Jointly published by Kluwer Academic Publishers, Dordrecht and Akadémiai Kiadó, Budapest

Scientometrics, Vol. 50, No. 2 (2001) 301–312

Two decades of "Scientometrics"
An interdisciplinary field represented by its leading journal

URS SCHOEPFLIN,[1,2] WOLFGANG GLÄNZEL[2,3]

[1]*Max Planck Institute for the History of Science, Berlin (Germany)*
[2]*Research Association for Science Communication and Information e. V. (RASCI) Frankfurt/Oder (Germany)*
[3]*Bibliometrics Service, Library of the Hungarian Academy of Sciences, Budapest (Hungary)*

The development of the field of bibliometric and scientometric research is analysed by quantitative methods to answer the following questions: (1) Is bibliometrics evolving from a soft science field towards rather hard (social) sciences (Schubert-Maczelka hypothesis)? (2) Can bibliometrics be characterised as a social science field with stable characteristics (Wouters-Leydesdorff hypothesis)? (3) Is bibliometrics a heterogeneous field, the sub-disciplines of which have their own characteristics? Are these sub-disciplines more and more consolidating, and are predominant sub-disciplines impressing their own characteristics upon the whole field (Glänzel-Schoepflin hypothesis)? The Price Index per paper, the percentage of references to serials, the mean references age, and the mean reference rate are calculated based on all articles and their respective references in *Scientometrics* in 1980, 1989, and 1997. The articles are classified in six categories. The findings suggest, that the field is in fact heterogeneous, and each sub-discipline has its own characteristics. While the contribution of these sub-disciplines in *Scientometrics* was still well-balanced in 1980, papers dealing with case studies and methodology became dominant by 1997.

Introduction

One of the tempting exercises in bibliometric and scientometric research may seem to be the analysis of the own field by bibliometric methods. On the one hand, it might seem worth studying the evolution of the discipline, its bibliometric characteristics, its position among related fields and its changing relations to other disciplines. On the other hand, such an exercise will immediately be confronted with and echoed by the cumulated opinions of the community of fellow scientometricians – very much in contrast e.g. to standard bibliometric research aimed at other fields, where the experts

0138–9130/2001/US $ 15.00

are more remote and feel less provoked to react on what bibliometricians claim to say about their respective disciplines.

In fact, many articles published so far stand for the lasting appeal of this topic. While most of them stay on a mere descriptive level, in the last decade several attempts have been made to characterise the evolving field as a more or less "hard" social science, thus trying to make a more analytical point. Among these attempts prominence shall be given to the papers by *Schubert* and *Maczelka* (1993) and by *Wouters* and *Leydesdorff* (1994). In both papers, the evolution of the field of bibliometrics has been studied based on literature published in the journal *Scientometrics*. Indeed, this journal covers almost the complete spectrum of bibliometric research. It publishes theoretical papers and papers on mathematical models as well as on the research evaluation of special fields and/or selected institutions, on science policy questions as well as articles on social studies of science and general discussions about the field.

Interestingly enough, the above-mentioned two papers came to quite different, even contradicting conclusions. While *Schubert* and *Maczelka* found a clear move from 'softer' towards 'harder' (social) sciences between the analysed time periods 1980-1981 and 1990-1991, respectively, *Wouters* and *Leydesdorff* concluded on the basis of the change of Price's Index in time that bibliometrics has not become a hard social science field in the observation period 1978-1992.

Glänzel and *Schoepflin* (1994) stated in their discussion paper that bibliometrics has become a heterogeneous field and sub-disciplines are drifting apart. Consequently, bibliometrics comprises sub-disciplines with distinctly different communication, citation and publication characteristics. If this statement holds, one could argue that the differences must be measurable by using of bibliometric tools. In doing so, instead of studying the journal *Scientometrics* as a whole a paper-by-paper analysis should therefore be applied. In earlier papers, *Glänzel* and *Schoepflin* (1995 and 1999) have used a set of indicators of reference literature to visualise differences between hard sciences/ medicine and other subject fields within the sciences and the social sciences, on one hand, and among different fields within the social sciences, on the other. In the following study, we will apply these indicators to prove an alternative hypothesis against the hypotheses by *Schubert* and *Maczelka* and *Wouters* and *Leydesdorff*, respectively. In particular, we will attempt to answer the following questions on the basis of measures defined on the reference literature of scientific papers.

1. Is bibliometrics evolving from a soft science field towards harder (social) sciences (Schubert-Maczelka hypothesis)?
2. Can bibliometrics be characterised as a social science field with stable characteristics (Wouters-Leydesdorff hypothesis)?

3. Is bibliometrics a heterogeneous field, the sub-disciplines of which have their own characteristics? Are these sub-disciplines more and more consolidating, and are predominant sub-disciplines impressing their own characteristics upon the whole field (Glänzel-Schoepflin hypothesis)?

Assuming that bibliometrics is an interdisciplinary field and that authors coming from different fields bring their specific communication behaviour into it, we have classified all papers published in *Scientometrics* into different categories representing the main field-specific approaches to bibliometrics. We will extend the analysis to six sub-disciplines as well as to the whole field represented by the contributions in the journal *Scientometrics* in three different publication years.

Methods

All source articles published in the journal *Scientometrics* in three sample years, 1980, 1989 and 1997, have been processed. All references cited in articles, notes and letters in the above three publication years were selected. Review articles have not been taken into consideration since the extent and structure of the reference lists of these documents are expected to distinctly differ from those of other research papers. Papers without references have been omitted. The age of the references was determined as the difference between the publication years of the reference and that of the citing article respectively. In case of missing publication years of the reference (e.g. unpublished material, or papers indicated as being 'in press'), the year was substituted by the publication date of the corresponding document.

References have been assigned to two categories, reference to serials (*S*) and reference to non-serials (*N*). All references have been classified manually.

The following statistics have been calculated.

1. *The Price Index per paper.* This index is defined as the percentage of references not older than five years in all references of an individual paper. This indicator has been introduced by *Moed* (1989).
2. *The percentage of references to serials.* The share of references assigned to category *S* in all references (*N*+*S*) cited by a journal or subfield expressed in percent.
3. *The mean references age.* The age of references cited in a journal or subfield are summed up and divided by the number of the references. This indicator can be determined also as a conditional mean, that is for both the subset of references in serials and non-serials separately.

4. *The mean reference rate.* This is the ratio of the number of references cited by a journal and the total number of papers published in the journal including those without references.

The "Price Index" is commonly used as a measure to distinguish between hard science, soft science, technology and non-science (see *Price*, 1970). For the present study and to allow a paper-by-paper analysis we have calculated the Price Index according to the original definition given by *Price* and the "Price Index per paper" as defined by *Moed* (1989).

According to the results of an earlier study (*Glänzel* and *Schoepflin*, 1999), the percentage of references to serials proved to be a sensitive measure to characterise typical differences in the communication behaviour between the sciences and the social sciences. The mean reference age also serves as an efficient measure of the "hardness" of science. In the paper by *Glänzel* and *Schoepflin* (1999), a comparison of the mean age of references and the Price Index has shown that the age of references is only in part reflected by the Price Index, in particular if the average age of references does not exceed about 15 years. In addition, we calculated the mean reference rate, that is, the "average size" of the reference list of a bibliometric paper published in *Scientometrics*. Although this cannot be considered a sensitive measure of the "hardness" of science, it reflects nevertheless additional field-specific characteristics (see *Glänzel* and *Schoepflin*, 1999).

In the next step, all selected source articles in *Scientometrics* have been assigned manually to one category out of a scheme of six. This scheme was defined to characterise the scope of bibliometrtic and scientometric research as a whole. It is designed to take into account not only the present state of research in the field, but to reflect fully the scope of research. The classification scheme used for this study is presented in Table 1. The classification permits to group the material in several ways: the categories can be regarded from the viewpoint of *core bibliometrics* (2, 3, and 4) and *background research* (1, 5, and 6), but also with respect to *theoretical* (1, 3, and 5) and *applied research* (2, 4, and 6).

The assignment of papers to one single category was in part not without difficulty and in a sense somewhat arbitrary. The borderlines between these categories were – particularly in the first volumes of *Scientometrics* – somewhat fuzzy. As a general rule, assignment was performed along the following lines: a paper was, for instance, assigned to category 2 if the methodology presented and applied in the paper is incidental to the purpose of the case study. Otherwise, if the methodology was new and a predominant

component of the paper, it was in spite of a possibly extensive application assigned to category 3. The same rules were applied to other categories, too. Table 2 presents the number of papers assigned to each category in 1980, 1989 and 1997.

Table 1
Classification of papers published in the journal *Scientometrics*

Category	Description
1	Bibliometric theory, mathematical models and formalisation of bibliometric laws
2	Case studies and empirical papers
3	Methodological papers including applications
4	Indicator engineering and data presentation
5	Sociological approach to bibliometrics, sociology of science
6	Science policy, science management and general or technical discussions

Table 2
Classification of papers published in the journal *Scientometrics* by categories

Category	1980	1989	1997
1	1	6	3
2	5	21	35
3	6	24	25
4	2	6	1
5	5	9	6
6	12	1	5
Total	31	67	75

Results

The indicators have been determined for all papers of the type Article, Letter and Note published in 1980, 1989 and 1997, that is for all categories combined. First, they have been calculated for the journal, and thereafter, for each paper, separately. The median of these publication-based indicators can then be compared with the corresponding journal indicators. This shows the direction of deviation of paper-based indicators from the corresponding journal indicators. The values are presented in Table 3.

Table 3
Reference-based indicators for three publication years of *Scientometrics*
(First row: journal indicators, second row: median of indicators for individual papers)

Year	Documents	References per paper	Price Index	Share of serials	Mean age of refs.
1980	31	16.2	33.4%	44.1%	12.5
		11.0	42.9%	50.0%	7.0
1989	67	19.6	35.2%	62.7%	16.1
		14.0	37.5%	64.3%	8.7
1997	75	17.8	35.7%	58.0%	10.4
		14.0	38.5%	60.0%	8.2

The data presented in Table 3 show a surprising similarity for the years 1989 and 1997. The values for 1989 are somewhat below those of the later years. The Price Index has been calculated as the share of references not older than 5 years in all references, that is, strictly according to the definition given by de Solla Price. In our case, the Price Index is thus the share of references to the years 1976-1980 in all references for the publication year 1980. The same applied *m. m.* to the publication years 1989 and 1997, too. Papers not assigned to the document types Article, Letter and Note have been omitted. The values for the journal indicator proved to be lower than the values presented by *Wouters* and *Leydesdorff* (1994). The reason can not be cleared up here, since the way in which they have calculated their index is not documented. However, the Price Index shows stability also in our case. Even more, all indicator values reflect patterns typical of social-science journals (c.f. *Glänzel* and *Schoepflin*, 1999). It is worth mentioning here, that the median of paper-based indicators is greater both for the Price Index and the share of serials than the corresponding journal indicators. This phenomenon allows only one possible interpretation: Numerous papers are 'harder' than expected on the basis of the overall journal indicators. The plot of the Price Index vs. the mean age of references is presented in Figure 1. The comparison of the distribution of the share of references to serials in the three years 1980, 1989 and 1997 (see Figure 2) with the overall patterns of all social science fields combined (see *Glänzel* and *Schoepflin*, 1999) reveals an interesting similarity. Except for the 'low end' (0%-10% in 1980) and the 'high end' (90%-100% in all three years), the serials-share distribution of *Scientometrics* is that of a typical social science journal.

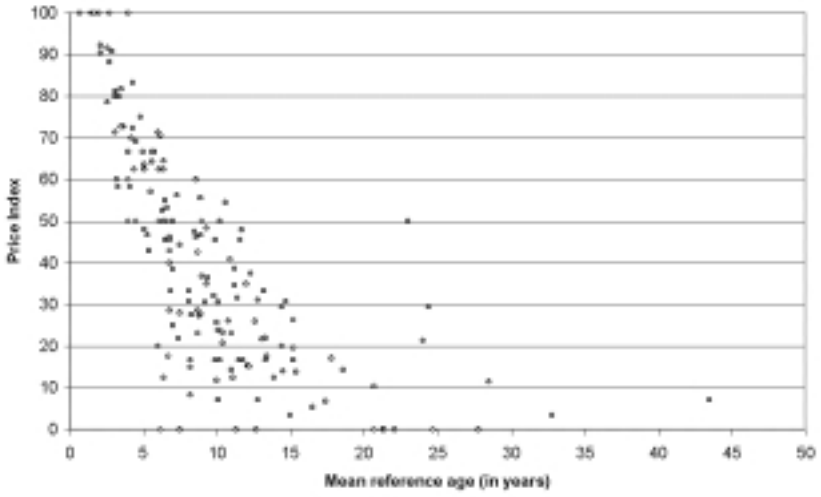

Figure 1. Plot of Price Index vs. 'mean reference age'

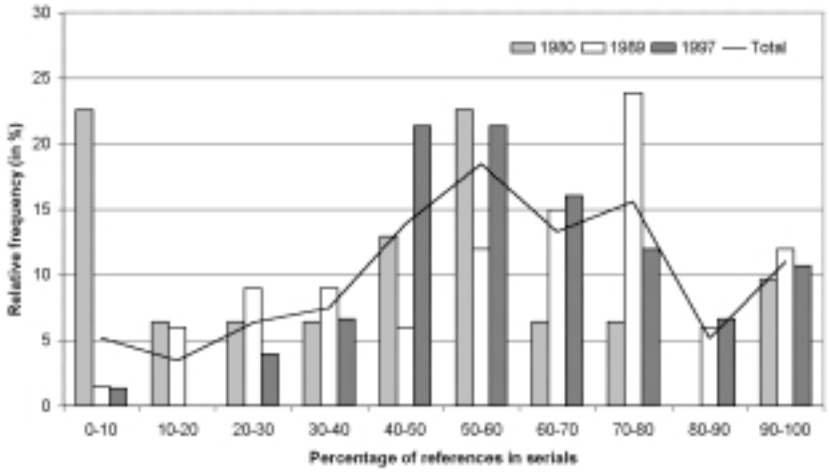

Figure 2. Distribution of the share of references to serials in 1980, 1989 and 1997

The influence of the individual categories on the journal total over time is of special interest for characterising the developments of the field. Therefore we will now have a look at the six categories. Figure 3 presents the distribution of papers over document categories in the three time periods. There are obviously two dramatic developments: first, there is an impressing and steady growth of Case Studies, from a forth position in 1980 to the predominant first position in 1997. Second, there is a similarly impressive loss of share of articles on Science Policy and Discussions (category 6) from the predominant first position in 1980 to a minor category in 1997. This goes along with a steady loss in material with a sociological approach, too (category 5). On the other hand, there is a certain increase in Methodology (category 3), while Theory and Indicator Engineering remain minor classes.

If we now take a look on Core Bibliometrics as defined by categories 2, 3, and 4, it becomes obvious, that this group is practically reduced to Case Studies and Methodology, and by far dominating the total output of research as represented by the journal *Scientometrtics*. The group characterised as Background Research (categories 1, 5, and 6) has continuously lost ground since 1980. Moreover, theoretical research in bibliometrics seems to be mainly a matter of methodology.

Of course, in comparing these output indicators over time, the shift in publication opportunities has also to be considered. Following the differentiation of the field, journals like *Social Studies of Science* or the newly founded *Research Evaluation* publish a considerable share of bibliometric research in categories 5 and 6. But also journals in information science (e.g. *JASIS*, *Information Processing & Management*, or *Journal of Information Science* to name just a few Anglo-Saxon titles) attract bibliometric research articles. On the detriment of a larger scope, *Scientometrics* has clearly become the forum for Case Studies and Methodology-oriented contributions.

This is also proven by the deviating figures for 1980 in Table 3 as well as the exceptionally high frequency of papers with very few references to serials in 1980. This might be caused by a shift from a science policy approach towards methodology and empirical studies. If this statement proves to be true, we will have evidence of structural changes in the field. The median of the values of the Indicators 2-4 shed some light on this phenomenon (see Table 4).

The above-mentioned deviating patterns in 1980 can be at least in part explained by the great share of papers in category 6 and their low share of serials and relatively low age of the references. Indeed, the indicator values of the six categories give evidence of specific characteristics of the corresponding sub-disciplines.

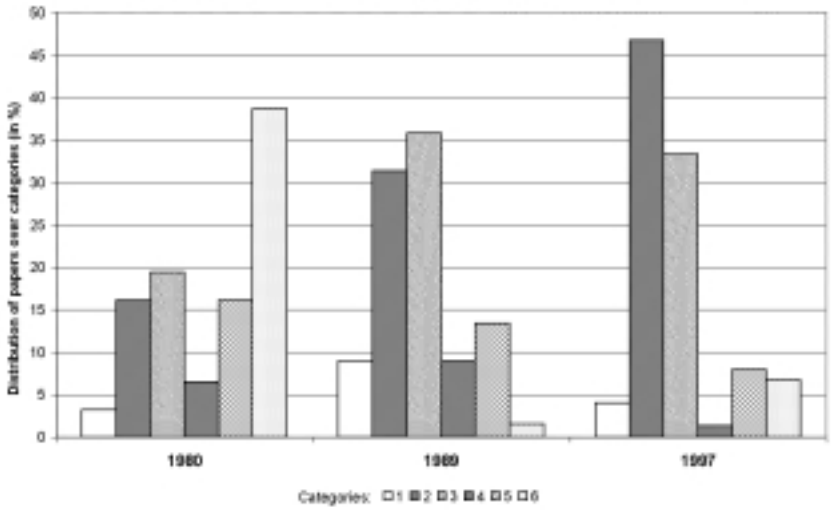

Figure 3. Distribution of papers published in 1980, 1989 and 1997 over document categories in the journal *Scientometrics*

Table 4
Reference-based indicators for six categories and all categories combined

Category	Percentage share of serials (median)	Mean reference age (median)	Mean reference rate
1	66.7	12.3	14.0
2	59.1	8.1	18.4
3	61.5	7.5	18.9
4	80.0	4.2	17.3
5	54.3	11.4	26.0
6	29.2	7.2	9.8
All categories	59.1	8.2	18.2

The medians of the indicator values of category 5 are close to those of Sociology in 1993 (cf. *Glänzel* and *Schoepflin*, 1999: 40.4% (share of serials), 12.5 (mean reference age), 32.7 (references per paper)). As expected, category 1 resembles to the field of Mathematics in 1993 (64.4% (share of serials), 11.3 (mean reference age), 16.2

(references per paper)). Categories 2 and in part 3 can be rather characterised as engineering (cf. Electronic Engineering in 1993: 62.2% (share of serials), 8.6 (mean reference age), 15.0 (references per paper)). However, the extreme share of serials in connection with fast ageing categories 4 and 6 already refer to technology and non-science oriented literature, respectively. We would like to stress that neither sub-discipline resembles to the social science field Library and Information Science in 1993 although bibliometrics was one of its subfields. The reason can be found in different characteristics of other journals in information science and distinctly different patterns of the other main subfield of Library Science.

The distribution of the mean reference age over papers published in *Scientometrics* in 1980, 1989 and 1997 in categories 2, 3 and 4 and in categories 1, 5 and 6 are shown in Figure 4. The difference in the shapes of distributions between the two groups but also within each group is remarkable. They support the hypothesis that in fact there are distinct sub-disciplines which have their own characteristics. The following interpretation seems plausible, but of course will need further evidence, which cannot be given here: the great share of slowly ageing papers in category 1 is in keeping with peculiarities in mathematics, the atypical share slowly ageing papers in category 2 is caused by publications devoted to the bibliometric analysis of selected subfields in the sciences.

Discussion

The above trends and figures tell unambiguously against an evolution of bibliometrics towards a discipline of 'hard' social science (Schubert-Maczelka hypothesis). On the other hand, we cannot speak of stable characteristics, either (Wouters-Leydesdorff hypothesis). The indicators allow only an interpretation in the sense of the third hypothesis. The field is indeed heterogeneous, and each sub-discipline has its own characteristics. This may, of course, be at least in part caused by the deviating field-specific communication behaviour of the authors who bring traditional organisation schemes from their own fields into bibliometrics. This is, however, contradicted by several authors with many published papers in the journal *Scientometrics*. The same is true for the special issues of the journal devoted to a specific topic or to a conference. The overall influence of these issues seem rather small.

In all, we can conclude that the sub-disciplines here presented by six categories have more and more consolidated. While the contribution of these sub-disciplines in *Scientometrics* was still well-balanced in 1980, case studies and methodology became dominant by 1997. Today, these sub-disciplines determine the characteristics of the whole field, whereas in 1980 bibliometrics was still an average and superposition of their relatively well-balanced components.

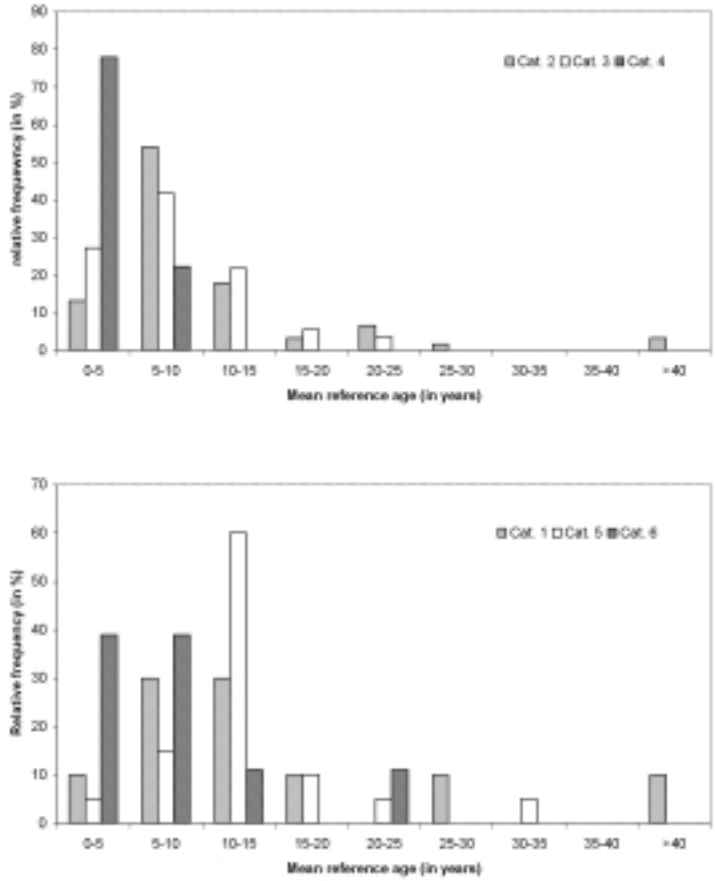

Figure 4. Distribution of the mean reference age in categories 2, 3, 4 (top) and 1, 5, 6 (bottom)

References

L. EGGHE (1997), Price Index and its relation to the mean and median reference age, *Journal of the American Society for Information Science*, 48 : 564–573.

W. GLÄNZEL, U. SCHOEPFLIN (1994), Little scientometrics – Big scientometrics ... and beyond, *Scientometrics*, 30 : 375–384.

W. GLÄNZEL, U. SCHOEPFLIN (1995), A bibliometric ageing study based on serial and non-serial reference literature in the sciences, *Proceedings of the 5th International Conference on Scientometrics and Informetrics*, held in River Forest, Illinois, June 7-10, Learned Information Inc., Medford, pp. 177–185

W. GLÄNZEL, U. SCHOEPFLIN (1999), A bibliometric study of reference literature in the sciences and social sciences, *Information Processing and Management,* 35 : 31–44

H. MOED (1989), Bibliometric measurement of research performance and Price's theory of differences among the sciences, *Scientometrics*, 15 : 473–483

A. SCHUBERT, H. MACZELKA (1993), Changes in Scientometrics, *Social Studies of Science*, 23 : 571–582

D. J. DE SOLLA PRICE (1970), Citation measures of hard science, soft science, technology and non-science. In: C. E. NELSON, D.K. POLLAK (Eds): *Communication among Scientists and Engineers*, Heat, Lexington, Mass., pp. 1–12

P. WOUTERS, L. LEYDESDORFF (1994), Has Price's dream come true? Is scientometrics a hard science?, 31 : 193–222

Received October 13, 2000.

Address for correspondence:
URS SCHOEPFLIN
Max Planck Institute for the History of Science
Wilhelmstraße 44, D-10117 Berlin, Germany
E-mail: schoepfl@mpiwg-berlin.mpg.de