

Short Communications

DOES IT PAY TO COOPERATE? A BIBLIOMETRIC CASE STUDY IN MOLECULAR BIOLOGY

H. HERBERTZ

Institute for Genetics, University of Cologne (Germany)

(Received August 30, 1994)

Various bibliometric studies report that multiinstitutional or multinational authored papers are more frequently cited than papers that come from a single institute. The conclusion, however, that there is a systematic improvement of scientific success by cooperation on every level of scientific research in leading or mediocre research institutes might be misleading: In a citation analysis of 13 well-known research institutes in molecular biology there was no difference in the average citations per paper with regard to cooperations. In a subsample of 7 German institutes that difference found could be explained by selfcitations. In another case, all articles of a two year sample of an excellent journal in molecular biology, the EMBO-Journal, the same phenomenon could be observed: Differences in the average citations per article with regard to cooperations could be explained by selfcitations.

Introduction

Bibliometric analysis about impact or visibility of scientific publications as a result of scientific collaboration seem to support the idea that it pays to cooperate.¹⁻³ The reported differences in the mean citation rates of multi-institutional authored papers and papers with no institutional collaboration seem to support the science policy to strengthen their efforts to increase – especially international – scientific cooperation. These bibliometric studies however have a methodological shortcoming which make us doubt the interpretation of the results. The topic in question is the selfcitation-rate. Bibliometric studies about scientific cooperations usually count all citations including selfcitations. Therefore one might expect the number of citations of collaborative work to be somewhat higher because of the higher probability of selfcitations. The main problem, however, is to estimate the selfcitation-rate.

Selfcitations in leading research institutes in molecular biology

We recently studied the work of 13 well-known international research institutes in biochemistry and molecular biology. They had a publication rate of 5291 articles in journals between 1980 and 1984.⁴ We counted the citations during a five year period. For our purpose – evaluation of success – we had to analyse selfcitations in a detailed way. For this we used the publication lists of the institutes.

Although there was no hint for a greater success of published research based on collaboration in the total sample we found a difference in the mean citation rates in a subsample of our data (7 German research institutes, $N = 2407$). A paper with no collaboration was cited 13.4 times ((including selfcitations) on the average, a paper with at least two groups was cited 17.8 times during a five year period. Statistically this difference is significant ($t = 4.98$ $p > 0.001$). Can this difference be explained by selfcitations? To deal with this question we estimated the selfcitation-rate according to our data. We assumed the average selfcitation-rate to be the same for every collaborating group. We looked at the 3057 articles produced by all the 13 institutes that were not the result of a cooperation as indicated in the publishing journals. We found 4.3 average selfcitations per articles in a five year period. therefore we would expect that the collaborative work of two groups on the average will be selfcited 8.6 times, in the case of 3 groups 12.9 times and so on in the same period. If we compare the actual average citation per paper with the expected values based on the estimated selfcitation rate (4.3 per group) we find no hint for a greater success of collaboration (Table 1). In the case of three collaborating the groups the actual number of citations is slightly better than the expected where as the results of two collaborating groups are slightly worse. The results of the four collaborating groups might be due to the small number of papers.

To get another set of papers we studied all articles of a two years sample of the EMBO-Journal (1988, 1989, $N = 1042$). We electronically counted how often these articles were cited up to 1993 (Table 2).⁵ Citations to the articles of 1988 ($N = 529$) therefore were counted during a six-year-period. At first sight the differences in the observed values in the mean citation rate seem to be too great to be explained by the selfcitation-rate. But here the estimated selfcitation-rate used is too small. The EMBO-Journal is a high impact journal. Because of the persuading function of citations⁶ one might expect that articles in high impact journals have a higher probability to be selfcited.

Table 1
Comparison of expected and actual results of various kinds of cooperation in the case of seven German research institutes in molecular biology

	Citation per article		Number of articles	Difference
	Found	Expected		
one group	(13.4)	(13.4)	1279	-
two groups	16.5	17.7	914	-1.2
three groups	24.2	22.0	175	+2.2
four groups	20.2	26.3	33	-5.9
			2401	

The citations for 2407 articles in journals published between 1980 and 1984 were counted in a five year period. Number of research groups according to the informations given in the journal. In six cases we found more than 4 collaborating research groups.

Table 2
Average citations per paper to articles of a two years sample of the EMBO-Journal (1988, 1989) in a five/six year period

	Average citation per paper	Number of articles
one group:	47.8	582
two groups:	52.9	302
three groups:	66.0	119
four groups:	70.5	27
		1030

1041 articles published in the EMBO-Journal in 1988 and 1989 were studied. Number of the research groups according to the informations given in the journal. Citations were counted up to December 1993. In 11 cases there was a collaboration of 5 and more groups.

We reanalysed our data of the 13 institutes to deal with this question. We used the Journal Impact Factor as reported by ISI as an indicator of the importance of a journal and compare the selfcitation-rate of the 3057 articles without any cooperation with the Journal Impact Factor (Table 3). The selfcitation-rate increases with growing Journal Impact Factor. For example those 439 articles that were published in journals with a Journal Impact Factor of 3 to 3.9 were selfcited 4.2 times on the

average, those 148 articles in journals with a Journal Impact Factor of 8 to 8.9 were selfcited 6.9 times on the average.

Table 3
Comparison of the selfcitation-rate if articles in journals with increasing Journal Impact Factor

Journal Impact Factor	Selfcitation-rate	Number of articles
1-1.9:	2.6	467
2-2.9:	3.6	606
3-3.9:	4.2	439
4-4.9:	4.7	201
5-5.9:	6.0	335
6-6.9:	4.7	322
7-7.9:	5.8	75
8-8.9:	6.9	148
9-9.9:	7.3	82
10-10.9:	9.6	19
>11	7.6	107
All		2801

Articles are the articles in journals produced by 13 research institutes in biochemistry and molecular biology that were not the result of a cooperation according to the informations of the publishing journals. Journal Impact Factor as reported by the Institute of Scientific Information. 256 articles were published in a journal with a Journal Impact Factor < 1 or in a journal that was not covered by ISI.

The EMBO-Journal started in 1982. Since that time it has become an eminent journal of the field. To get a more adequate estimation of the selfcitation-rate we used the Journal Impact Factor. In 1988 the Journal Impact Factor of the EMBO-Journal was 10.9. In our sample of institutes' articles we only found 19 articles in journals with an Journal Impact Factor of 10 to 10.9 and a selfcitation-rate of 9.6. But with regard to the selfcitation-rate of articles combined with a Journal Impact Factor of 9 to 9.9 (Selfcitation-rate: 7.3) and above 11 (Selfcitation-rate: 7.6) it might be justified to take 8 average selfcitations per paper per collaborative group (see Table 3) as a rough estimation. If we compare the average citation per paper of various forms of corporation found in the sample of the EMBO-articles with the average citation per paper we would expect with regard to the selfcitation-rate we find no substantial difference (about ± 5 percent). But again the results of three cooperating

groups are slightly better, whereas the results of two cooperating groups are slightly worse than expected.

*Narin et al.*⁷ concluded that internationally coauthored work published in 1977 was cited twice as often on the average than work from only a single institute in the field of biomedical research. Nationally collaboration was placed in the middle. In this case the estimated selfcitation-rate can only reduce the difference. But another reason might explain the difference. Our analysis is based on data from leading research institutes and a well-known journal. Our conclusion is that there seems to be no systematic improvement of the success of scientific research in the case of the best. *Narin et al.* studied the scientific research efforts on a national level. Therefore their sample includes the work of the best as well as the work of the mediocre which will outweigh the best. Agreeing with *Narin et al.* we would argue that excellent research is per se international and cooperative to a certain extent. It would be interesting to see whether there is a difference in the observed development of increasing cooperations and the reported success comparing the mediocre and the leading institutes.

Table 4
Comparison of expected and actual results of various kinds of cooperation in the case of all articles of a two-years sample of the EMBO-Journal (1988, 1989)

	Average citations per article		Number of articles	Difference found/expected
	Found	Expected		
one group:	(47.8)	(47.8)	582	-
two groups:	52.9	55.8	302	-2.9
three groups:	66.0	63.8	119	+2.2
four groups:	70.5	71.8	27	-1.3
		1030		

For detailed information see legend of Table 2.

Conclusions

The selfcitation-rate – in most bibliometric studies a neglectible variable – seem to be important if the success of scientific collaboration is measured with citations. The case study in molecular biology has shown that differences in the average citations per paper can be explained to a certain extent by the higher probability of

self-citations in the case of collaborating groups. Since the number and the rules of citations vary from field to field this result can only be applied to the field of molecular biology. Furthermore it is the result of the examination of the best. *Moed et al.*² asked the question whether the indicator used (coauthorship as reported in the journals) is valid. The indicator neglects all differences in collaboration, the technical help through providing material for experiments on one side and fruitful discussion on the other side. There might be another problem which had not been analyzed yet. The rules of becoming author of a particular scientific paper may have changed in molecular biology and other fields under the pressure of "publish or perish". To provide material for experiments under the condition of becoming author of the paper, for example, is nowadays usual, but it was not in the past, as molecular biologists report.

*

I acknowledge with thanks the encouragement given by Professor Benno *Müller-Hill* and Professor Hans-Dieter *Daniel*. I thank Dr. *Kayser*, DIMDI, Cologne, for providing the data of the EMBO-Journal. This work was supported by Deutsche Forschungsgemeinschaft through Forschungsschwerpunkt "Wissenschaftsforschung".

References

1. F. MÜNZINGER, H. D. DANIEL, Biochemie und Molekularbiologie in der Bundesrepublik Deutschland. Internationale Ko-Autorenschaft und deren Einfluss auf die Zitierhäufigkeit von Veröffentlichungen 1973 bis 1987, *Futura*, 1 (1994) 20–25.
2. H. F. MOED, R. E. DE BRUIN, A. J. NEDERHOF, R. J. W. TUISSEN, International scientific cooperation and awareness within the European Community: Problems and perspectives, *Scientometrics*, 21 (1991) 291–311.
3. F. NARIN, K. STEVENS, E. S. WHITLOW, Scientific co-operation in Europe and the citation of multinationally authored papers, *Scientometrics*, 21 (1991) 313–323.
4. H. HERBERTZ, B. MÜLLER-HILL, Quality and efficiency of basic research in molecular biology: A bibliometric analysis of thirteen excellent research institutes, *Research Policy*, submitted.
5. The data were provided by Dr. *Kayser*, DIMDI, Cologne.
6. M. H. MACROBERTS, B. R. MACROBERTS, Quantitative measures of communication in science: A study of the formal level, *Social Studies of Science*, 16 (1986) 151–172.