quite different fields: biochemistry and sociology. We selected these fields in part because of the availability of sampling frames that gave university affiliations and other relevant information. We expected the relationship stated in hypothesis 1 to apply in each field, but on the basis of evidence that the natural sciences exhibit a higher level of consensus than the social sciences,[2] we also expected certain differences between them. Specifically, studies of the social organization of research work (Lodahl and Gordon, 1972; Hargens, 1975), the evaluation of scholars (Yoels, 1974; Hargens and Hagstrom, 1982), competition for priority in reporting research findings as indexed by both the incidence of being anticipated before publication and publication in the form of articles rather than books (Hagstrom, 1965), and the evaluation of research proposals (Cole and Cole, 1981) and papers submitted to journals (Zuckerman and Merton, 1972; Pfeffer, Leong, and Strell, 1977) all show higher levels of consensus in the natural than the social sciences. Evidence also suggests that differences in overall levels of consensus affect scholars' attributional patterns; Rubin (1975) found that chemists who had been denied tenure at Ph.D.-granting departments were more likely to blame themselves for their failure than sociologists, who more

[2]A number of concepts roughly correspond to our "level of consensus," including "paradigm status" (Lodahl and Gordon, 1972), "degree of codification" (Zuckerman and Merton, 1972), and the "hard-soft" dimension (Biglan, 1973; Smart and Elton, 1982).

TABLE I
Characteristics of the Sample Strata and Response Rates, by Discipline and Department Prestige Level

Discipline-prestige combination	Range of prestige scores	No. of Associate and Full Professors	No. Sampled	No. Responded	Response Rate
Biochemistry High prestige	74-65	234	66	46	70%
Biochemistry Low prestige	45-33	200	67	52	78%
Sociology High prestige	71-63	195	64	49	77%
Sociology Low prestige	43-28	209	69	57	83%

Source: Jones *et al.* (1982).

often disputed the validity of the criteria by which they were judged. These considerations led to two more hypotheses:

> 2a. Scientists in fields with relatively low levels of consensus on appropriate research questions and techniques are more likely to use citation counts to measure individuals' scholarly contributions than scientists in fields with relatively high levels of consensus. The former will also evaluate such use of citation counts more favorably than the latter.

This relationship derives from Hypothesis II of Festinger's (1954) statement of social comparison theory: when more objective means of evaluation are unavailable, people evaluate themselves by comparison with others. In this case, lack of consensus about the importance of contributions in a field should lead its members to be less certain about the value of their own and others' research contributions than in fields with high levels of consensus, and this should lead them to seek means of gauging contributions more than members of high-consensus fields.[3] Furthermore,

> 2b. The relationship between one's own citation level and one's use of citation counts to measure scholarly contributions will be stronger in fields with less consensus than in fields with more consensus. Similarly, the relation between one's own citation level and one's evaluation of citation counts as a measure of scholarly contributions will be greater in low- than in high-consensus fields.

The predictions in hypothesis 2b follow from those in hypotheses 1 and 2a. Highly cited biochemists should feel less need to use citation counts for evaluation since at best they would be redundant with widely shared evaluations among others in the field. As a result, the validation citation counts afford to those whose work is highly cited should be less in biochemistry than in sociology. Moreover, infrequently cited sociologists should be more negative toward citation counts than infrequently cited biochemists because the former are more likely to be able to argue that the citation-count "evidence" is inconsistent with other evaluations of their work. Indeed, in sociology, having one's work infrequently cited is sometimes viewed as a sign that one rejects current research fads and instead concentrates on more important, although unfashionable, projects.

We also developed an hypothesis that is unrelated to social comparison theory but which stems from scientists', perhaps especially social scientists, skepticism about trying to measure scholarly contributions. Sociologists often disagree about whether quantitative data can contribute significantly to understanding social behavior. Therefore, we reasoned that those who doubt the value

[3]Hargens and Hagstrom (1982) studied the link between consensus and the ability to gauge research potential and past contributions, and found results consistent with their predictions about how status-attainment patterns should vary across fields with differing levels of consensus.

of quantitative data generally should have a low opinion of citation counts quite apart from other factors. Thus, even if sociologists are more positive toward citation counts than biochemists as a result of social comparison processes, the fact that a subset of sociologists denigrate any form of quantitative evidence could obscure the field differences.

Each of the above hypotheses specifies a relationship that should hold independently of other possible causes of scientists' use and evaluation of citation counts. To evaluate the accuracy of the predictions, an analysis must, insofar as possible, include other causes that may be correlated with the independent variables at issue. Thus, we gathered data on other variables that might affect the use and evaluation of citation counts beyond the effects discussed above.

Sampling and Data Collection

We sampled from the lists of biochemistry graduate faculty in the American Chemical Society's *Directory of Graduate Research* (1984) and sociology graduate faculty in the American Sociological Association's *Guide to Graduate Departments in Sociology* (1985). We decided to draw the samples from high- and low-prestige departments, as measured by departments' reputational rankings reported by Jones, Lindsey, and Coggeshall (1982), because the reputational rankings of departments are substantially associated with measures of the eminence of their members (Cole and Cole, 1973; Long, 1978).

We sought responses from at least 50 associate and full professors in each discipline-prestige combination, and expected a response rate of about 75% given the brevity of our questionnaire, which we designed to fit on a postcard (our questionnaire is reproduced in the Appendix). In addition, we wanted to include no more than one-third of the members of any one department in our sample. Accordingly, we began by determining the number of high-prestige biochemistry departments required to produce a sampling frame of at least 200 persons, the number of low-prestige biochemistry departments which met the same condition, etc.[4] Next, we determined the sampling fraction for each group that would yield a sample of approximately 66 members. We then randomly selected the four samples and mailed explanatory letters plus questionnaires in late April 1985. Three weeks later we mailed follow up questionnaires to nonrespondents. Table 1 gives, for each of the four groups, the range of prestige scores of the departments, the number of associate and full professors, the numbers we selected for our samples, the numbers who returned questionnaires, and the return rates.[5]

In addition to questionnaire data, we collected biographical data on the members of our samples. We obtained information on their sex, academic rank, and year of Ph.D. (or M.D. for a few biochemists) from the directories we sampled from. For a few sample members for whom the directories did not include these data, we used the most recent edition of *American Men and Women of Science*. We also collected bibliometric data, including each sample member's number of citations in the 1984 *SCI* or *Social Sciences Citation Index®*, and the median number of citations for all of the associate and full professors in each sample member's department. We collected data on the latter variable to assess the possibility that researchers' perceptions of their relative eminence are based on their relative standing among the members of their own de-

[4]We needed at least 200 members in each of the four groups because (1/3) (3/4) 200 = 50. We excluded from our sampling frame persons with ranks below associate professor because their typically low citation levels only reflect their professional youth. We also omitted professors emeriti.

[5]Table 1 shows that members of highly ranked departments were less likely to return questionnaires than members of low ranked departments. In addition, within each field citations to sample members' work was negatively correlated with whether they responded: for biochemistry $r = -.17$ and for sociology $r = -.18$. Thus, eminent scholars are slightly underrepresented in our samples.

partments as well as on their relative standing among all the members of their disciplines. After gathering these and other data, we worked only with identification numbers to protect our respondents' confidentiality.

Finally, for our sociology sample we constructed a measure of whether a respondent is likely to view the quantitative analysis of empirical data favorably by using information about the specialties they listed in the 1985 *Guide to Graduate Departments in Sociology*.[6] Our measure classified 34% of the sociologists in our sample as quantitatively oriented, 46% as mixed, and 20% as nonquantitatively oriented.[7]

The concluding sections of this article, including the **Results** and **Conclusions** of the study, will appear in the January 21, 1991, issue of *CC*.

[6]We began by listing specialities whose members are, in our experience, typically either favorably or unfavorably disposed toward using quantitative data. Our list of quantitatively oriented specialities included "quantitative methods," "statistics," "research methods," "evaluation research," "demography," and "population." Our list of nonquantitatively oriented specialities included "theory," "interpretive sociology," "comparative and historical sociology," "macro sociology," "religion," "culture," "cultural change," "field methods," "psychoanalytic sociology," "Marxist sociology," and "mathematical theory and modeling" (members of this last specialty often emphasize the importance of formal models for analyzing social phenomena and express skepticism about the value of statistical analyses of empirical data). We classified specialities not included in either of these two lists as "mixed." Next, we examined each sample member's list of specialities. We classified sample members as quantitatively oriented if they listed only quantitative or both quantitative and mixed specialties. We classified sample members as nonquantitatively oriented if they listed only nonquantitative or nonquantitative and mixed specialties. We classified as mixed sample members with all other combinations. Note that since sociologists typically listed three or four specialties in their entries in the *Guide to Graduate Departments in Sociology*, the validity of our classification of individuals is probably greater than that of our classification of specialties.

[7]We each classified the sociologists in our sample independently and obtained discrepant classifications for only 13 of the 133 sociologists (and resolved the discrepancies on a case-by-case basis). The association between our independent classifications, when we treat the three categories as an ordinal measure of orientations toward quantitative data, yielded a coefficient of .997.

REFERENCES

Biglan, A. E. (1973). "The characteristics of subject matter in different academic areas," Journal of Applied Psychology **57**, 195-203.

Braxton, J. M., and Bayer, A. E. (1986). "Assessing faculty scholarly performance," *in* Measuring Faculty Research Performance (J. W. Creswell, Ed.), pp. 25-42. Jossey-Bass, San Francisco.

Cole, J. R., and Cole, S. (1973). Social Stratification in Science. Univ. of Chicago Press, Chicago, IL.

Cole, J. R., and Cole, S. (1981). Peer Review in the National Science Foundation: Phase Two of a Study. National Academy Press, Washington, DC.

Festinger, L. (1954). "A theory of social comparison processes," Human Relations **7**, 117-140.

Garfield, E. (1979). Citation Indexing—Its Theory and Application in Science, Technology and Humanities. Wiley, New York.

Goethals, G. R., and Darley, J. M. (1977). "Social comparison theory: an attributional approach," *in* Social Comparison Processes (J. M. Suls and R. L. Miller, Eds.), pp. 259-278. Wiley, New York.

Hagstrom, W. O. (1965). The Scientific Community. Basic Books, New York.

Hargens, L. L. (1975). Patterns of Scientific Research: A Comparative Analysis of Research in Three Scientific Fields. American Sociological Assoc., Washington, DC.

Hargens, L. L., and Hagstrom, W. O. (1982). "Scientific consensus and status attainment patterns," Sociology of Education **55**, 183-196.

Jones, L. V., Lindsey, G., and Coggeshall, P. E. (1982). An Assessment of Research-Doctorate Programs in the United States. National Academy Press, Washington, DC.

Lewis, L. S. (1975). Scaling the Ivory Tower: Merit and Its Limits in Academic Careers. Johns Hopkins Univ. Press, Baltimore, MD.

(continued on next page)

Lodahl, J. B., and Gordon, G. (1972). "The structure of scientific fields and the functioning of university graduate departments," American Sociological Review **37**, 57-72.
Long, J. S. (1878). "Productivity and academic position in the scientific career," American Sociological Review **43**, 889-908.
Pfeffer, J., Leong, A., and Strehl, K. (1977). "Paradigm development and particularism: Journal publication in three scientific disciplines," Social Forces **55**, 938-951.
Rubin, L. C. (1975). "The Dynamics of Tenure in Two Academic Disciplines," Unpublished Ph.D. dissertation. SUNY at Stony Brook.
Smart, J. C., and Elton, C. F. (1982). "Validation of the Biglan model," Research in Higher Education **17**, 213-229.
Stigler, G. J. (1984). The Intellectual and the Marketplace, 2nd ed. Harvard Univ. Press, Cambridge, MA.
Wade, N. (1975). "Citation analysis: A new tool for science administrators," Science **188**, 429-432.
Whitley, R. (1984). The Intellecutal and Social Organization of the Sciences. Oxford Univ. Press, Oxford, UK.
Yoels, W. C. (1974). "The structure of scientific fields and the allocation of editorships on scientific journals: Some observations on the politics of knowledge," Sociological Quarterly **15**, 264-276.
Zuckerman, H., and Merton, R. K. (1972). "Age, aging and age structure in science," *in* Aging and Society (M. W. Riley *et al.*, Eds.), Russell Sage, New York.

APPENDIX

The Survey Questionnaire

1. Are you at all familiar with the *Science Citation Index* ® (*Social Sciences Citation Index* ®), which lists individuals alphabetically and shows the citations to each of their publications during a given year?

_____1. Yes _____2. No, never heard of it (please return post card)

2. Have you ever consulted the *Science Citation Index* (*Social Sciences Citation Index*)?

_____1. Yes _____2. No (go to Q. 3)

For what purpose? (Check all that apply)

_____ 1. To use citations to an earlier work to locate more recent work on that topic.
_____ 2. To determine how frequently particular individuals have been cited during a certain period.
_____ 3. Other (please specify)

3. Has your department ever made use of citation counts in making decisions about hiring, promotion or salaries?

_____ 1. Yes _____ 2. No _____ 3. Don't know

4. Overall, how useful do you think a citation count is in evaluating the contributions of someone in your field? (check one point on the line)

Not useful at all	1	2	3	4	5	6	7	8	9	10	Extremely useful