HUMAN AFFAIRS, 5, 1995, 1, 9-18

A CONTRIBUTION TO THE EVALUATION OF RESEARCH FINDINGS

Eva KRIŠTOFIČOVÁ
Department of Library and Information Science, Faculty of Arts, Comenius University, Gondova 2, 818 01 Bratislava, Slovakia

The issue of scientific assessment and the evaluation of the results of scientific research by means of citations has become the subject of many discussions. Numerous problems associated with publication practices and citation ethics are well known. This paper is directly connected with the reliability of citations and their objectiveness. The focus of the article is just on this circle of issues. The author highlights the increasingly topical problems of citation in the area of creation and the use of electronic documents. This paper contains some results obtained from surveys carried out among university students and research workers.

A scientist is obliged to inform the scientific public on his/her achievements usually by writing a scientific publication. It is meaningful if other specialists apply their results in their own work. Citation analysis shows the measure of the use of the results. That means that the publication of research findings is accepted as part of the scientific system. As a result, citations serve in a sense as evidence of cognition. This approach had a decisive influence on the importance and the formation of citation analysis (Garfield, 79).

The development of citation analysis is supported by the formation of citation indexing. These tools of retrieval created in the sixties were often more informative and more effective than the previous subject indexes. With regard to the practical significance of citation indexes, citations as the essential object are significant and are interesting from the perspective of research, historiography of science, mapping of science, evaluation of the productivity of research workers, etc.

CITATIONS - AN OBJECT OF CITATION ANALYSIS

Citations as the object of the research of citation analysis were studied by several authors. Cronin (1984) studied citation process in detail on a theoretical basis primarily from the point of view of its tasks and content. In spite of an extensive criticism of Cronin and other authors, it is generally accepted that the document is cited to such an extent and to what extent it is relevant to a particular theme (Smith,

1981). Of course, there are citations which are random, superficial, or not substantial but they do not occur frequently. A study might be cited in a negative sense, but such cases are rare, and, finally, it does not necessarily mean that such a study would be of a lower value. It is in particular cases, where we take account of the length of the period of time during which the article is cited. Such a citation can be used as a kind of indicator of the "survival" of the publication. These assumptions change the outlook on citations and, simultaneously, they might be helpful in avoiding numerous mistakes and noise.

We need not emphasize that citations serve as important links among information sources, wrong citations complicate the research of the scientific community and affect it by their unavailability of original information sources.

All these aspects have to be estimated in the aggregate. Such an approach to the evaluation of citations shows that citations are legitimate objects of research, and, consequently, citation analysis is a significant research tool. Citation analysis has not yet got rid of the problems associated with concepts and methodology. Many of the problems connected with the specific features of citations are described in this paper.

CITATION ANALYSIS TODAY

Citation analysis might be said to be undergoing the process of maturation and search. New possibilities and forms of its application are investigated. There are evident efforts to

(a) define citation analysis, its object and the set aims of its research;

(b) make a choice and define variables, including sources of the data examined, methods used, etc.

In his study of the structural model of all information sources, Zunde differentiates three application areas of citation analysis (Zunde, 1971):

1. qualitative and quantitative evaluation of scientists, publications, and scientific institutions,
2. modelling of the historical development of science and technology,
3. information research and survey.

Within citation analysis there are certain problems that come to the fore:

- citation indexing and its principles
- citation theory
- research of citation motivation
- citation networks and citation matrices
- co-authorship
- obsolescence of literature
- bibliographic coupling and co-citation analysis
- citation analysis of scientific journals
- acquisition policy - citation approach.

Attention of experts has recently turned to the issues of citation in modern electronic documents. These are rather serious problems.

We see that the scope of the use of citation analysis is relatively wide and difficult to be covered in one study. Therefore, we direct our attention towards citation and some related problems.

CITATION PROCESS - CITATION MOTIVATION

A more profound processing of the theory of citation is increasingly urgent within citation analysis. Its main task is to study the factors that are incorporated into citations from the perspective of the links among the known facts, among documents, scientists' behaviour, etc.

The process of designating the use of the results of brainwork of others by means of references and citations, it is the citation process.

The citation process is subjective and thus is hardly able to be standardized. We have to rely on the abilities of the one who cites or on information that ensues from the context of the cited work. We hardly ever will be able to understand why an author cites just in that particular way and at that time. Many factors characterize the author's approach to citation. It is primarily connected with his/her personality and hardly can become his/her public domain. The ideal situation would be if citations would be understood as the author constructed them.

Several authors tried to elucidate problems associated with explaining the motivations and abilities of citation. There are many reasons why the authors cite the works of other scientists. Garfield (1965) names several reasons. These are often used as a basis in argumentation of different opinions on the results of citation analyses. There is a necessity to shape studies of the methods of citations from informal habits (in art usually also at present) into strict forms required by the ethics of contemporary scientists from the region of growth and development. Research into citation motivation is usually much more complicated if it is presented by individual authors since there is a number of motivations and methods for studying them. The individual authors usually employ one of the possible approaches to the research of citation motivation. The problem of self-citations, which is also strongly associated with various citation motivations, is rather serious.

THE SIGNIFICANCE OF CITATION MOTIVATION RESEARCH

Why is it important to deal with these problems and why is their solution so significant? We often face the problem of how to evaluate a publication: what new did it bring, was a method or a theory promoted, improved or discussed in the world of science? It is difficult to answer the question having just the absolute number of citations for one publication, the number of citations of relevant publications or the calculated average number of citations per article at your disposal. For the purpose

of significance, such averages should be compared with those of other relevant publications that were published in similar journals at the same time. It concerns thus the choice of a control set of documents with a comparing function - see Narin (1977), Garfield (1979).

There are some methods that are applicable for this purpose, e.g. the method of bibliographic coupling. Although these approaches are adequate in many cases, one can hardly be satisfied with the results. At the level of reconstruction of research programmes in terms of cognition, citations can be used as a picture of the bonds among documents. Citations link documents but we do not know exactly how. Citation theory will at this level be the result of understanding relations among various documents and functions of various textual elements in these relations. Some researchers working in this field deny the possibility of using citations as an identifier of the development of research programmes. At the level of a factor in research, many theories concerning citations are valid but at the level of the executive dimension (evaluation of the performance of a scientist, an institution, etc.) the theory is rather unintelligible. A series of questions are raised here, such as: Do citations at this level imply "quality", "influence", "intellectual reassuming" or an "institutional affiliation"? Many of the problems indicated are also connected with the issue of insufficient elaboration of the theory of citations. It leads to a stage of disorganization that, instead of asking "what can be expressed by citation", we ask "what does a citation express". Consequently, further research into the theory of citation should lead to a better understanding of the factors included in the citations, mainly from the point of view of bonds between knowledge, among documents, behaviour of scientists, etc. These problems are directly associated with citation behaviour, whose part is the problem of citation motivations, as well as with the whole citation theory. Citation theory among others also investigates citation behaviour.

Scientists' sociological motivations and their practical sides of citation behaviour are often described in literature. Publication practices and citation ethics differ from one scientific field to another. Information obtained in citations can be used in various ways. Their impartiality and reliability are also diverse. A number of problems emerge that are suitable for the research of the motivation of citation behaviour, which would probably answer questions like:

- Are citations really transitive? If so, to what extent?
- Can individual documents be objectively distinguished with respect to transitiveness?
- Is the bond among documents cited in the same way stronger than other citations? Is it always valid?
- How should an author's share in the documents written by several authors be evaluated?

How should the measure of the contribution expressed by the order of the authors' names be decided? Should there not be an adopted rule for determining the order of the names in case of co-authorship?

These are some of the numerous problems pertaining to citations which indicate an urgent need of a more comprehensive theory of citation.

The following authors tried to shed light on the problems associated with explaining motivations and abilities of citing:

Ravetz, who said that citations create a form of the reward or remuneration (Ravetz, 1971),

- Kaplan argued that citations are used for determining the priority of demands (Kaplan, 1965)
- Gilbert maintained that citations are a tool of persuasion (Gilbert, 1985)
- Garfield gave a sum of arguments for citation (Garfield, 1965).
- Brooks used a simple project to research the motives which lead an author to cite the works of other authors (Brooks, 1986). He asked twenty authors working in 15 different scientific disciplines to estimate the intensity of motivation to cite using a four-point scale with 7 motives for each reference (437 altogether). He processed the results by factor analysis, making correlations among the citation motives of 437 references. Brooks' findings met with a positive response in the work of MacRoberts and MacRoberts (1987) who used them as one of their arguments which reduced the validity of the research of the Ortega hypothesis as tested by Cole and Cole (1972).

Benkovič published (1988) a project on citation motivation, where he drafted three research problems:

(a) What are the motives of citation applied by an author in all his publications?

(b) What are the motives of citation in general, what is their structure, which hypothetic (latent) variables are in the background of these motives?

(c) What are the motives of citation of one and the same publication in the works of other authors? To resolve these research problems, he posed research hypotheses which should be tested by statistical methods, such as: factor analysis as a means for finding the motivation structure, cluster analysis of persons and publications according to the intensity of citation motives, testing hypotheses and the equality of mean values in vectors plus regression analysis.

If we take into account the known fact that the scientific community producing scientific results, is not identical with the user community, the direct consequence of this inequality could be that the journals, in which the results from a particular research area have been published, are different from those in which they have been cited. New research methods in some disciplines can be published in one type of journal whereas their applications can appear in other groups of periodicals. The quality of the citations associated with the object of the cited document may differ from the citations expressing their use or application. It is a de facto dilemma - which methodological documents are cited more, theoretical documents or those of

an empirical character? In such a case the presentation of results should be divided between citations from the two source documents. That means, the variable "the number of citations" will often be divided into two parts:

(a) citations referring to the use or application of the method reported in the source document

(b) all other citations.

Such a result can be achieved in three ways:

(a) by classifying the tasks of citations in the citing document

(b) by classifying source documents, cited in the citing document

(c) according to the classification of the cited journals.

The process (a) is evidently the best but also the most costly, whereas the process (c) is easiest and the most inaccurate. These processes have one thing in common. They require classification which should be focused on the content of citations in the documents, and on the documents themselves or journals.

The survey of the problem of citation motivation of scientists can be done by several methods:

(a) by asking the authors for publications directly, i.e. finding their motivation to cite directly from them

(b) by observing the incidence of the individual types of citations, or

(c) by examining relations among citations of individual authors, looking for connections among them, i.e. subjecting citations in the documents to research.

In the area of citation of scientists, their behaviour can be followed from several perspectives:

- the strategy of a scientist for competition or other reasons
- non-understanding of the ethics of citations
- deliberate violation of the ethics of citation (for what reasons).

It is important to realize that citation rate is affected by various factors:

- the properties of the journal where a study or an article is published
- its popularity, availability, language barrier, etc. These properties can be expressed by various quantitative characteristics, such as an impact factor, immediacy index, etc.,
- the type of document, in which the paper was published (we can assume that studies published in journals are better accessible whether from the point of view of time or in view of its exposure, etc., e.g. results published in a monograph or in a thesis),
- last but not least, citation rate can also be influenced by the type of study done. Then a question already posed arises; Are methodological studies and articles cited more than empirical or theoretical?
- the scientific discipline or the scientific problem also plays a role. There are areas which are objects of interest for a number of the scientific public but there are also fields of regional character where the problem is very specific and at-

tracts the attention of just a small group of scientists, or perhaps it will only be interesting in the future.

CITATION PROBLEMS IN ELECTRONIC DOCUMENTS

At this point I would like to mention at least some problems, which will arise in our country in the near future in connection with the advent of computers and their products, i.e. electronic documents. What I have described so far, has primarily concerned citations in traditional printed documents. But nowadays, scientists publish more and more through electronic documents. This is the area where ethical and legal regulations for exploitation of intelectual property are also valid. However the practical application is different, in spite of the existence of the Act on literary and scientific works of art No. 247/1990 coll., which also regards computer programmes as objects of protection as long as they satisfy the conceptual signs of works of art by virtue of this Act. This fact is often forgotten. The users copy computer programmes or computer products without the authors' permission (Kristofičová, 1992 and 1994).

Information sources are of different origins, topicality and reliability. Therefore, the user has to be able to verify the source of the material. The processed electronic documents of high quality might be helpful. It mainly concerns the user's ability to orientate within the electronic document.

The existence of numerous ways of presenting citations confuses not only the users of traditional documents, it is also a serious problem even for the users of online information.

Another aspect that can cause a citation problem in modern documents, is the need of a clearer differentiation of individual information sources, on the computer screen. The text of an electronic document should be adapted to the form of a printed document.

The invention of computer documents will certainly soon be accompanied by further problems in the field of citation. In addition to better legislation and education concerning ethics in publication, the thorough upholding of the existing author's law might help. I am sure that the realization of the existence of problems in this field is one of the first steps towards a solution.

SOME RESULTS OF THE SURVEY IN THE FIELD OF CITATION

Several articles published in scientific journals (Sweetland, 1989) are concerned with the issue of incorrect citations. Their estimates of the amount of citation errors varies between 7 and 50% (depending on criteria used to define the error). The average estimate of wrong citations reported in textbooks is 20-30%. It should be emphasized that those errors are made by experienced researchers, not by students.

The students being primary users often make even more mistakes. There are many reasons for underestimating citation problems, from simple errors in transcription or even forgetting them, up to deliberate falsifying and fraud. The use of several types of the rules of citations also leads to citation errors.

These facts provided a challenge on how to do a survey of students and how to show at what stage of preparation they were at in their scientific work. The results from the survey were described in detail in a recent study (Krištofičová - Lepeňová, 1995). I will present just some of them.

The sample consisted of 120 fresh university graduates (40 respondents graduated from the Faculty of Natural Sciences, 40 from technical university and 40 were graduates from the social sciences). The question, "Do you know citation regulations?", was answered as follows: 46% responded that they knew of them only partially and 13% answered that they did not know how to cite at all. Only 41% of respondents knew how to cite. In their responses to the question, "Where have you been informed about the principles of citation?", only 22% of respondents said that they had been informed at university, twenty (17%) answered that they had not been informed about such principles at all. These are certainly interesting data, especially if we realize that many of them would probably want to publish their scientific results soon.

In our further survey, in a sample of randomly selected research workers, we tried to illuminate the relation of classification of tasks of citation and their popularity and/or the preference of individual functions. The survey was done by the students of the library and information science of the Bratislava Faculty of Arts of Comenius University within the framework of their lectures on bibliometrics. We focused on the behaviour of scientists, whilst also examining their citation policy. The questionnaire method was used. The core of items in the questionnaire was based on Garfield's (1965) citation arguments. They were adapted and amended for our purposes. The interviewed scientific worker could rate his/her reasons for citations with the numbers 1 - 6. The detailed results were published (Krištofičová, 1994) but I would like to alert you to our conclusions achieved. The total number of questionnaires filled in was 96. The majority of respondents (51) reported that they refer to publications containing essential arguments about the particular issue. As many as 31 respondents put this reason in first place, 44 did not report that reason at all, the rest rated its significance using 2-5 scale. After scoring, it reached priority with 245 points. A high incidence (46) was recorded for the citation reason - identification of original research communications. This reason was also rated highly, it had a total of 188 points. (The maximum points to be achieved was 576.)

Only four interviewed respondents gave the reason for citation motivation as paying debts to the founders of the field, eleven alerted to their citations prepared in the papers, three respondents placed this reason even in the first place (older medical doctors). As many as 38 respondents reported the extent of reasons offered in the questionnaire as insufficient. The item 'other' contained most frequently:

- support of one's own argument (theory) by the authority in the discipline or just reported 'other' without giving any reason.

They often gave high preference to their own proposed reasons (12 times 1st place, 11 times 2nd place). Twelve respondents gave among their reasons also the reason of correcting their own publications. According to these results, the majority of authors are content with their published papers. Almost the same number gave as the reason for the correction of the publications of other authors, criticism of other papers occurred 25 times, which might indicate that the respondents use more regularly a critical approach instead of submitting a concrete proposal for the correction of the publication.

I will also offer some other remarks that ensued from the survey.

- The respondents often conditioned their responses by not giving their names, although it had not been required in the questionnaire.
- Some of them said that the prioritization of individual reasons is misleading since such a priority varies with the type of study - we agreed with this approach.
- The individual reasons for citations used in their papers seem to be equally important to several scientists - it is not necessary to examine their priority.

A total of 96 respondents represented twelve disciplines, from medicine, through to economy, law, philosophy, theory of culture, and mathematics. Objectively speaking, the sample was incongruous from this aspect and the results would be insignificant from the perspective of scientific disciplines.

As for the age of respondents, they were mostly older scientists - 62 respondents were over 45 years. It was quite helpful, as we could expect that they were scientists already experienced in publication activities. The survey did not show any significant difference in the preference of the reasons for citation motivation in the groups below 45 and above 46 years of age. Interestingly, of 21 negative citations, eight belonged to the younger group.

Our conclusion is that there are many problems in the field of citation motivation within citation analysis that should be elucidated. We see some progress in the fact that a number of problems have been put forward, but a resolution is needed. This would be an asset to several scientific disciplines, e.g. informetrics, scientometrics, etc. In spite of all this the fact remains that such a research is deemed to be purely academic.

REFERENCES

Benkovič, P.: *Výskum motivácie citovania; otvorené problémy.* In: *Moderní využití citačních indexů.* Symposium Proceedings, VTEI MR ČSVTS Praha, 10 April 1988, pp. 85-89.

Brooks, T.A.: *Evidence of Complex Citer Motivations.* Journal of the American Society for Information Science, Vol. 37, No. 1, 1986, pp. 34-36.

Cole, J.R. - Cole, S.: *The Ortega Hypothesis.* Science, 178, 1972, pp. 368-371.

CRONIN, B.: *The Citation Process.* London, Taylor Graham 1984.

GILBERT, G.N.: *Referencing as Persuasion.* Social Studies of Science, 7, 1985, p. 113.

GARFIELD, E.: *Can Citation Indexing be Automated?* In: *Statistical Association Methods for Mechanized Documentation.* Symposium Proceedings, National Bureau of Standards, 269, 1965, pp. 189-192.

GARFIELD, E.: *Citation Indexing - its Theory and Application in Science, Technology and Humanities.* New York, Wiley 1979.

KAPLAN, N.: *The Norms of Citation Behavior: Prolegomena to the Footnote.* American Documentation, 16, 1965, p. 179.

KRIŠTOFIČOVÁ, E.: *Citácie - meradlo úrovne vedeckého výskumu?* KZ. MS Martin, 1991, pp. 72 to 78.

KRIŠTOFIČOVÁ, E.: *Bibliometria. (Hypertextová učebnica bibliometrie).* KKVI FiF UK Bratislava 1992. 1 floppy disk. 1 MB.

KRIŠTOFIČOVÁ, E.: *Etické a právne problémy využívania elektronických dokumentov.* In: *Aktuálne problémy spracovania a využívania dokumentov a informácií v praxi.* ZSVTS Bratislava 1992, pp. 23 -29.

KRIŠTOFIČOVÁ, E. - PIAČEK, J.: *Hypertextový osobný informačný systém vedeckého a riadiaceho pracovníka.* Paper presented at the symposium *Aktuálne problémy spracovania a využívania dokumentov a informácií v praxi.* ZSVS, 12 November 1992, Bratislava.

KRIŠTOFIČOVÁ, E.: *Teoreticky a prakticky o motivácii citovania vo vede.* Kal, 1994, No. 8-9, pp. 333-339.

KRIŠTOFIČOVÁ, E. - LEPEŇOVÁ, M.: *Citačný proces v praxi.* Kal, 1995 (in press).

MACROBERTS, M.H. - MACROBERTS, B.R.: *Testing the Ortega Hypothesis: Facts and Artifacts.* Scientometrics, 12, 1987, No. 5-6, pp. 293-295.

NARIN, F.: *Evaluative Bibliometrics: The Use of Publication and Citation Analysis in the Evaluation of Scientific Activity.* Cherry Hill, NJ, Computer Horizons 1976.

PERITZ, B.C.: *On the Objectives of Citation Analysis: Problems of Theory and Method.* Journal of the American Society for Information Science. July 1992, pp. 448-451.

PIAČEK, J.: *Aplikácie hypertextového systému INTERES (CREDIT) vo filozofii.* In: *Problémy výstavby informačných systémov v spoločenských vedách a v oblasti ich aplikácií.* Bratislava, DT ZSVTS 1992, pp. 101-106.

RAVETZ, J.R.: *Scientific Knowledge and its Social Problems.* Oxford, Clarendon Press 1971, pp. 245-259.

SMITII, I.: *Citation Analysis.* Library Trends. 1981, 30, pp. 83-106.

SWEETLAND, J.H.: *Errors in Bibliographic Citations: A Continuing Problem.* Library Quarterly, Vol. 59, Oct. 1989, pp. 291-304.

Zákon o literárnych, vedeckých a umeleckých dielach (autorský zákon). Zb. z. č. 247/1990.

ZUNDE, P.: *Structural Models of Complex Information Sources.* Information Storage and Retrieval, No. 7, 1971, p. 1.