Social Science Information Studies (1981), *1* (67–87)

THE STRUCTURE OF SOCIAL SCIENCE LITERATURE AS SHOWN BY A LARGE-SCALE CITATION ANALYSIS

MAURICE B. LINE*

Director General, British Library Lending Division

ABSTRACT

To collect information relevant to the improvement of secondary services, an analysis of 59,000 citations in the social sciences was conducted—11,000 taken from 300 monographs, and 48,000 from 140 serials, including 47 highly cited titles and 47 taken at random. The analyses, the main results of which are summarized, included concentration and scatter, rank lists of cited serials, date distributions (crudely corrected for literature growth), subject relationships, country and language links, and forms of material cited. Large differences were apparent between analyses of references taken from serials and those from monographs, and smaller differences between analyses of references from highly cited and randomly chosen serials. This suggests that analyses based only on references from highly cited serials give a very incomplete picture. There were also great differences between subjects; most of these differences have implications for the planning of secondary services. The results imply that secondary services in the social sciences are deficient in the range of forms of material they cover, in the number and possibly the selection of serials covered, in the subject spread of material of possible relevance, and probably also of foreign language material and material published in many other countries.

Citation analyses have become a popular exercise in recent years, especially since the advent of the *Social Science Citation Index* made citation counting much easier. It is difficult to avoid the suspicion that many of the analyses, which have by no means been conducted only by information scientists, have been mere academic exercises, but citation analyses have also been seen as a relatively easy, and certainly unobtrusive, way of shedding light on the structure of knowledge, research and literature usage—part of the new discipline of scientometrics. They can help to answer such questions as to how much use researchers and other writers make of the literature of the past, of literature in languages other than their own (and in which languages), of subjects beyond their own immediate discipline, and so on. By helping to indicate the present state of affairs citation analyses can raise other more fundamental questions: if, for example, they reveal geographical or linguistic parochialism, especially if this characteristic is more

* Maurice Line is Director General of the British Library Lending Division. He was formerly Librarian of the National Central Library before its merger with the British Library and, before that, Librarian of Bath University where he directed the two major social science information projects INFROSS and DISISS. He retains his interest in the information needs of social scientists as well as in bibliometrics and directs IFLA's programme of Universal Availability of Publications.

prominent among researchers in one country than in others, it may be useful to ask why and to seek remedies. There may be more specific and immediately practical reasons why citations are collected and analysed, as in the present case; what is certainly true is that information of this kind is very hard to gather in other ways.

Several writers have expressed reservations about the validity of citation analyses, pointing out that not all that is cited is read, that not all that is read is cited, and that some citations are negative. Further questions concern the practical value of citation analyses; in particular, it is hard to see how they can be put to practical use in libraries (Line and Sandison, 1975; Line, 1977). However, they can, *if properly planned and executed*, reveal a good deal about the way information is used, by showing relationships between countries of publication, languages of publication, and subjects. They can also produce rank lists of serials in order of citations received, and group serials into families by clustering the citations they receive in common. Such information can help in the design of secondary services in several ways. It can suggest how serials might be allocated among them, by indicating empirical subject boundaries (which may be different from those imposed by library classifications and existing secondary services). It can help to identify the most important serials in a subject field to be covered by a selective service. It can show what subject areas are likely to contribute most relevant items to a discipline covered by a secondary service. Language and country relationships are harder to interpret: weak links may merely show under-use of relevant material, and suggest that coverage of publications of other countries and in other languages should be improved.

As a large part of the research programme Design of Information Systems in the Social Sciences (DISISS), which aimed to collect information relevant to the improvement of secondary services in the social sciences, 59,000 citations in the social sciences were collected and analysed. Two reports on the analyses have been issued (Design of Information Systems in the Social Sciences, 1979, 1978): this paper summarizes some of the main results, and considers some of their implications for the planning of secondary services.

SOURCES OF REFERENCES/CITATIONS

A deliberate attempt was made to avoid deficiencies found in many or most citation studies. References analysed hitherto have nearly always been taken solely from serials, and these have generally consisted of a few highly rated titles. This is particularly serious in the social sciences, where monographs form a very large part of the literature. A previous study (Earle and Vickery, 1969) analysed references from both serials and monographs in the social sciences, but did not produce separate analyses for the two collections, so that no comparisons can be made. One or two analyses have been made of references from monographs only (e.g. Oromaner, 1968).

For DISISS, over 11,000 references were taken from nearly 300 British and American monographs. To avoid the bias that might be caused by using only a few source serials, 140 were used, including 47 high-ranking serials and 47 selected at random, yielding in all nearly 48,000 references.

The monographs used as sources were published exclusively in the U.K. or U.S.A. The reason for this was the strictly practical one that if a random sample of non-English language monographs had been drawn, so few of them would have

been available in British libraries that the resulting collection of references would have been very unrepresentative. 148 British monographs published in 1971 were sampled from the social science sections of the *British National Bibliography*, and 149 U.S. monographs published in the same year from the *American Book Publishing Record*. Every fifth reference was sampled from these. The U.K. monographs yielded 4,409 references, and the U.S. monographs 6,632–11,041 in all. It is interesting to note that the U.S. monographs were 50 per cent more productive of references than the British monographs; the average numbers of references per monograph were 223 and 149 respectively.

The bias caused by the limitation of sources to British and U.S. monographs is unknown, but it is less important than might have been expected for purposes of comparison with serial references, since nearly three quarters of the latter were drawn from British and U.S. serials. It is probable that by no means all of the monographs could be considered 'research' monographs. Some of them will have been textbooks and general reviews of a subject, and some will have been 'popular' works. Popular monographs cannot have introduced much distortion into the findings, since they will have contained relatively very few references. Textbooks and reviews of subjects are rather a different matter; they undoubtedly represent serious and academic use, but possibly of a different kind from research monographs.

One weakness in the monograph sources is that some subjects were represented by only one or two monographs. The number of references yielded by them was in some cases quite large, but this does not of course increase the validity of the analyses based on them. Ideally, a much larger number of monographs should have been used as sources, but the time and cost involved in data collection would have been prohibitive.

The 140 serials used as sources of references were chosen in several ways. 47 were selected as being highly cited in a pilot citation analysis. 23 more were chosen by specialists to improve subject coverage, and 18 to improve language coverage. A random sample of 47 was then taken from the Check List of Social Science Serials compiled specially for DISISS (Bradshaw *et al.*, 1974); 2 of these 47 were also among the highly cited titles. Finally, 7 titles in criminology were added for the special purpose of comparing citation patterns over time (a purpose for which the sample did not in fact prove adequate).

The volume year for most of the source journals was 1970; 101 (68 per cent) of the source journals were published in this year. 1970 volumes of the other 39 were not available, and the volumes used belong to years as near as possible to 1970, most of them to the previous year. In contrast to the monographs, all references from the 140 journals in the appropriate year were used.

The sample of serial sources, and hence of references, was unbalanced by an over-representation of some subjects (criminology and psychology in particular) and under-representation of others (especially political science, economics and social science [general]). Unlike the monograph sources, the serial sources excluded history altogether. The sample was also heavily biased towards British and U.S. serials (which represented 69 per cent of the titles), and towards English-language titles. These imbalances reduced the validity of some of the analyses, particularly those based on the whole sample of references.

As in the case of monographs, U.S. serial titles were much more productive of references than British titles; the average yields of references were 447 and 261 respectively, though this difference must be due partly to the fact that the U.S. serials contained more articles on average. References were not taken from forms

of publication other than serials and monographs, such as report literature, unpublished dissertations and theses, and official publications. It would be interesting to see how analyses of references from such sources compare with those from serials and monographs. It might be speculated that report literature references resemble serial references, that official publication references are not dissimilar to monograph references, and that references in theses are considerably more exhaustive than either.

Inevitably the coding of sources and cited items by subject, and to a lesser extent by form, is an uncertain and somewhat subjective procedure. Even if subject boundaries and forms of material could be defined, and the definitions were accepted, views would differ as to the placing of particular items within them. The procedures used are explained in DISISS Research reports B4 and B5 (Design of Information Systems in the Social Sciences, 1975, 1976).

The collection of references analysed, although very large in total and drawn from a much broader range of sources than previous collections, had, as has been noted, several imperfections, of varying degrees of seriousness, and the analyses must be read with these qualifications in mind. It should be pointed out, however, that these imperfections are likely to be greatly magnified in analyses based, as the vast majority are, on a much smaller number and range of citations.

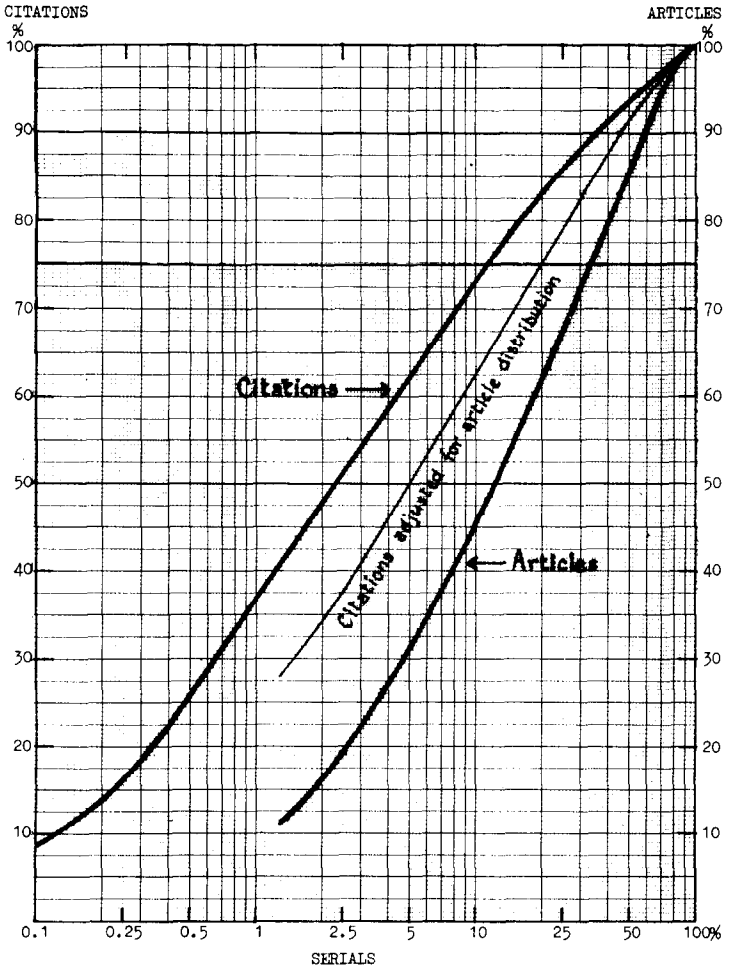
Numerous citation analyses were carried out on the references collected: concentration and scatter of titles cited, rank lists of titles cited, date distributions, subject relationships, and form, country of publication and language of cited items. Most of these analyses were made for both serial and monograph references, but the references from two sets of sources were not combined for any analyses, because there were sufficient differences in the two sets of sources to cast doubt on the validity of the results. (The cluster analyses that were carried out have been discussed elsewhere [Arms and Arms, 1978].)

CONCENTRATION AND SCATTER

The concentration of citations received from serials by serials was high (Figure 1). Some concentration of citations is due to the fact that there is a concentration of articles among serials as published; Wood and Ferguson (1974) found that 34 per cent of social science titles contained 75 per cent of articles published. The length of time a serial has been in existence also affects the number of citations it receives; this was allowed for in the present study by analysing separately citations received by 1965–1970 serials.

Twelve per cent of all serials cited by serials received 75 per cent of all citations, and 6 per cent received 65 per cent. (For comparison, 6 per cent of serials cited by monographs received only 24 per cent of citations by monographs.) Differences between subjects (Figure 2) are very large: 90 per cent of citations were received by 30 per cent of the serials cited by psychology, while in political science 90 per cent were received by 80 per cent. These differences are much too large to be accounted for by differences between the size of serials in the subjects. Economics, linguistics and education also showed high concentration; while environmental planning had a very wide scatter. Citations received by 1965–1970 volumes showed, as expected, less concentration than all citations. The shapes of the distribution curves for different subjects differed widely. There was some similarity between the rank order of subjects according to concentration and

FIG. 1. Distribution of citations among serials compared with distribution of articles among serials.

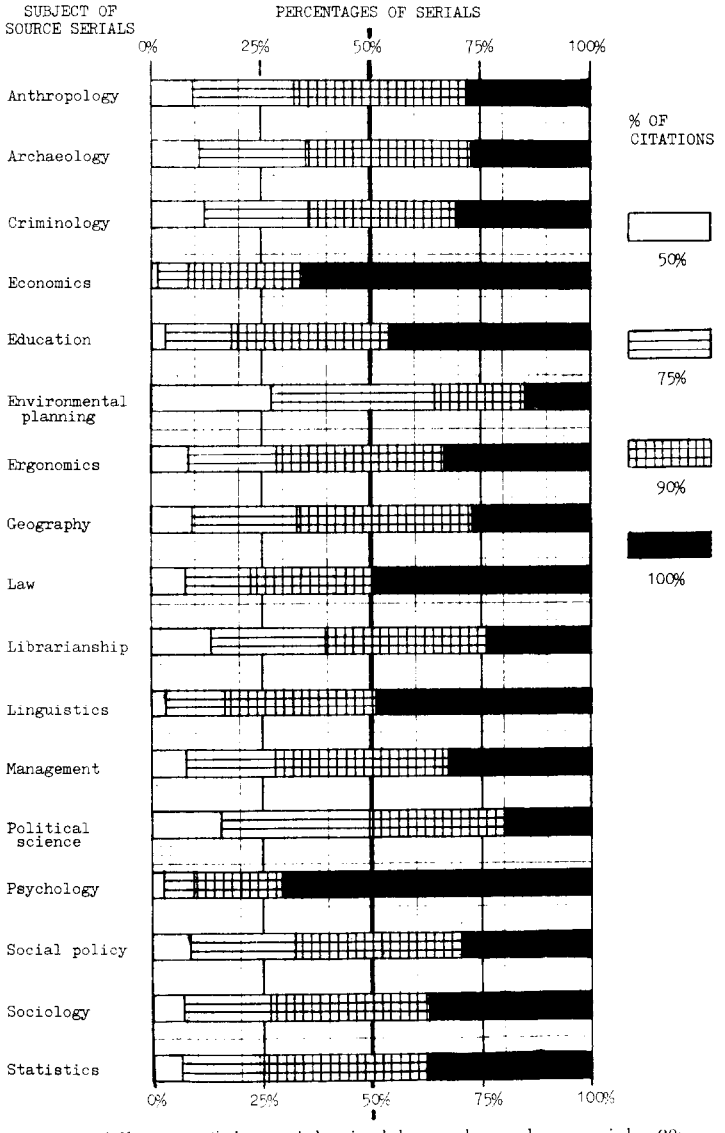


To be read as follows: 12% of serials cited by all social science serials received 75% of citations of social science articles published. If all social science serials contained an equal number of articles, about 20% of serials would receive 75% of citations.

according to the proportion of references they made to themselves—not surprisingly, since most of the cited serials would be expected to be within the subject.

For the planners of secondary services, the degree of concentration is of less practical interest than the *number* of serials required to give a high degree of coverage. For example, although sociology showed a higher concentration than political science, only 33 per cent of titles accounting for 75 per cent of citations 1965–1970 compared with 53 per cent, the number of titles involved was 148 compared with only 107. Also, a subject can have quite a high concentration and a very long 'tail' of little-cited serials. Subjects with long tails of little-cited serials were sociology and psychology; although the latter has *relatively* the shortest tail

FIG. 2. Percentages of serials in different subjects accounting for 50%, 75% and 90% of references made.



To be read as follows: Of the serials cited by anthropology serials, 9% received 50% of the citations made, 34% received 75% and 72% received 90%.

of all, the total number of serials cited is so large that a relatively very short tail is a very long one in absolute terms (Table 1).

Obviously, good coverage of relevant material in serials can be achieved most easily in subjects where there is a very high concentration on a limited number of titles or a short 'tail' of cited serials. High concentration favours services that do not aim at higher than 75 per cent or 80 per cent coverage, while short tails tend to favour those that aim at full coverage. High concentration and short tails can coincide only in subjects in which the total number of cited serials is small, and these subjects clearly have the least problems of all.

TABLE 1. Percentages of serials rarely cited, by subject

Subject of source serials	Serials cited				Total no. of serials cited
	1 or 2 times		3–5 times		
	% of citations	(No. of serials)	% of citations	(No. of serials)	
Anthropology	34.7	(339)	18.0	(71)	537
Criminology	28.8	(163)	22.4	(43)	232
Economics	11.1	(454)	15.5	(101)	650
Education	19.9	(228)	13.6	(42)	308
Geography	34.3	(468)	10.7	(63)	590
Linguistics	17.1	(192)	11.8	(31)	263
Political science	47.4	(261)	12.7	(41)	324
Psychology	8.1	(559)	14.2	(123)	867
Social policy	30.1	(248)	17.1	(57)	333
Sociology	22.8	(537)	16.6	(126)	760
Statistics	24.0	(222)	13.3	(40)	301

TABLE 2. Serials in rank order of citations received from economics serial sources

Rank order			Serial
All years (no. of citations in brackets)		1965–1970 (incl. self-citations)	
Incl. self-citations	Excl. self-citations		
1 (543)	1 (350)	1	*Amer. Economic Rev.* 1911–
2 (387)	2 (297)	2	*J. Political Economy* 1892–
3 (327)	3 (274)	7	*Econometrica* 1933–
4 (262)	4 (209)	4	*Q. J. Economics* 1887–
(262)	5 (198)	5	*Rev. Economic Studies* 1933–
6 (250)	6 (182)	3	*Rev. Economics & Statistics* 1919–
7 (202)	7 (169)	9	*Economic J.* 1891–
8 (132)	8 (100)	10	*Economica* 1921–
9 (118)	30 (16)	6	*Monthly Labor Rev.* 1915–
10 (115)	—	8	*Wirtschaft und Statistik* 1949–
11 (73)	13 (38)	11	*J. Business* 1928–
12 (68)	9	13	*Int. Economic Rev.* 1961–
13 (63)	10	12	*J. Finance* 1946–
(63)	10	20	*J. Amer. Statistical Assoc.* 1888–
(63)	10	19	*Oxford Economic Papers* 1938–
16 (39)	—	14	*Agrarwirtschaft* 1952–
17 (38)	—	15	*J. Marketing Research* 1964–
(38)	17 (27)	15	*Southern Economic J.* 1933–
19 (36)	23 (22)	17	*Voprosy Ekonomiki* 1948–
20 (35)	14		*Management Science* 1954–
21 (33)	15	31	*Amer. J. Agricultural Economics* 1919–
22 (30)	—	18	**Canadian J. Economics* 1968–
23 (28)	16	30	*Economic Development & Cultural Change* 1952–
24 (27)	17	22	*Kyklos* 1948–
25 (25)	19	62	*J. Royal Statistical Society Series A (General)* 1838–

In column 2, where no figure appears in brackets, the number of citations received is as in the first column, because the serial in question was not a source serial.

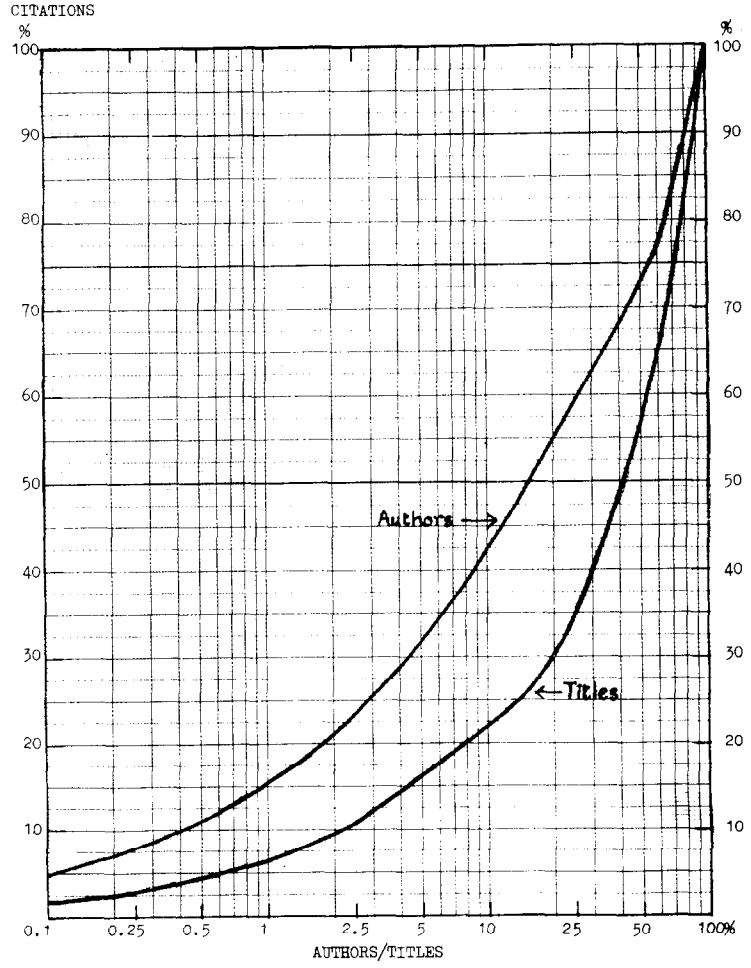
* *Canadian J. Economics and Political Science* split from 1968 into *Canadian J. Economics* and *Canadian J. Political Science*.

Citation analyses can reveal little of direct practical value about other forms of cited material than serials, since the scatter of references is too wide, and the relatively few titles identified as cited highly would both be already well known and have been published some time (often a long time) before. In the present case, 80 per cent of the monographs cited by serials were cited only once, and only 27 titles received more than 10 citations; most of these are classics and methodological works. Twenty-five per cent of monograph *authors* received 60 per cent of citations by serials (Figure 3).

RANK LISTS

The general rank lists of serials cited are seriously affected by the subject imbalance of the sources, and in some subjects the rank lists are of relatively little

FIG. 3. Distribution of citations from serials among monograph titles and authors.



To be read as follows: 4% of authors received 29% of citations from serials; 8% of titles received 20% of citations from serials.

value because the sample of citations was small. However, the majority of the subject lists are more reliable than those produced by most previous citation

TABLE 3. Serials in rank order of citations received from psychology serial sources

Rank order			Serial
All years (no. of citations in brackets)		1965–1970 (incl. self-citations)	
Incl. self-citations	Excl. self-citations		
1 (782)	1 (577)	1	*J. Experimental Psychology* 1916–
2 (468)	4 (239)	4	*J. Comparative & Physiological Psychology* 1921
3 (333)	2 (298)	6	*Psychological Rev.* 1894–
4 (324)	10 (115)	2	*J. Verbal Learning & Verbal Behaviour* 1962–
5 (275)	7 (162)	3	*Psychonomic Science* 1964–
6 (263)	3	—	**J. Abnormal & Social Psychology* 1921–64
7 (231)	5 (213)	6	*Psychological Bull.* 1904–
8 (189)	6	8	*Science* 1880–
9 (170)	14 (84)	9	*Brit. J. Psychiatry* 1858–
10 (161)	8	5	**J. Personality & Social Psychology* 1965–
11 (139)	9	10	*Psychological Reports* 1955–
12 (130)	13 (86)	12	*J. Consulting & Clinical Psychology* 1937–
13 (114)	11	15	*Archives of General Psychiatry* 1919–
14 (111)	21 (69)	11	*Child Development* 1930–
15 (99)	22 (68)	25	*J. Psychology* 1936–
16 (94)	17 (80)	18	*Amer. Psychologist* 1946–
17 (93)	16 (83)	38	*Amer. J. Psychology* 1887–
18 (92)	12	17	*Amer. J. Psychiatry* 1844–
19 (91)	18 (76)	21	*J. Personality* 1932–
20 (89)	—	23	*Int. J. Psychoanalysis* 1920–
21 (84)	14	21	*J. Experimental Analysis of Behavior* 1958–
22 (76)	18	31	*J. Nervous & Mental Disease* 1874–
23 (72)	20	42	*Brit. J. Psychology* 1904–
24 (63)	28 (52)	24	*Canadian J. Psychology* 1947–
25 (61)	37 (41)	42	*J. Clinical Psychology* 1945–
26 (60)	23	28	*Perceptual & Motor Skills* 1949–
27 (59)	35 (42)	12	**J. Abnormal Psychology* 1965–
(59)	24	14	*Perception & Psychophysics* 1966–
29 (58)	25	31	*J. Applied Psychology* 1917–
30 (56)	26	94	*Psychoanalytic Study of the Child* 1945–
(56)	26	28	*Q. J. Experimental Psychology* 1948–
32 (51)	37 (41)	42	*J. Social Psychology* 1929–
(51)	—	58	*Occupational Psychology* 1932–
(51)	29	16	*Physiology & Behavior* 1966–
33 (49)	30	69	*J. Genetic Psychology* 1935–
36 (47)	31	35	*J. Educational Psychology* 1910–
(47)	31	74	*J. Amer. Psychoanalytic Assoc.* 1953–
38 (46)	33	33	*Nature* 1869–
39 (43)	34	20	*Behaviour Research & Therapy* 1963–
(43)	—	25	*J. Psychosomatic Research* 1956–
41 (42)	35	45	*Psychometrika* 1936–
42 (41)	37	37	*J. Experimental Child Psychology* 1964–
(41)	37	36	*Psychosomatic Medicine* 1938–
44 (40)	50 (30)	55	*Amer. J. Orthopsychiatry* 1930–

**J. Abnormal & Social Psychology* split from 1965 into *J. Abnormal Psychology* and *J. Personality & Social Psychology*.

In column 2, where no figure appears in brackets, the number of citations received is as in the first column, because the serial in question was not a source serial.

analyses because they are based on many more, and more carefully selected, source serials. Subject rank lists for economics and psychology, the two subjects with much the largest number of citations, are given in Tables 2 and 3. (Lists for all subjects are given in DISISS Research report A3.)

There are quite large differences in rank order between the lists based on all citations and those based on 1965–1970 citations only; the latter are more valid as guides to selection for coverage by current secondary services. Self-citations had a large effect on the rank order of some source serials, a few of which were cited only by themselves. The large differences between the rank lists produced from references in high-cited serials and those in the serials selected at random strongly suggest that rank lists based (as nearly all are) solely on the former are unreliable.

A list of monograph authors and editors most cited by serials is given in Table 4.

TABLE 4. Monograph authors (or editors) cited by serials 25 or more times

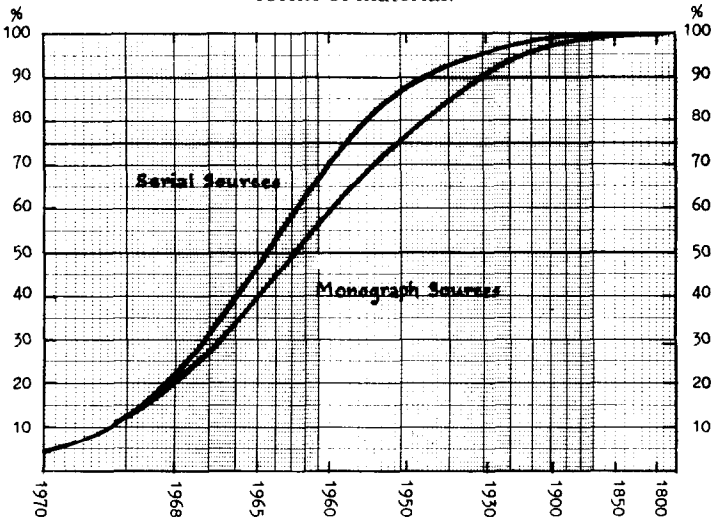
No. of times cited		No. of times cited	
265	Lenin, V. I.	37	Mill, J. S.
156	Marx, K.	36	Edwards, A. L.
142	Engels, F.	35	Duncan, O. D.
127	Freud, S.	34	Samuelson, P. A.
90	Parsons, T.	33	Durkheim, E.
67	Friedman, M.	33	Thorndike, E. L.
66	Pavlov, T.	32	Spence, K. W.
63	Ricardo, D.	31	Blalock, H. M. jnr
55	Merton, R. K.	31	Johnston, J.
53	Blau, P. M.	30	Goffman, E.
52	Piaget, J.	29	Bruner, J. S.
49	Lipset, S. M.	27	Homans, G. C.
47	Weber, M.	26	Festinger, L.
45	Eysenck, H. J.	26	Horton, D. L.
45	Siegel, S.	26	Simon, H. A.
45	Winer, B. J.	25	Arrow, K. J.
44	Lazarsfeld, P. F.	25	Atkinson, J. W.
42	Hicks, J. R.	25	Dixon, T. R.
41	Coleman, J. S.	25	Etzioni, A.
40	Chomsky, N.	25	Musgrave, R. A.
38	Johnson, H. G.	25	Popper, K. R.

The fact that the rank lists are now some years out of date is unlikely to affect the relative standing of the most highly cited titles. Those lower down the order may have gone either up or down in standing and new titles will have appeared. Regular rank lists should be produced every three or four years, to provide an up-to-date picture. These could most easily be produced from the SSCI data base, even though the references in this are taken mainly from fairly high ranking serials. The present general rank lists prepared by ISI are of little value to planners of secondary services in particular subjects, who need rank lists by subject.

DATE DISTRIBUTIONS

Date distributions of citations need very careful interpretation (Line and Sandison, 1974). Crude correction factors were applied to indicate the kind of

FIG. 4. Dates of references by serials and monographs to all subjects and forms of material.



To be read as follows: 53% of all references by serials, and 45% of all references by monographs, were to items published 1964–71.

TABLE 5. Annual serial citation decay factors, by subject

Subject of source serial	1966/69[1]–1900/29[2]		1966/69[1]–1950/53[3]		
	Uncorrected	Corrected for title growth[4]	Uncorrected	Corrected for title growth[4]	Corrected for title and article growth[5]
Anthropology	.928	.956	.877	.911	
Economics	.901	.928	.898	.919	.949
Education	.926	.963	.893	.937	
Geography	.925	.951	.880	.890	
Linguistics[6]	.927	.966	.907	.930	
Political science[6]	.899	.926	.846	.871	
Psychology	.905	.938	.880	.910	.936
Social policy[6]	.927	.950	.852	.865	
Social sciences (general)[6]	.920	.949	.889	.919	
Sociology	.909	.945	.900	.936	.928
All cited serials	.911	.941	.873	.902	

[1] Based on average of citations received (by serials only) during the years 1966–1969 (average used to avoid distortion due to high or low figures in particular years).

[2] Based on average of citations received by serials during the years 1900–1929 (see note 1).

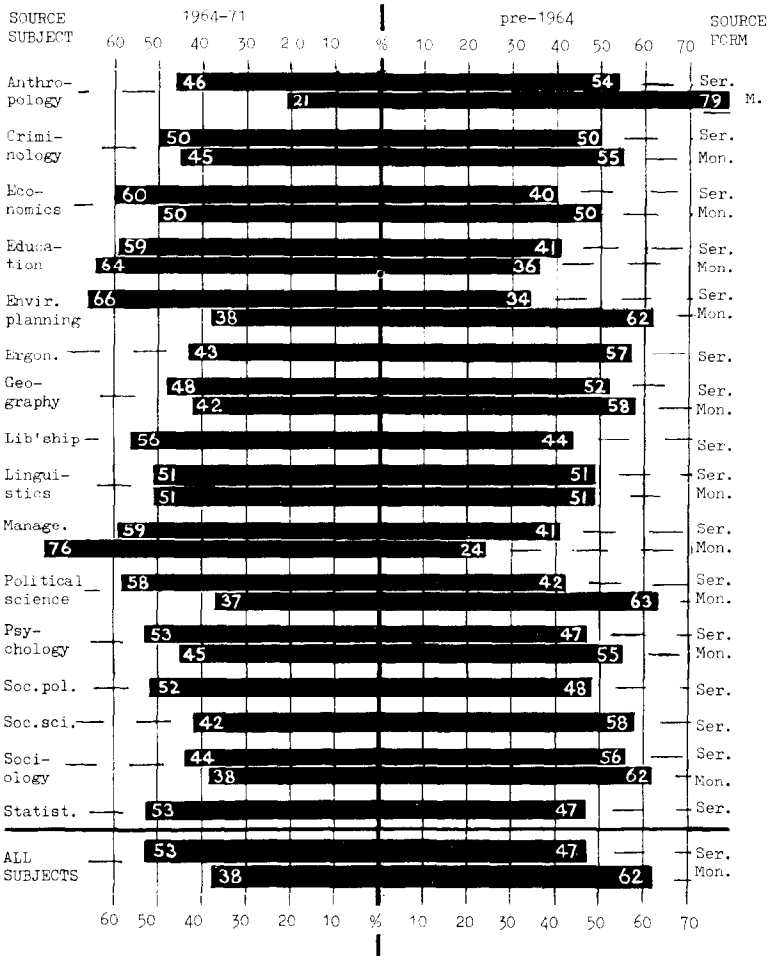
[3] Based on average of citations received by serials during the years 1950–1953 (see note 1).

[4] The correction factors used (DISISS Report A3, Table 17) relate to growth rates in various subjects; the citation decay factors, on the other hand, are calculated for the *serials cited by particular subjects*, which of course extend well beyond the subjects themselves.

[5] Correction factors as follows: Economics 1.0327 (based on 20 journals 1950–1968); Psychology 1.0282 (based on 8 journals 1950–1968); Sociology 0.991 (based on 5 journals 1950–1968). The serials in question are relatively significant ones, and their growth rates may not be typical. It may be noted that the number of articles in the 5 sociology serials actually declined between 1950 and 1968—but the number of *pages* increased by an average annual factor of 1.0077.

[6] Number of citations relatively small, especially in earlier years; results therefore less reliable.

FIG. 5. Proportion of references in different subjects, made by serials and monographs, pre-1964 and 1964–71.



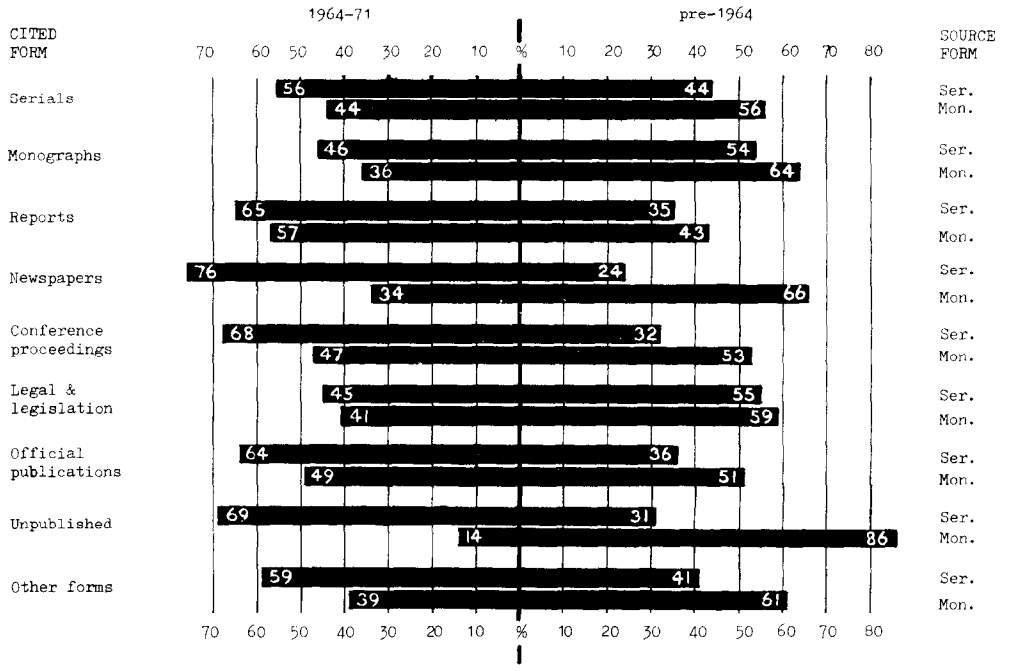
To be read as follows: 46% of all references by anthropology serials, and 21% of all references by anthropology monographs, were to items published 1964–71.

effect literature growth is likely to have had on citation decay. A higher percentage of citations from highly cited serials were to recent years than of citations from serials shown at random. References by serials showed a much faster decay than references by monographs in nearly all subjects and for all forms of material cited (Line, 1979) (Figures 4–6). Citations from serials to serials only decayed more rapidly than citations from serials to all forms of material, in all subjects but one.

Decay factors varied considerably according to the period and length of time over which they were calculated; in some subjects citations decayed rapidly in one period and slowly in another. However, economics and political science had fast decay rates over all periods calculated. Few subjects referred much to material published before 1940.

Differences between the date distributions for different forms of material cited were very large; for example, the period 1964–1971 accounted for 76 per cent of

FIG. 6. Proportion of references by serials and monographs to different forms of material, pre-1964 and 1964–71.



To be read as follows: 56% of references by serials to serials, and 44% of references by monographs to serials, were to items published 1964–71.

TABLE 6. Annual citation decay factors, by form

	1966/69[1]–1900/29[2]		1966/69[1]–1950/53[3]	
	Uncorrected	Corrected for growth[4]	Uncorrected	Corrected for growth
Serials	.911	.941	.873	.902[4]
Monographs	.927	—	.898	.950[5]
Reports	.917	—	.835	—
Newspapers	.925	—	.777	—
Conference proceedings	.891	—	.838	—
Legal and legislation	.930	—	.877	—
Official publications	.918	—	.846	—
Unpublished	.921	—	.833	—
All forms	.919	—	.879	—

[1–4] See notes 1–4 to Table 5.

[5] Correction factor 1.0578, calculated from statistics of book production 1960–1970 in UDC class 3 of the 45 top book producing countries. The books produced by these countries seem rather more likely to be representative of the population of cited books than the books produced by all countries (growth factor 1.0524) or those produced by the top 12 producing countries (growth factor 1.0708).

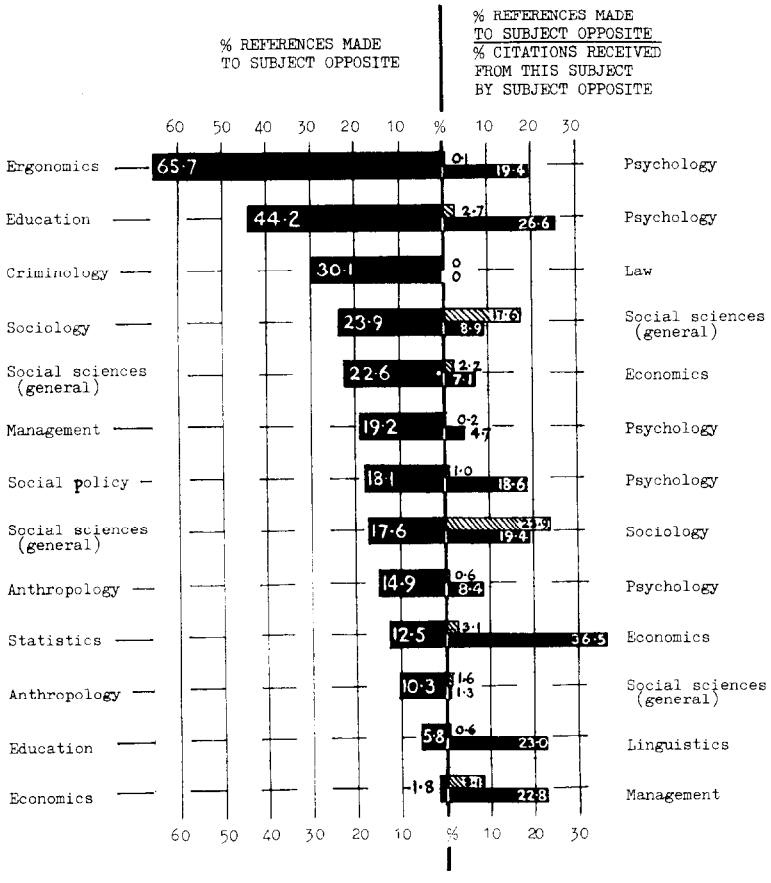
serial references to newspapers at one extreme, and for 45 per cent of references to legal and legislative material at the other.

Crude correction factors were applied to the citation decay factors, to allow for the different size of the citable serial literature at different times. These, shown in

Tables 5 and 6, are intended as indications of the sort of effect that changes (mainly growth) in the size of the literature can have on so-called 'obsolescence', the rate of which has often been grossly exaggerated.

Date distributions of citations have little relevance to current published secondary services, though they may be of some interest in the preparation of retrospective bibliographies. They do have considerable interest for the maintenance of computer-based retrieval services, however, which can rarely afford to keep more than a limited amount of material available on-line and therefore need some programme of retirement of less used references. Ideally, retirement should be carried out on the basis of actual use (or non-use), title by title; this should be possible if use is automatically monitored. Data such as those provided by the present study can serve as guides to the proportion and type of material that may need to be retained on current files.

FIG. 7. Links between selected subjects as shown by serial references and citations.



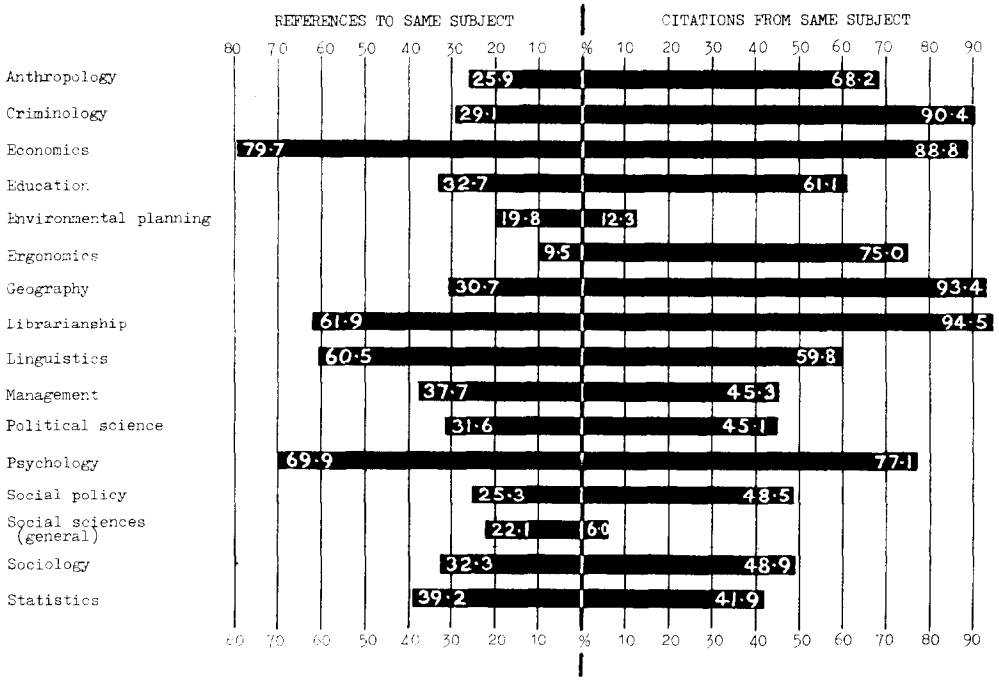
To be read as follows: 65.7% of references to serials made by ergonomics serials were to psychology serials; 0.1% of references to serials made by psychology serials were to ergonomics serials; 19.4% of serials citations received by ergonomics serials were from psychology serials.

SUBJECT RELATIONSHIPS

Subject relationships are shown both by references made (indicating dependence of a subject on other subjects) and by citations received (indicating the extent to which the subject is drawn on by other subjects). There was a heavy general dependence on psychology, and to a lesser extent on economics. Dependences by one subject on another were rarely reciprocated (Figure 7).

Subjects differed greatly in their self-dependence (Figure 8). Economics and psychology made a very high percentage of their references to themselves and a high percentage of the citations they received were from themselves. Criminology, geography, social policy, sociology, political science and education made relatively few references to themselves. References from monographs showed in most subjects a much higher self-dependence than references from serials, but in economics and psychology there were large differences in the opposite direction (Line, 1979).

FIG. 8. Subject 'inbreeding' as indicated by serial references made and citations received.

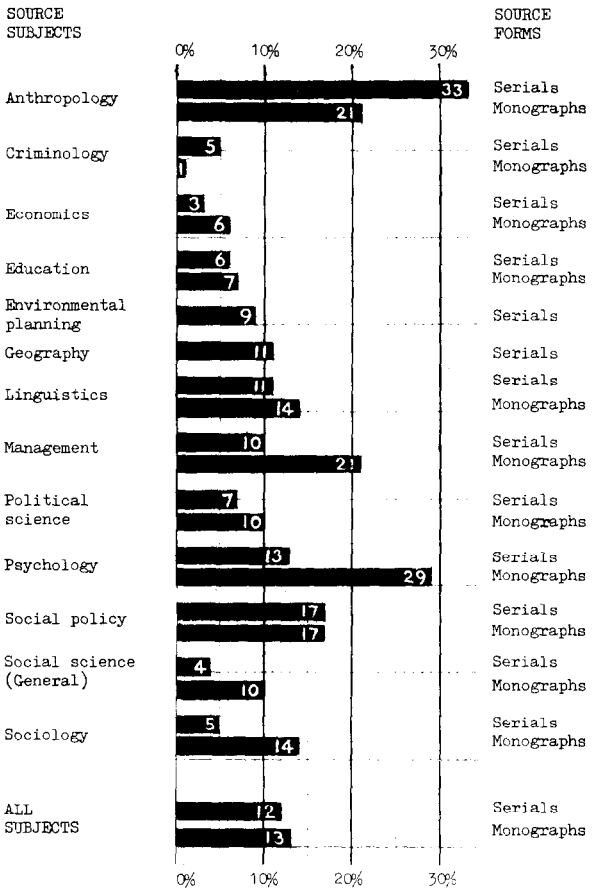


To be read as follows: 25.9% of references to serials made by anthropology serials were to anthropology serials; 68.2% of citations from serials received by anthropology serials were from anthropology serials.

High self-sufficiency can indicate 'inbreeding'—an unwillingness to seek enrichment from other subjects—or maturity and coherence. There would be little disagreement that psychology and economics are the social sciences with the clearest identities. Sociology and political science are of their nature less clear cut, while the dependence of geography, criminology and education on other subjects is understandable.

Whatever the conceptual implications of self-sufficiency, it obviously aids the planning of secondary services, since it is much easier to draw up a list of serials in a subject than of serials on which a subject draws. Serious problems are faced by subjects like political science that not only draw heavily on themselves but draw on a large number of other disciplines; and perhaps the worst problems of all face those subjects that rely a good deal on subjects beyond the social sciences, such as anthropology, geography and social policy (Figure 9).

FIG. 9. References by serials and monographs beyond the social sciences.

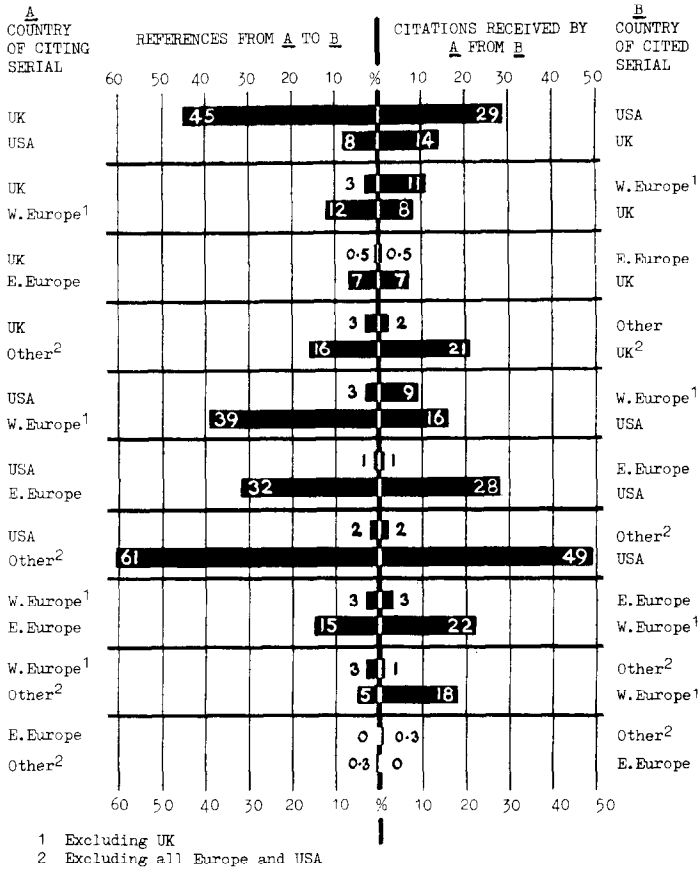


To be read as follows: 33% of references by anthropology serials, and 21% of references by anthropology monographs, were to subjects other than the social sciences.

COUNTRY AND LANGUAGE ANALYSES

Citation links between different areas of the world are shown in Figure 10. There was a strong tendency for most countries to refer to and be cited by their own publications, though U.S. (and to a lesser extent U.K.) serials accounted for a

FIG. 10. Links between areas of the world as shown by serial references and citations.

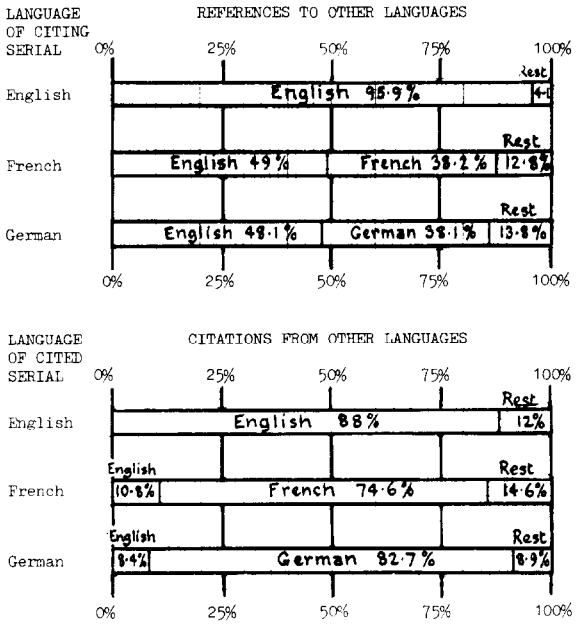


To be read as follows: 45% of references to serials by U.K. serials were to U.S.A. serials, and 29% of citations from serials received by U.K. serials were from U.S.A. serials.

higher percentage of citations than would be expected, except in the case of Slavonic serials. This characteristic may have an element of parochialism, but most social sciences are inevitably more concerned with local conditions and circumstances than pure and applied science, which is much less affected by national boundaries. That said, it does appear that social scientists look less beyond their frontiers than they might. Whether this should be accepted as a fact—a welcome fact—by secondary services, or whether they should aim deliberately to 'improve' the situation, is a question of much interest and importance.

The language analyses (Figure 11) similarly showed that writers in each language tend to refer to works in their own language, but serials in nearly all languages referred even more to writings in English. Similar considerations apply as were mentioned with regard to the country analyses. The relatively low use of material in other languages may be in large part due to the poor coverage of such

FIG. 11. Links between selected languages as shown by serial references and citations.



To be read as follows: 95.9% of references to serials by English language serials were to serials in English, 4.1% to serials in other languages.

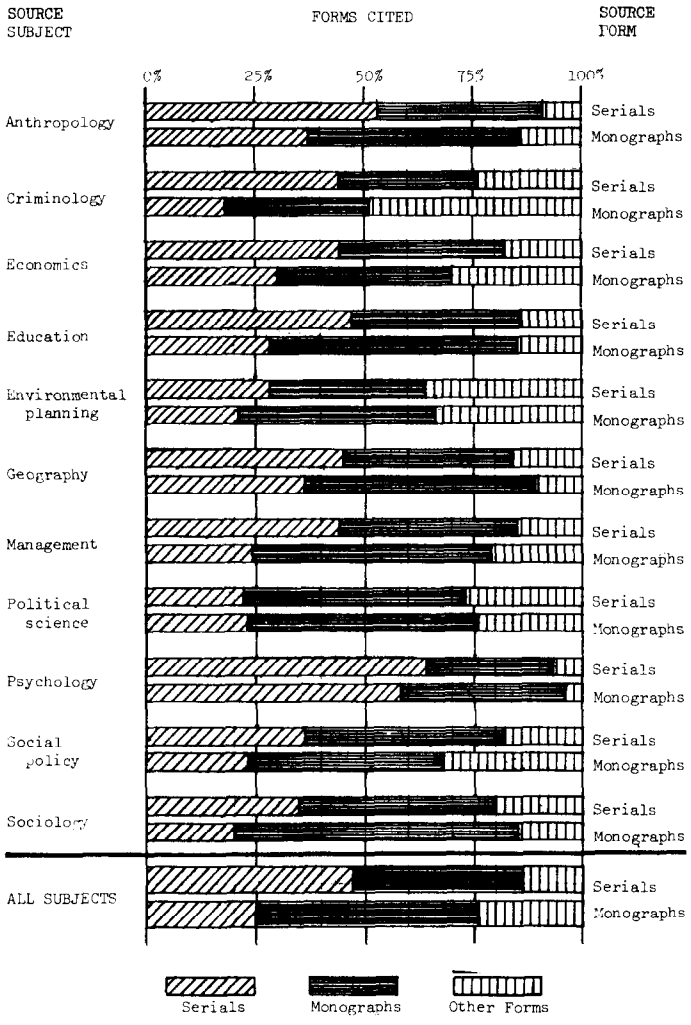
material by secondary services. On the other hand, the linguistic ability of social scientists is known to be poor (Line, 1971; Hutchins and others, 1971), and without much better translating services than at present exist—including maybe serials in particular subjects consisting of selected articles translated from foreign languages—deliberate extension of coverage of foreign language material by secondary services could be wasteful of resources that might be used to improve them in other, and perhaps more necessary, ways.

FORMS OF MATERIAL CITED

There were very large differences in the forms of material cited by serials and monographs (Line, 1979). Over all subjects serials made nearly twice as many references to serials as did monographs (47 per cent and 25 per cent respectively) and considerably fewer references to monographs (39 per cent and 51 per cent respectively).

There were also great differences between subjects in the forms cited by them (Figure 12). As would be expected, the highest proportion of references to serials was made by psychology serials (64 per cent); the lowest was by political science serials (22 per cent). The highest proportion of references to monographs by serials was in sociology (56 per cent). Miscellaneous forms of material (reports, official publications, newspapers, etc.) accounted for high proportions in environmental planning, criminology, social policy, political science and economics.

FIG. 12. References by serials and monographs to different forms of material.



To be read as follows: 53% of references by anthropology serials were to serials, 38% to monographs, and 9% to other forms.

Secondary services devoted exclusively to serials, as many are, would, even if their coverage of serials was comprehensive, cover more than half of the relevant literature only in psychology. In political science they would cover less than a quarter. Problems of coverage would be particularly severe in subjects in which 'miscellaneous' forms of material are heavily used. It could of course be argued that the high proportion of references made to non-serials shows that their coverage by secondary services is unnecessary, since these other forms of material must be readily traced in order to be used and cited. In some fields, certainly, researchers on certain topics do not need to be led to such forms as newspapers or official publications, which may even constitute their basic working sources.

However, this does not prove that improved coverage of different forms of material by secondary services would not make access much easier and draw attention to relevant and important material that would otherwise be missed.

IMPLICATIONS FOR THE COLLECTION OF REFERENCES FOR ANALYSIS

Differences between the patterns of references made by monographs and by serials, and by high ranking serials and serials chosen at random, have been mentioned several times. Those mentioned by no means indicate all the differences that emerged from the analyses. Different forms of publication are undoubtedly used for different purposes, which the references they make will reflect. Consequently, references taken from particular forms may well be appropriate for particular purposes—for example, high-ranking serials may well constitute the most satisfactory sources for studying uses made at the frontiers of research.

If a citation analysis has a specific purpose this should be clearly stated, and the references most appropriate to the purpose collected. Unfortunately the purpose is often not stated, although the results are frequently used to draw conclusions that are highly questionable. For example, a collection of references from a few high ranking serials may show a fairly rapid date decay, and this is held to prove that libraries can readily dispose of all serial volumes prior to a certain date. The objections to such a conclusion are first that the sources may not be numerous or varied enough to represent use adequately (even if a rough general equivalence between citations and uses is assumed), secondly that different serials obsolesce at very different rates, and thirdly that if the object is to save shelf space, use must be related not to volumes but to shelf space occupied—earlier volumes may be much thinner than later ones. (It may be that in science references made by high ranking serials are much more representative of all references than in the social sciences: tests should be carried out to establish this.)

GENERAL COMMENT

Against the background of the citation analyses reported here, most secondary services in the social sciences appear to be seriously deficient in the range of forms of material they cover, in the number and possibly the selection of serials they cover, in the subject spread of material of possible relevance, and probably also in coverage of foreign language material and of material published in many other countries. However, no service can attempt, let alone attain, full coverage in all these aspects, and if it did it would be both very costly and probably unusable. The best that an individual service of a conventional nature can attempt is to optimize its coverage, bearing in mind the market and the usability of the end product.

REFERENCES

ARMS, W. Y. and ARMS, C. R. (1978). Cluster analysis used on social science journal citations. *Journal of Documentation*, *34*, 1–11.

BRADSHAW, R. G. and others (1974). CLOSSS (Check List of Social Science Serials): a machine readable data base of social science serials. *Program*, *8*, 29–39.

DESIGN OF INFORMATION SYSTEMS IN THE SOCIAL SCIENCES (1975). *Research report B4. Characteristics of social science serials.* London: Polytechnic of North London *and* Bath: University Library. (BL R&D Report No 5269.)

DESIGN OF INFORMATION SYSTEMS IN THE SOCIAL SCIENCES (1976). *Research report B5. The collection and construction of a file of citation data.* London: Polytechnic of North London *and* Bath: University Library. (BL R&D Report No 5334.)

DESIGN OF INFORMATION SYSTEMS IN THE SOCIAL SCIENCES (1978). *Research report B7. Literature usage and interrelationships in the social sciences as shown by citations in monographs.* London: Polytechnic of North London *and* Bath: University Library. (BL R&D Report No. 5478.)

DESIGN OF INFORMATION SYSTEMS IN THE SOCIAL SCIENCES (1979). *Research report A3. The structure of social science literature as shown by citations.* Bath: University Library. (BL R&D Report No. 5524.)

EARLE, P. and VICKERY, B. (1969). Social science literature use in the UK as indicated by citations. *Journal of Documentation, 25*, 123–142.

HUTCHINS, W. J., PARGETER, L. J. AND SAUNDERS, W. L. (1971). *The language barrier: a study in depth of the place of foreign language materials in the research activity of an academic community.* Sheffield: University, Postgraduate School of Librarianship and Information Science.

LINE, M. B. (1971). The information uses and needs of social scientists: an overview of INFROSS. *Aslib Proceedings, 23*, 412–434.

LINE, M. B. (1977). On the irrelevance of citation analyses to practical librarianship, *In: EURIM II: a European conference on the application of research in information services and libraries, 23–25 March 1976*, pp. 51–56. London: Aslib.

LINE M. B. (1979). The influence of the type of sources used on the results of citation analyses. *Journal of Documentation, 35*, 265–284.

LINE, M. B. AND SANDISON, A. (1974). 'Obsolescence' and changes in the use of literature with time. *Journal of Documentation, 30*, 283–350.

LINE, M. B. and SANDISON, A. (1975). Practical interpretation of citation and library use studies. *College and Research Libraries, 36*, 393–396.

OROMANER, M. J. (1968). The most cited sociologists: an analysis of introductory text citations. *American Sociologist, 3*, 124–126.

WOOD, D. N. and FERGUSON, J. (1974). Statistics of social science periodicals. *BLL Review, 2*, 92–95.

(Received 26 February 1980)