This is a preprint of an article accepted for publication in
European Science Editing

Functional use of frequently and infrequently cited articles in citing publications.

A content analysis of citations to articles with low and high citation counts

Lutz Bornmann,[a] Hans-Dieter Daniel[a,b]

[a] Professorship for Social Psychology and Research on Higher Education, Swiss Federal Institute of Technology Zurich, Zurich (Switzerland)

[b] Evaluation Office, University of Zurich, Zurich (Switzerland)

Address for correspondence:

Dr. Lutz Bornmann

ETH Zurich

Professorship for Social Psychology and Research on Higher Education

Zaehringerstr. 24

CH-8092 Zurich

E-mail: bornmann@gess.ethz.ch

Abstract

Using publication and citation data from a study on the selection procedure of the Boehringer Ingelheim Fonds (B.I.F.), this study investigated the extent to which frequently and infrequently cited articles were used differently by the scientists that cited them. The data set consisted of 31 articles by B.I.F. grant applicants that had received 451 citations in 270 citing publications. In a comprehensive content analysis each reference to the B.I.F. article in the citing publication was classified according to two categories: 1) the location of the citation within the citing publication (section of the paper in which the citation appears) and 2) meaningful or cursory mentioning of the article in the citing publication. The results showed statistically significant differences between the B.I.F. applicants' articles with low or high citation counts. All in all, the results indicate that an article with high citation counts had greater relevance for the citing author than an article with low citation counts.

Keywords

citation, citation behavior, citation content analysis, citation context analysis, evaluative bibliometrics

Introduction

The central problem in the use of citation counts to evaluate scientific work is that it is not certain what is being measured by the citations (COZZENS, 1989). Are frequently cited articles used by a citing author differently than articles that are infrequently cited? For citing authors, does a frequently cited article have greater relevance – in terms of 'intellectual influence' and 'contribution to scholarly progress' (MOED, 2005, p. 221) – than an infrequently cited article? According to the social constructivist sociology of science (LATOUR & WOOLGAR, 1979) the significance of an article depends largely on the manner in which it is used by other scientists. If scientists intensively use the content of an article, knowledge claims that are made in this article become scientific facts and are gradually integrated into the stock of scientific knowledge (AMSTERDAMSKA & LEYDESDORFF, 1989). Although a number of studies on citing behavior have already been published (BORNMANN & DANIEL, accepted for publication, were able to find about 30 studies for a literature review on citing behavior), very few of these studies investigated the extent to which articles with different citations counts are used differently by the citing authors.

The present study investigated to what extent frequently and infrequently cited articles were used differently by the scientists that cited them. In a comprehensive content analysis we classified citations to cited articles in citing publications using two different categorizations of citations. Firstly, we noted the location of citations with respect to one of the sections of the citing publication: introduction, methods, results, and discussion. According to VOOS and DAGAEV (1976) there are obvious indications that it is possible to calculate the value of a cited article for the author of the citing publication using its location in the citing publication. In a citation content analysis, MARICIC et al. (1998) attach the highest importance to citations in the methods or results section of a cited publication. Citations in the discussion section are rated somewhat lower, and citations in the introduction section are ascribed the lowest importance. For CANO (1989), citations located in introductory sections represent a

"setting of the stage" (p. 288) and have little informational utility to the authors of the citing publications.

Secondly, we classified citations according to intensity of mentioning of the cited article by the citing authors. We followed BONZI (1982) and chose a simple three-level distinction that captured both cursory mentioning and more meaningful mentioning of the cited article (HOOTEN, 1991; MARICIC et al., 1998). Other schemes that was used for citation content analyses also include cursory or meaningful mentioning of cited articles in the citation categories (on this, see MARICIC et al., 1998): cursory citation is called perfunctory (MURUGESAN & MORAVCSIK, 1978), peripheral (MCCAIN & TURNER, 1989), or non-essential (CANO, 1989), and meaningful citation is called organic (MURUGESAN & MORAVCSIK, 1978), central (MCCAIN & TURNER, 1989), or essential (CANO, 1989).

Methods

The sample of articles cited in the publications analyzed

We previously investigated committee peer review for awarding long-term fellowships to young researchers as practiced by the Boehringer Ingelheim Fonds (B.I.F.) – a foundation for the promotion of basic research in biomedicine (BORNMANN & DANIEL, 2005a, 2005b, 2005c, 2006a, 2006b, 2007). Assessing the validity of the B.I.F. selection decisions, bibliometric analyses for articles published previous to the post-doctoral applicants' approval or rejection for a B.I.F. fellowship were conducted. All in all, 1,586 articles had been published by 397 applicants previous to their applications to the B.I.F. (on average four articles).

Using the same data set of articles used to evaluate the B.I.F. selection procedure (B.I.F. applicants' articles and their citing publications), the present study examined to what extent frequently and infrequently cited papers were differently used by scientists who cited them. As content analysis of citations with different classifications is very time-consuming (it entails finding the citation in the article, reading the whole sentence, incorporating the

reference, and classifying the citation several times) we did not include all B.I.F. applicant articles (and their citations) in the analysis but instead draw a stratified random sample from the total data set of articles, selecting a separate random sample from each of two strata. The stratification variable was the decision by the B.I.F. Board of Trustees to approve or reject an applicant for a post-doctoral fellowship, as it can be assumed that the articles published by approved applicants were of higher quality than the articles published by rejected applicants. In total, 34 articles written by B.I.F. applicants with comparable citation windows of at least eight years were selected randomly: 17 articles each by approved and rejected applicants.

The citations to the articles published by the B.I.F. applicants

The 34 articles of the B.I.F. applicants in our sample were cited by 308 citing publications, with an average of 11 citing publications per cited article (median). The sample of the citing publications was adjusted by excluding those that listed the B.I.F. applicants' articles only in a bibliography without mention in the text (n=2) and those that were published in non-English language journals (n=5). In order to test the extent to which the number of citations to articles by the B.I.F fellowship applicants correspond with the categories of both citation classifications, we divided the total of 34 articles into two groups by using the citations' median value as threshold (see PREACHER et al., 2005): 1) articles with low citation counts (n=24), that is, articles with fewer than 11 citations (3 to 10 citations), and 2) articles with high citation counts (n=7), that is, articles cited 12 to 23 times. Three articles with citation counts equal to the median value were not included into the statistical analyses. In so doing, the sample for the statistical analyses consisted of 270 citing publications. As some articles of the B.I.F. applicants were cited multiple times in one citing publication, the total number of citations was 451. On average, one article by the B.I.F. applicants was cited 1.7 times in one of the 270 citing publications.

Statistical methods

The associations between the categorical variables low or high citation counts for the articles by the B.I.F. applicants and the categories of the categorizations were calculated using the Cochran Mantel-Haenszel test (AGRESTI, 2002, section 7.5.3-7.5.6; CYTEL SOFTWARE CORPORATION, 2007). The test adjusted for potential effects of qualitative differences between articles of the approved and rejected B.I.F. applicants in the statistical analyses. Since the result of the statistical significance test is dependent on sample size and "statistical significance does not mean real life importance" (CONROY, 2002, p. 290), it is the strength of the association that is more interesting and important for interpreting the empirical finding. For calculating strength we have to employ an additional measure of association, here *Cramer's V* coefficient (CRAMÉR, 1980).

Results

The location of the citations to the articles by the B.I.F. applicants with respect to one of the sections of the citing publication

Table 1 shows the sections in the citing publications where the B.I.F. applicant articles are cited: a total of 32% of the articles are cited in the introduction, 24% in the methods, 13% in the results, and 31% in the discussion section. This result agrees approximately with citation distributions reported by VOOS and DAGAEV (1976) and CANO (1989). Their findings indicate that the largest concentration of citations is located in the beginning sections of the citing publications.

A look at the differences in the percentages of citations in the different sections of the citing publications between articles by B.I.F. applicants with low or high citation counts in Table 1 shows expected differences between the methods, results, and discussion sections. As expected, articles with high citation counts are more frequently cited in the methods (27% of the citing publications) and results (15% of the citing publications) sections than articles with low citation counts (methods: 20% of the citing publications; results: 11% of the citing

publications). Articles with low citation counts (39% of the citing publications) are more frequently cited in the discussion section than articles with high citation counts (25% of the citing publications). But contrary to our expectations, articles with high citation counts are more frequently cited in the introduction section (34% of the citing publications) than articles with low citation counts (30% of the citing publications). The differences in the distribution of the citations in sections of the citing publications between articles by the B.I.F. applicants with low or high citation counts are statistically significant; T (n=350) = 8.82, p=.03; with small effect size, *Cramer's V*=.17. (see Table 1).

Table 1. Sections in the citing publications containing the citations to the B.I.F. applicants' articles with low or high citation counts

Section of citing publication	Articles with low citation counts (3 to 10 citations)	Articles with high citation counts (12 to 23 citations)	Total
Introduction	30	34	32
Methods	20	27	24
Results	11	15	13
Discussion	39	25	31
Total	100	100	100
Number of classified citations	162	188	350

Notes. T (n=350) = 8.82, p=.03 (Cochran Mantel-Haenszel test adjusted for potential effects of qualitative differences between articles of the approved and rejected B.I.F. applicants); *Cramer's V*=.17. A total of only 350 citations could be assigned to a section, because 101 citations to the B.I.F. applicants' articles were located in citing publications that had no (clear) section headings.

Cursory or meaningful mentioning of the B.I.F. applicants' articles in the citing publications

The citation content categories provided by BONZI (1982) are based on the premise that one measure of true relevance to a citing publication is the extent of treatment of the cited article in the citing publication. An article simply mentioned in a citation can be expected to be less relevant for the author than a citation where the cited article is discussed in depth within the citing publication. For our citation content analysis we used three categories

provided by BONZI (1982) to measure citation relevance: (1) not specifically mentioned in text (e.g., "Several studies have dealt with …"), (2) barely mentioned in text (e.g., "Smith has studied the impact of …"), and (3) one quotation or discussion of one point in text (e.g., "Smith found that …").

For this type of content analyses it is customary for two persons to conduct the coding of text material for purposes of determining the interjudgmental reliability of the codings, using measures of agreement (MCCAIN & TURNER, 1989). In the present study two independent coders classified the citations as to cursory or meaningful mentioning of the cited article in the citing publications. The reliability of the two coders' ratings was very high, kappa coefficient = .93 (on interpreting the coefficient, see VON EYE & MUN, 2005, pp. 5-6).

Table 2. Cursory or meaningful mentioning of the B.I.F. applicants' articles with low or high citation counts

Citation content category	Articles with low citation counts (3 to 10 citations)	Articles with high citation counts (12 to 23 citations)	Total
(1) Not specifically mentioned in text (e.g., "Several studies have dealt with …")	31	13	22
(2) Barely mentioned in text (e.g., "Smith has studied the impact of …")	32	48	40
(3) One quotation or discussion of one point in text (e.g., "Smith found that …")	37	39	38
Total	100	100	100
Number of classified citations	226	223	449

Notes. T (n=449) = 22.84, p=.00 (Cochran Mantel-Haenszel test adjusted for potential effects of qualitative differences between articles of the approved and rejected B.I.F. applicants); *Cramer's V*=.22

The distribution of the citations to the article of the B.I.F. applicants across the three citation content categories in Table 2 shows that the greatest percentage of articles (40%) are barely mentioned in the citing publications (second citation content category: e.g., "Smith has

studied the impact of …"). In another 38% of the citations, either a passage from a B.I.F. applicant's article is cited directly or the content of the article is discussed (third citation content category: e.g., "Smith found that …"). Twenty-two percent of the citations to the articles are simple mentions, with no discussion of the content of the cited article (first citation content category: e.g., "several studies have dealt with …").

The results in Table 2 show, as expected, that B.I.F. applicants' articles with low citation counts are clearly used more frequently (31% of the citing publications) by the citing authors for cursory mentioning (first citation content category) than articles with high citation counts (13% of the citing publications). B.I.F. applicants' articles with high citation counts are more frequently barely mentioned in the citing publication (48% of the citing publications) and quoted directly or discussed in the citing publication (39% of the citing publications) than articles with low citation counts (second citation content category: 32% of the citing publications; third citation content category: 37% of the citing publications).

The differences between the frequencies are statistically significant; T (n=449) = 22.84, p=.00; the association between both variables has a medium effect size; *Cramer's V*=.22. These findings suggest that when infrequently cited B.I.F. applicants' articles were used, they tended to have lower relevance than frequently cited articles for the authors of the citing publication (and vice versa).

Discussion

Using publication and citation data from the fellow selection procedure of the B.I.F., the present study investigated to what extent frequently and infrequently cited articles were differently used by the scientists that cited them. We utilized two different categorizations to capture the functional use of the articles by the authors in the citing publications: 1) the location of the citation to the article within the citing publication (section), and 2) meaningful or cursory mentioning of the article in the citing publication.

Our results show that for both classifications of citations in the citing publications there are statistically significant differences between B.I.F. applicants' articles with low or high citation counts. B.I.F. applicants' articles with high citation counts were more frequently cited within the citing publications in the methods and results sections than articles with low citation counts. Articles with high citation counts were more frequently cited in meaningful mentions in the citing publications than articles with low citation counts. We proved whether these associations are still hold when the threshold for the categorizations of the citations is changed. Using three groups (low, medium, and high citation counts) instead of two in the statistical analyses, we got nearly the same results.

All in all, our findings suggest that the more an article is cited the more intensively its content is used by the citing scientists. Therefore, citation counts are not only an indication of the (superficial) relevance of research but are also an indicator for the relevance of this research for scientific work in a research field.

References

AGRESTI, A. (2002), *Categorical data analysis*, Hoboken, NJ, USA, John Wiley & Sons, Inc.

AMSTERDAMSKA, O., LEYDESDORFF, L. (1989), Citations: indicators of significance? *Scientometrics*, 15(5-6): 449-471.

BONZI, S. (1982), Characteristics of a literature as predictors of relatedness between cited and citing works. *Journal of the American Society for Information Science*, 33(4): 208-216.

BORNMANN, L., DANIEL, H.-D. (2005a), Committee peer review at an international research foundation: predictive validity and fairness of selection decisions on post-graduate fellowship applications. *Research Evaluation*, 14(1): 15-20.

BORNMANN, L., DANIEL, H.-D. (2005b), Criteria used by a peer review committee for selection of research fellows - a boolean probit analysis. *International Journal of Selection and Assessment*, 13(4): 296-303.

BORNMANN, L., DANIEL, H.-D. (2005c), Selection of research fellowship recipients by committee peer review. Analysis of reliability, fairness and predictive validity of Board of Trustees' decisions. *Scientometrics*, 63(2): 297-320.

BORNMANN, L., DANIEL, H.-D. (2006a), Potential sources of bias in research fellowship assessments. Effects of university prestige and field of study on approval and rejection of fellowship applications. *Research Evaluation*, 15(3): 209-219.

BORNMANN, L., DANIEL, H.-D. (2006b), Selecting scientific excellence through committee peer review – a citation analysis of publications previously published to approval or rejection of post-doctoral research fellowship applicants. *Scientometrics*, 68(3): 427-440.

BORNMANN, L., DANIEL, H.-D. (2007), Gatekeepers of science – effects of external reviewers' attributes on the assessments of fellowship applications. *Journal of Informetrics*, 1(1): 83-91.

BORNMANN, L., DANIEL, H.-D. (accepted for publication), What do citation counts measure? A review of studies on citing behavior. *Journal of Documentation*.

CANO, V. (1989), Citation behavior: classification, utility, and location. *Journal of the American Society for Information Science*, 40(4): 284-290.

CONROY, R. M. (2002), Choosing an appropriate real-life measure of effect size: the case of a continuous predictor and a binary outcome. *The Stata Journal*, 2(3): 290-295.

COZZENS, S. E. (1989), What do citations count? The rhetoric-first model. *Scientometrics*, 15(5-6): 437-447.

CRAMÉR, H. (1980), *Mathematical methods of statistics*, Princeton, NJ, USA, Princeton University Press.

CYTEL SOFTWARE CORPORATION (2007), *StatXact: version 8*, Cambridge, MA, USA, Cytel Software Corporation.

HOOTEN, P. A. (1991), Frequency and functional use of cited documents in information science. *Journal of the American Society for Information Science*, 42(6): 397-404.

LATOUR, B., WOOLGAR, S. (1979), *Laboratory life: the social construction of scientific facts*, London, UK, Sage.

MARICIC, S., SPAVENTI, J., PAVICIC, L., PIFAT-MRZLJAK, G. (1998), Citation context versus the frequency counts of citation histories. *Journal of the American Society for Information Science*, 49(6): 530-540.

MCCAIN, K. W., TURNER, K. (1989), Citation context analysis and aging patterns of journal articles in molecular genetics. *Scientometrics*, 17(1-2): 127-163.

MOED, H. F. (2005), *Citation analysis in research evaluation*, Dordrecht, The Netherlands, Springer.

MURUGESAN, P., MORAVCSIK, M. J. (1978), Variation of nature of citation measures with journals and scientific specialties. *Journal of the American Society for Information Science*, 29(3): 141-147.

PREACHER, K. J., RUCKER, D. D., MACCALLUM, R. C., NICEWANDER, W. A. (2005), Use of the extreme groups approach: a critical reexamination and new recommendations. *Psychological Methods*, 10(2): 178-192.

VON EYE, A., MUN, E. Y. (2005), *Analyzing rater agreement. Manifest variable methods*, Mahwah, NJ, USA, Lawrence Erlbaum Associates.

VOOS, H., DAGAEV, K. S. (1976), Are all citations equal? Or, did we *op. cit.* your *idem*? *Journal of Academic Librarianship*, 1: 19-21.