

BIBLIOMETRIC LITERATURE: A QUANTITATIVE ANALYSIS

Bluma C. PERITZ

Senior Lecturer, Hebrew Univ. Jerusalem, Graduate Library School, P.O.
Box 503, Jerusalem 91904, Israel

Abstract

A preliminary analysis of the literature of bibliometrics for the period 1960-1985, based on several secondary sources, is presented. In spite of the inherent limitations of such investigations a number of relevant variables were extracted: 1) the year of publication, which allows the study of growth or other changes over time, 2) language, 3) the field of the journal in which the paper was published, 4) the subject area within bibliometrics with categories such as growth, scatter, obsolescence, journal cores, bibliometric laws, citation analysis, etc. Furthermore, a substantial number of papers in this body of data are concerned with the bibliometric analysis of some specific discipline; the distribution of these papers according to the discipline in question is an additional variable in this study. Frequency distributions of these variables and selected cross-tabulations, in particular those highlighting changes over time, are presented and analysed statistically. Some tentative causal interpretations are also suggested.

1. INTRODUCTION

In the introduction to the supplement of his Bibliography of Bibliometrics and Citation Indexing and Analysis, Hjerppe states that he avoided "the temptation to apply various bibliometric techniques to the present bibliography". Before you stands a person who has succumbed to this temptation. True, the moment may not be the most propitious for such an attempt. We hear, indeed, that Pritchard is working on a comprehensive multi-volume bibliography, the first volume of which has already appeared [1], while the second is available in machine-readable form. When completed, such a bibliography will, of course, hold great promise for a definitive self-study of bibliometrics. Nevertheless, our first international conference (as well as my own curiosity) have prompted me to attempt at least a first, preliminary analysis of Hjerppe's and his successor's work. This analysis will cover the years 1960 to 1985, thus leaving aside the "pioneering" period covered in Pritchard's first volume: 1874-1959. The objective of such a preliminary analysis will be to help us ask the right questions, rather than obtain the correct answers.

2. METHODS

The population studied in this survey consists of: a) Hjerppe's bibliography [2], covering the period 1960-1979, b) the supplement published by Hjerppe [3] which includes additional material, mainly from May 1978 until the end of 1979 and goes on to cover the year 1980, and c) the bibliographies of Schubert [4] which supplement Hjerppe for 1980 and go on to the end of the study period. The last bibliography of Schubert included in the present analysis dates from 1986. No doubt, later issues of this periodical bibliography still include material from 1985

and before, since the delay from publication to inclusion in the bibliography is about two years.

From Hjerpe's first bibliography the following were excluded: 93 publications published before 1960 and 19 items listed without year of publication. The final number of publications for the present analysis is 3225. Out of these publications 2675, or 83%, are papers in journals and proceedings.

The following variables are included in the analysis: period of publication, language, and three content-related variables:

a) The field of the journal in which the paper was published. This variable was evidently taken into account only for papers published in journals (or proceedings). A simple ad hoc classification into fields was devised; whenever the field was not evident from the title of the journal Ulrich's directories [5,6] were perused. A fairly large number of papers were published in the general scientific press (e.g. Science, Nature, etc.); they appear as "science, general" in this classification.

b) The discipline studied. This refers to the scientific discipline which is the object of bibliometric investigation in the publication in question. The categories were made as similar as possible to those of the field of the journal, but the method used was slightly different: the present variable is derived by analyzing the title of the publication. Thus, an item whose title includes the term, say, "journals in physics" is included among the publications whose "discipline studied" is physics. Consequently, a comparative study dealing with two disciplines will be included in each of them. Although the title publication will list more than one discipline only very rarely, in principle a certain number of "double counts" will appear under this variable. On the other hand, many bibliometric studies do not deal with a specific discipline - or at least do not mention any discipline in the title; obviously, such studies do not appear in the classification by "discipline studied".

c) The subject within bibliometrics. This variable refers to the area of bibliometrics used in the study in question. Here too, the title is used as the source of data. Only a section of specifically "bibliometric" subjects are used for this variable, and therefore the enumeration is far from exhaustive. On the other hand, double counts abound in this variable: thus, a study may use citation analysis in order to investigate obsolescence. Such a study will be counted both under citation analysis and under obsolescence. Whenever relevant, synonyms and other related terms (such as Bradford's law and Zipf's law) are brought together in the same subject category. When two such terms appear in the same title they are counted only once.

Cross-tabulations of the three content-related variables are likely to yield interesting results. We did not, however, include them in the present preliminary analysis.

The sources on which this analysis is based are, necessarily, selective. It is assumed that the criteria for selection, both explicit and implicit, have remained constant throughout the period covered, although in 1983 Schubert changed the title of the bibliography to "Quantitative Studies of Science". The completeness of the bibliography does fall off in the years near 1985, when data collection was closed. Nevertheless, it is not clear a priori whether this affected the internal distribution of the items according to the variables under study.

A very partial check relating to this issue was introduced by means of the variable "bibliographical source". Briefly, this variable allows one to see whether or not the exclusion of the Schubert bibliographies, or even of Hjerpe's supplement, has any important effect on the distributions according to the other variables.

3. RESULTS

Bibliometrics is evidently still a young field (although we hardly need any data to remind us of this!). According to Table 1 the total number of publications increased from 138 in 1960-65 to 1147 in 1976-80. If we take into account the fact that the first period numbered six years against five in all the other periods, this represents a tenfold increase in roughly fifteen years. The increase in the numbers of papers published in journals (also in conference proceedings) was even steeper. From the growth indices in the second half of the table one sees that there was a 44% increase in journal papers in the five years between 1971-75 and 1976-80.

Table 1. Publications in bibliometrics by period and type of publication 1960-1985
Absolute numbers and growth indices (1971-1975 = 100)

	All Periods	1960- 1965	1966- 1970	1971- 1975	1976- 1980	1981- 1985
Absolute numbers						
All publications:	3225	138	452	840	1147	648
Papers in journals and proceedings	2675	94	245	696	1000	640
Books, reports, diss., coll. works, occ. papers, etc.	550	44	207	144	147	8
Growth indices						
All publications		16	54	100	137	77
Papers in journals and proceedings		14	35	100	144	92
Books, reports, diss. coll. works, occ. papers, etc.		30	144	100	102	6

After 1980 there was an apparent decrease in the number of publications, in particular for books, reports, dissertations. It seems, however, that at least part of this is artefactual. The year-by-year counts of the total numbers of publications from 1980 till 1985 throw some light on this issue.

Table 2. Publication in bibliometrics
Single years of publication 1980-1985

Year	Absolute no.	%
1980	285	30.5
1981	160	17.1
1982	218	23.4
1983	183	19.6
1984	70	7.5
1985	17	1.8
Total	933	100.0

In 1980 the increasing trend from previous years is continuing; there are 285 publications, as against a yearly average of 229.4 for 1976-80. In 1981 there is an unexplained drop, followed by a slight rise; from 1983 onwards the numbers are decreasing again - very sharply so since 1984. Some, if not most, of the decrease in the last two or three years is due to the usual delay between publication time and being recorded in a bibliography. The earlier decrease, however, between 1980 and 1981-1982 is more difficult to explain; it might even reflect a genuine reduction in the publication of new bibliometric research.

Turning now to the fields of the journals that publish bibliometric studies (we included here the relatively small number of conference proceedings) one finds that about one half of them are in library and information science. This percentage stayed about the same since 1971 and was only slightly less (about one third) before that year. Another important outlet for bibliometric papers is the general (non-specialized) scientific press. About 9% of the papers are published there. Journals in history or sociology of science also publish a substantial number of such papers.

Table 3. Bibliometrics papers in journals and proceedings by field of journal and period 1960-1985

Field of Journal	PERIOD					All Periods	
	1960 -65	1966 -70	1971 -75	1976 -80	1981 -85	Absolute numbers	%
Library and information sciences 1)	39	75	323	511	332	1280	47.9
History and sociology of science 2)	1	6	25	115	35	182	6.8
Mathematical sciences 3)	6	5	12	14	-	37	1.4
Physics, astronomy	2	7	13	24	37	83	3.1
Chemistry, pharmacology	3	9	15	28	20	75	2.8
Technology, engineering	3	15	30	21	9	78	2.9
Life sciences, genetics	1	4	7	19	13	44	1.6
Agriculture	-	1	3	5	3	12	0.4
Medicine 4)	2	8	23	48	31	112	4.2
Earth Sciences 5)	3	3	9	8	6	29	1.1
Economics, business, management 6)	1	6	36	18	24	85	3.1
Sociology, communications 7)	5	30	45	39	34	153	5.7
Psychology	4	7	15	56	30	112	4.2
Education	1	-	8	12	13	34	1.3
Social Sciences, others and general (law, criminology)	1	3	28	20	16	68	2.5
Humanities	-	1	3	6	-	10	0.3
Science, general	21	62	94	44	26	247	9.2
General, other, unknown	1	3	7	12	11	34	1.3
Total	94	245	696	1000	640	2675	100.0

1. Including documentation, publishing
2. Including scientometrics
3. Including statistics, computer science
4. Including clinical psychology, public health
5. Including geology and geography
6. Including operations research
7. Excluding sociology of science

Papers published in journals from other disciplines are often studies in self-evaluation of these fields. The main exceptions to this rule are sociology and psychology: journals in these areas also publish material which does not pertain directly to self-evaluation. In any case, the number of bibliometrics papers appearing in these journals is substantial. The percentage of papers from all the social sciences was 17%; if one excludes from this count economics and management, it was still 14%. A surprisingly small number of publications comes from journals in education.

In the science proper, a substantial share was published by medical journals, with 112 papers (4%). The two other leading "scientific" contributors to bibliometric literature were journals in physics and chemistry. The number of papers from all the specialized scientific journals (including engineering and technology) was 470, or 18%. If one takes all scientific journals, both general and specialized, one gets no less than 717 papers, or 27%.

Table 4 gives us some idea about the trends in the bibliometric papers published by the various kinds of journals. The growth indices for papers from library and information science journals were about the same as those for all journals (see Table 1). History and sociology of science, on the other hand, show a particular rapid growth. Other areas of steep increase were psychology, medicine, physics and chemistry, while the impact of the general scientific press is decreasing.

Table 4. Bibliometrics papers in journals and proceedings by selected field of journals and period 1960-1985
Indices of Growth (1971-1975 = 100)

Field of journal	1960 -65	1966 -70	1971 -75	1976 -80	1981 -85
Library and information science	12	23	100	158	103
History and sociology of science	4	24	100	460	140
Physics, astronomy	15	54	100	185	285
Chemistry, pharmacology	20	60	100	187	133
Technology, engineering	10	50	100	70	30
Medicine	9	35	100	209	135
Economics, business, management	3	17	100	50	66
Sociology, communications	11	67	100	87	76
Psychology	27	47	100	373	200
Science general	22	66	100	47	28

The data on the specific disciplines investigated by bibliometric studies are given in Table 5. Here we deal with all publications, not just journal papers. The frequencies are given separately for each of the three bibliographic sources.

There were 1363 publications on specific disciplines investigated, making up for 42.3% of all publications. This percentage does not vary much from one bibliographic source to another. The main disciplines studied were medicine, physics, engineering and technology, chemistry and psychology, in that order.

It is worth noting that the numbers of publications dealing with sociology is substantially less than the number of bibliometric papers published in sociology journals. This is so because many papers in sociology of science and other related areas (e.g. all those dealing with the so-called "Matthew effect") use bibliometric techniques and appear in sociological journals and yet do not relate in any way to sociology itself. The same is true, although to a lesser extent, for psychology journals.

Table 5. Bibliometrics publications by bibliographical source and by discipline studied, 1960-1985

Discipline	HJERPPE I 1960- 1970	HJERPPE SUPPL. May 1978- Dec. 1980	SCHUBERT 1980-1985	TOTAL	TOTAL %
Mathematics (incl. computer sci.)	39	15	5	59	1.8
Physics (incl. astronomy)	106	22	70	198	6.1
Chemistry (incl. pharmacology)	64	29	38	131	4.1
Engineering, technology	108	37	12	157	4.9
Life sciences, agriculture	39	14	30	83	2.6
Medicine	150	50	48	248	7.7
Earth sciences (geology, geography)	31	9	5	45	1.4
Sociology	76	6	11	93	2.9
Psychology	56	22	28	106	3.3
Education	32	12	4	48	1.5
Economics, business, management	46	15	15	76	2.3
Other social sciences	11	23	27	71	1.9
Humanities	28	22	8	58	1.8
Total disciplines studied	786	276	301	1363	42.3
Grand Total	1920	518	787	3225	100.0

The specifically bibliometrical subjects treated in these publications are shown in Table 6. As mentioned before, the only source of information for this breakdown is the title. Also not infrequently more than one such subject is mentioned in the title. The list of subjects selected for this table is limited to specifically bibliometric issues. Citation indexes per se, and papers about them, for instance, are not included, although there is a vast literature on this subject. Selected subjects in bibliometrics were mentioned 1324 times in the titles of these publications. Ignoring the fact that there are "double counts", those mentioned are 41.1% of the total. Slightly over one half of them deal with the various aspects of citation analysis. Bibliometrics laws, i.e. mostly Bradford's law and its "relatives", account for 200 mentions, or 6.2%. The generic group of "theory and models" comes next (in this group there is a relatively large number of Russian authors). "Growth", "obsolescence", "core" and finally "scatter" represent, in this order, most of the preoccupations of practising bibliometrists. For this preliminary analysis I selected only one so-called "external" characteristic of the publications, namely language. The data are presented in Table 7.

Table 6. Publications on selected subjects in bibliometrics by bibliographical source

Subject	HJERPPE I	HJERPPE Suppl.	SCHUBERT	Total	Total %
Growth	61	9	48	118	3.7
Scatter	20	2	11	33	1.0
Core	26	4	38	68	2.1
Obsolescence (incl. aging, half-life, decay)	55	17	19	91	2.8
Laws (incl. Bradford, Zipf, Lotka, others)	114	32	54	200	6.2
Theory, models	92	6	20	118	3.7
Citation analysis (1)	419	80	197	696	21.6
Totals of above subjects	787	150	387	1324	41.1
Grand Total	1920	518	787	3225	100.1

1 Includes co-citations, coupling, clustering, but does not include citation indexes as such.

As in many other discipline, English is the language most frequently used. Since 1966 about three quarters of the publications were in English, in each of the quinquennial periods. During the first period under study, however, the percentage was 97,8%, in the early period of Pritchard's first volume the percentage was also over 90%.

The second language is Russian, followed by German. There were also substantial contributions in Spanish (particularly in recent years); Czech and Hungarian publications follow next. In all, 246 publications (7.6%) come from East-European countries. Finally, Japanese and French are also represented.

It is of some significance to notice in this context that the language of the publication is not necessarily the language of the country in which the journal is being published. In fact, increasing numbers of journals in non - English speaking countries are published in English.

4. DISCUSSION

The few data presented here can serve only as a first orientation within the field of bibliometrics. They are essential descriptive and the comments needed at this point are mostly on methodological limitations.

The bibliographical sources used in this paper, are, of course, selective, and consciously so: thus, for instance, use and user studies are generally excluded and scientometrics and related fields are included only in part. Perhaps the more comprehensive bibliography of Pritchard [1] will give us a more complete picture. In general, the bibliographic control of our field is not yet satisfactory and we shall need a good "bibliometric population base" even if eventually the self-study of our field will turn to the analysis of more stringently defined research materials in specific subfields.

Highly productive authors pose another problem of selection: Should Garfield's 808 Essays of an Information Scientist have been included entirely or selectively (as Hjerpe actually did)? Should some of his briefer contributions have been excluded? In this instance, at least, decisions on a single author may have an impact on the statistical data for the entire field.

It seems that at least the problem of definition of our field is now settled. In 1987 the term bibliometrics has finally been introduced into the Encyclopedia of Library and Information Science [7] although it had been proposed in 1969 and has been heavily used since then [9]. Scientometrics has become even more clearly related to bibliometrics - almost to the point of synonymity - as attested by Library and Information Science Abstract (LISA).

The main question for this audience relates of course to the direction in which bibliometrics is developing. A few hints on this can be gleaned from the present material. Firstly, citation analysis remains a major tool, whose potential is far from exhausted. Secondly, the literature on bibliometric laws is vast and growing, both at the theoretical and the empirical level, including applications in librarianship and elsewhere. Thirdly, the bibliometric self-study of various scientific disciplines is a constant and growing preoccupation. Budgetary restrictions have made it even more imperative now for scientists to look critically at their own field. This is also true for librarians and information scientists. (As a bibliometrist I feel compelled to remark that this probably leads to a very high scatter of our literature).

Finally, one feels impressed by the truly international nature of our "invisible college". Statistics by language do not do justice to this aspect of our discipline.

Table 7. Bibliometrics publications, by language and period 1960-1985

Language	1960-1965	1966-1970	1971-1975	1976-1980	1981-1985	Total	%
English	135	335	618	899	502	2489	77.1
Russian	1	73	124	62	43	303	9.4
German	2	4	21	63	29	119	3.7
Spanish	-	-	9	27	27	63	1.9
Czech	-	16	16	16	2	50	1.6
Hungarian	-	7	8	15	9	39	1.2
Japanese	-	3	8	20	7	38	1.2
French	-	3	5	15	6	29	0.9
Portuguese	-	-	4	11	11	26	0.8
All others	-	11	27	19	12	69	2.1
Total	138	452	840	1147	648	3225	100.0

Indian scholars are very active in our field and so are lately scientists from Latin American countries and from Nigeria, etc. Eastern European countries have a long and respectable tradition in scientometrics and bibliometrics. Many countries and many disciplines come together in the creation of our discipline. This gives us the encouraging feeling of being part of a significant worldwide endeavour.

REFERENCES

- [1] Pritchard, A. and Wittig, G.R., *Bibliometrics: a Bibliography and Index*, Vol. 1: 1874-1959. (Watford: ALLM Books, 1981).
- [2] Hjerpe, R., *A Bibliography of Bibliometrics and Citation Indexing and Analysis*, (Stockholm: The Royal Institute of Technology, Library, 1980), (Report TRITA-LIB-2013)
- [3] Hjerpe, R., Supplement to a Bibliography of Bibliometrics and Citation Analysis. *Scientometrics*, 4 (1982), pp.241-281.
- [4] Schubert, A., Quantitative Studies of Science: A Current Bibliography. *Scientometrics*, 5 (1983), pp.125-133, pp.397-403, *Scientometrics*, 6 (1984), pp. 55-59, pp.197-202, pp.359-367, *Scientometrics*, 7 (1985), pp.111-117, *Scientometrics*, 8 (1985), pp.137-140, *Scientometrics*, 9 (1986), pp.293-304.
- [5] *Ulrich's International Periodicals Directory*, (New York: Bowker, various editions).
- [6] *Ulrich's Irregular Serials and Annuals*, (New York: Bowker, various editions).
- [7] Hertz, D.H., *Bibliometrics, History of the Development of Ideas in*: Kent, A., (ed.) *Encyclopedia of Library and Information Science*, 42, pp.144-219, (New York: Marcel Dekker, 1987).
- [8] Peritz, B.C., On the Careers of Terminologies: the Case of Bibliometrics, *Libri*, 34 (1984), pp.233-242.