Essays of an Information Scientist, Vol1, p.527-544, 1962-73 Reprinted from :Science, (178):471-479, 1972

Citation Analysis as a Tool in Journal Evaluation

Journals can be ranked by frequency and impact of citations for science policy studies.

Also see: Citation frequency and citation impact -- and the role they play in journal selection for Current Contents and other ISI services.

Eugene Garfield

(*NOTE: The article reprinted here was referenced in the essay which begins on page 409 in Volume 1. Its inadvertent omission was discovered too late to include it at its proper location, immediately following the essay.*)

As a communications system, the network of journals that play a paramount role in the exchange of scientific and technical information is little understood. Periodically since 1927, when Gross and Gross published their study (*1*) of references in 1 year's issues of the *Journal of the American Chemical Society*, pieces of the network have been illuminated by the work of Bradford (*2*), Allen (*3*), Gross and Woodford (*4*), Hooker (*5*), Henkle (*6*), Fussler (*7*), Brown (*8*), and others (*9*). Nevertheless, there is still no map of the journal network as a whole. To date, studies of the network and of the interrelation of its components have been limited in the number of journals, the areas of scientific study, and the periods of time their authors were able to consider. Such shortcomings have not been due to any lack of purpose, insight, or energy on the part of investigators, but to the practical difficulty of compiling and manipulating manually the enormous amount of necessary data.

A solution to this problem of data is available in the data base used to produce the *Science Citation Index* (SCI) (*10*). The coverage of the SCI is international and multidisciplinary; it has grown from 600 journals in 1964 to 2400 journals in 1972, and now includes the world's most important scientific and technical journals in most disciplines. The SCI is published quarterly and is cumulated annually and quinquennially, but the data base from which the volumes are compiled is maintained on magnetic tape and is updated weekly. At the end of 1971, this data base contained more than 27 million references to about 10 million different published items. These references appeared over the past decade in the footnotes and bibliographies of more than 2 million journal articles, communications, letters, and so on. The data base is, thus, not only multidisciplinary, it covers a substantial period of time and, being in machine-readable form, is amenable to extensive manipulation by computer.

In 1971, the Institute for Scientific Information (ISI) decided to undertake a systematic analysis of journal citation patterns across the whole of science and technology. It began by extracting from the data base all references published during the last quarter of 1969 in the 2200 journals then covered by the SCI. The resultant sample was about 1 million citations of journals, books, reports, theses, and so forth. To test whether this 3-month sample was representative of the year as a whole, it was matched against another sample made by selecting every 27th reference from the approximately 4 million references collected over the entire year. The two samples were similar enough in scope (number of different items cited) and detail (relative frequency of their citation by different journals) to

ITEM NO.	CITED JOURNAL	TOTAL	NUMBER OF TIMES CITED										
			1969	1968	1967	1966	1965	1964	1963	1962	1961	1960	REST
00243	ACTA PATH JAP	36	1	3	3	4	6	7	3	·	·	3	6
00244	ACTA PATH MICROBIOL	736	29	69	87	59	56	59	44	48	20	31	234
00909	AM J ANAT	637	7	27	37	56	32	41	15	26	15	21	360
00910	AM J BOT	1171	13	74	87	68	73	66	57	57	47	49	580
00911	AM J CANCER	103	·	·	·	·	·	·	·	·	·	·	103
00912	AM J CARDIOL	1238	73	201	199	247	134	70	78	66	53	73	44
03591	CAN J SOIL SCI	33	·	2	1	4	1	2	6	6	3	2	6
03592	CAN J SURG	61	2	6	4	3	13	11	3	3	5	1	10
03593	CAN J TECH	3	·	·	·	·	·	·	·	·	·	·	3
03594	CAN J ZOOL	356	46	38	40	28	24	20	19	29	17	22	73
08990	ISRAEL J AGR RES	29	1	1	7	·	1	8	7	2	1	·	1
08991	ISRAEL J BOT	16	·	·	3	5	4	3	1	·	·	·	·
08992	ISRAEL J CHEM	91	14	25	18	10	11	6	7	·	·	·	·
09651	J INVEST DERM	695	24	78	81	69	65	46	30	31	34	22	215
09652	J IOWA MED SOC	13	·	·	·	5	·	·	1	2	·	·	5
09653	J IRISH MED ASS	16	1	3	4	3	3	·	·	·	·	·	2
13390	P CALIF ACAD SCI	18	·	·	1	4	3	·	1	·	·	·	9
13391	P CAMBRIDGE PHIL SOC	389	8	22	23	11	12	9	13	11	17	3	260
19755	Z ANGEW CHEM	47	·	·	1	·	·	1	·	·	1	1	43
19756	Z ANGEW ENT	35	·	·	1	1	4	2	4	1	1	5	16
19757	Z ANGEW GEOL	49	2	7	5	8	5	5	4	4	2	1	6
19758	Z ANGEW MATH	10	1	·	1	·	1	·	1	1	1	·	4

Fig. 1. Journal citation frequencies. The data show the total number of times each journal was cited during the last quarter of 1969 and the distribution by publication date of the particular issues cited. The journals shown were taken from a list of more than 20,000 items (journals, books, reports, theses, and so on) cited during the last quarter of 1969 in journals covered by the SCI.

convince us that the 3-month data constitute a valid sample.

With this data from the last quarter of 1969, ISI produced three listings that should greatly further efforts to map the network of journal information transfer. The first listing is a cumulation of all citations of the same titles. It gives the number of times each different title was cited during the last quarter of 1969 and distributes that total over the years in which cited issues, editions, and so on were published. This distribution is shown by year from 1969 back to 1960 and in aggregate form for earlier years (*11*). Sample items from this listing are shown in Fig. 1.

The second listing is a detailed citation history of each cited title. It shows how frequently each title was cited and, as above, gives subtotals by year of publication. These citation totals and subtotals are further broken down to show how frequently each journal covered by SCI cited the title in question and the distribution by year of publication of cited items. A portion of this listing is shown in Fig. 2.

The third listing is similar to the second, but it arranges the data by citing journal rather than by cited title. The listing shows, for each journal covered by the SCI, the number of references published in issues processed for the SCI during the last quarter of 1969, and it breaks that total down by publication date of the items to which reference was made. The listing further identifies all titles cited in those references and the frequency with which they are cited. As in Fig. 2, the counts for each cited title are broken down by year of publication. A portion of this listing is shown in Fig. 3.

These listings afford, I believe, a new view of the literature in scientific and technical journals. Before discussing some of its implications, however, I note possible limitations of the data and problems encountered in analyzing

CITED JOURNAL / CITING JRNL	TOTAL	1969	1968	1967	1966	1965	1964	1963	1962	1961	1960	REST
J LINN SOC BOT	17*	0	3	0	3	1	0	0	0	1	1	8
NEW PHYTOL	5	0	1	0	1	0	0	0	0	1	0	2
ALL OTHER (10)	12	0	2	0	2	1	0	0	0	0	1	6
J LIPID RES	902*	33	108	109	98	133	121	75	83	60	58	24
BIOC BIOP A	83	5	8	16	6	13	15	4	6	5	1	4
BIOCHEM J	39	0	4	7	3	6	7	3	3	4	1	1
J BIOL CHEM	39	0	6	7	3	8	3	6	2	2	2	0
J CHROMAT	30	2	3	4	4	2	6	0	4	2	1	2
J CLIN INV	29	1	3	3	3	7	2	2	3	3	1	1
J LIPID RES	28	5	5	4	2	4	1	2	0	3	1	1
P SOC EXP M	25	0	3	1	3	4	3	2	5	1	1	2
MILIT MED	5	0	0	2	0	0	0	1	1	0	1	0
NY ST J MED	5	0	1	1	2	0	0	0	1	0	0	0
ALL OTHER (123)	219	6	19	16	29	43	27	23	15	15	23	3
J LOND MATH SOC	173*	16	19	19	11	8	10	2	8	4	8	68
J LOND MATH	71	8	12	7	5	4	4	0	4	2	4	21
T AM MATH S	11	2	1	2	1	0	0	0	1	0	0	4
P CAMB PHIL	8	1	0	3	0	0	0	0	0	0	0	4

Fig. 2. Statistics on cited journals. The data on cited journals show the total number of times each journal was cited in the last quarter of 1969 and the distribution by publication date of cited issues. For each cited journal, the figure identifies all other journals ("citing journals") that referred to it five times or more during the quarter year (and the distribution of these citations by publication date of cited issues). Journals that referred to the cited journal less than five times are grouped as "all other." The data were taken from a complete list of journals cited during the last quarter of 1969 in journals covered by the SCI.

the sample.

The SCI data base is, to my knowledge, the largest and most extensive of its kind. It does not, however, cover every scientific and technical journal. (Nor, as seems likely in view of findings discussed later in this article, need it attempt to.) The list of most frequently cited journals (the first 152 of these journals are given in Fig. 4) shows that the SCI has been remarkably successful in covering all "significant" and "important" journals, insofar as citation counts can be considered a reliable measure of "importance" and "significance." It is, of course, possible that one or more journals not covered by the SCI, and thus not represented in the data here, may cite themselves and other journals so frequently that their inclusion would alter their own and other journals' rankings. Such may be the case, for example, with certain journals in foreign languages, particularly those that do not use the Roman alphabet. As is true of most secondary services, the SCI is less likely to cover a journal that presents problems of transliteration (or transcription) and translation than one that does not. It may be, therefore, that this fact has adversely influenced the ranking of Russian and Japanese journals, for example, which probably cite other Russian and Japanese journals more frequently than do journals in other languages. Whether such an underrepresentation exists, and, if so, to what extent, is difficult to determine. Nevertheless, Soviet information scientists have reported that the SCI does a surprisingly good job of covering Soviet journals. In fact, Soviet scientists seem to have made more use of citation analysis in studies of science policy than have any other scientists (*12–14*).

One must remember that the listings were prepared from a 3-month sample of journal issues. The size of the sample is certainly more than adequate for statistical purposes, and, as noted, the sample has been matched against another sample of more than adequate size. It is nevertheless possible that random events in journal publishing have introduced some degree of distortion. For example, a journal may have, in the time period covered by the sample, departed from its usual policy of accepting only original research communications and published one or two review articles with extensive bibliographies. In the sample, that journal would appear to cite other publications more widely and more frequently than it actually does on the average. Or, a journal may have published an article that has since been cited with extraordinary frequency (*15*). In such a case, a single article will have had an inordinate influence on the ranking of the journal (*16*). Finally, a journal that publishes relatively few articles, but articles of high quality that both cite and are cited frequently, may seem to have considerably less impact than it actually does, particularly if the journal appears infrequently or irregularly and thus escapes representative inclusion in a sample of this type.

In analysis of the sample, an immensely irksome problem was the inconsistency with which different authors and editors abbreviate journal titles in their references. As far as possible, this

SOURCE JOURNAL / REFERENCE JOURNAL	TOTAL	NUMBER OF TIMES CITED 1969	1968	1967	1966	1965	1964	1963	1962	1961	1960	REST
J LIBR AUT ============>	168*	15	50	35	19	8	14	5	6	3	1	12
PROGRAM	15	3	6	5	1	0	0	0	0	0	0	0
J LIBRARY AUTOMATION	7	3	4	0	0	0	0	0	0	0	0	0
ALL OTHER (120)	146	9	40	30	18	8	14	5	6	3	1	12
J LIPID RES ============>	313*	15	36	35	26	25	29	20	10	15	15	87
J BIOL CHEM	43	2	6	3	6	1	3	1	0	1	1	19
J LIPID RES	28	5	5	4	2	4	1	2	0	3	1	1
BIOCHIM BIOPHYS ACTA	19	1	3	1	3	2	3	1	0	2	2	1
BIOCHEM J	13	1	1	3	0	0	0	1	2	1	1	3
J AMER CHEM SOC	12	0	0	0	0	0	0	0	0	1	0	11
BIOCHEMISTRY	9	0	3	0	2	1	2	1	0	0	0	0
J CHEM SOC LONDON	5	0	0	1	0	0	0	0	0	1	0	3
J CHROMATOGR	5	0	1	1	1	0	1	0	1	0	0	0
METHODS ENZYMOL	5	0	0	0	0	0	0	0	1	0	0	4
ALL OTHER (89)	124	5	12	19	10	8	16	10	2	5	6	31
J LOND MATH ==========>	743*	21	44	58	52	42	34	31	38	29	33	361
J LOND MATH	52	6	12	5	5	4	3	0	2	1	3	11
P LONDON MATH SOC	37	4	5	7	1	2	2	0	1	1	0	14

Fig. 3. Statistics on citing journals. The data on each citing journal ("source journal") show the total number of references each journal contained in the last quarter of 1969 and the distribution of that total by publication date of journal issues referred to. For each citing journal, the figure identifies all journals ("reference journal") cited five times or more during the quarter year (and gives the distribution by publication data of cited issue). Journals cited less than five times are grouped as "all other." The data are taken from a list of journals processed for the SCI during the last quarter of 1969.

inconsistency has been minimized by standardizing all variants of the same titles and their abbreviations. Some idea of the work involved in this standardization can be had from the fact that there were more than 100,000 different abbreviations for the 12,000 individual journal titles cited in the 3-month sample (*17*). Inconsistency was made worse by inaccuracy. In some cases, it was possible to disentangle the results of bibliographic carelessness—as, for example, when *Sol. St. Phys.* proved to have been used indiscriminately to identify *Solid State Physics* (Academic Press); *Solid State Physics, Proceedings of the Physical Society, London; Soviet Physics Solid State* (a cover-to-cover translation of *Fizika Tverdogo Tela*); and even *Physica Status Solidi* (Akademie-Verlag). In other cases, however, it was impossible, without going to inordinate expense, to determine exactly which journal was being cited—for example, when *Ann. Phys.* was found to have been used for *Annalen der Physik, Annals of Physics,* and *Annales de Physique*. It is not surprising that the editors of at least one publishing house, having decided that the problem of unique and unambiguous journal title abbreviations is simply insoluble, now use full titles in every reference to a journal.

Finally, it was necessary to make some arbitrary decisions in order to avoid unduly complex bibliographical technicalities. Journals merge; they split into new journals, or into "sections" that may be published separately or together. They change titles, with or without continuing their numbering of volumes and issues. Some journals appear in one or more translations; some such translations are complete, others selective, and some are similarly, others differently, numbered. Supplements outside a regularly enumerated series must be accounted for. In a few cases, journals periodically change their titles on single issues to note special subject matter (*18*). Serials librarianship abounds in difficulties of this type, and there is frequently disagreement on how best to handle them. Briefly: a journal published in sections is considered a single journal; translations of journals are identified with the original versions; changes of title have been ignored and previous volumes attributed to the current title; journals absorbed or incorporated by other journals have been credited to a new or remaining title; supplements have been considered as issues in the regular series. For the purposes of this analysis, such arbitrary decisions seem justified; as required in the future, the raw citation data can, of course, be compiled and manipulated in such a way as to differentiate between changed titles, sections, translations, and so forth.

Some Preliminary Analyses

Figure 4 shows the result of a familiar application of citation analysis. It is a listing of journal titles ranked by the frequency with which they were cited in the references of journals indexed for SCI. This partial listing gives the top 152 of the 565 most frequently cited journals of science and technology. (The top 152 account for 50 percent of all references to journals.)

It is apparent, even from the makeup of this partial listing, that a good multidisciplinary journal collection need contain no more than a few hundred titles. That is not to say that larger collections cannot be justified, but it does say something indisputable, in terms of cost and benefit, about how large a jour-

Item No. (1)	Cited Journal (2)	Times Cited Last Quarter 1969 (3)	1969 Citations to 1967 and 1968 Articles (4)	Articles Published in 1967 and 1968 (5)	Impact Factor (6)
1	J AM CHEM SOC	26323	22156	3946	5.614
2	PHYS REV	20674	20740	5767	3.596
3	J BIOL CHEM	17112	10768	1777	6.059
4	NATURE LONDON	15325	15956	6811	2.342
5	J CHEM SOC	14028	17764	5827	3.048
6	J CHEM PHYS	13690	11696	3738	3.128
7	SCIENCE	9752	11880	3968	2.993
8	BIOCHIM BIOPHYS ACTA	9550	10956	3531	3.102
9	P NAT ACAD SCI USA	8260	11548	1348	8.566
10	BIOCHEM J	7638	6348	2074	3.060
11	LANCET	7617	8164	5496	1.485
12	PHYS REV LETT	6581	11380	2317	4.911
13	CR ACAD SCI	5789	6576	8345	0.788
14	AM J PHYSIOL	5420	3156	1013	3.115
15	J ORG CHEM	5401	5756	2475	2.325
16	J APPL PHYS	5190	5072	2880	1.761
17	P SOC EXP BIOL MED	5079	3468	1920	1.806
18	J MOL BIOL	4982	7340	833	8.811
19	J PHYSIOL LOND	4966	3036	1248	2.432
20	P ROY SOC LOND	4864	1916	621	3.085
21	J CELL BIOL	4813	4596	1357	3.386
22	J CLIN INVEST	4785	3652	1086	3.362
23	J PHYS CHEM	4703	4516	1939	2.329
24	CHEM BER	4541	2128	1037	2.052
25	NEW ENGL J MED	4512	5252	2226	2.359
26	J AM MED ASS	4492	3980	3787	1.050
27	BRIT MED J	4304	4224	6238	0.677
28	SOV PHYS JETP	4295	3400	754	4.509
29	ASTROPHYS J	4271	5440	1167	4.661
30	ANALYT CHEM	4259	2424	1510	1.605
31	J BACTERIOL	4147	4712	1410	3.341
32	BIOCHEMISTRY	4076	6344	1114	5.694
33	NUCL PHYS	4034	6716	2345	2.863
34	PHYS LETT	3943	7160	3034	2.359
35	TETRAHEDRON LETT	3937	8252	2902	2.843
36	J EXP MED	3871	2700	325	8.307
37	ANN NY ACAD SCI	3787	2344	1216	1.927
38	ARCH BIOCHEM BIOPHYS	3689	3776	1169	3.230
39	J GEOPHYS RES	3537	5312	1569	3.385
40	J POLYM SCI	3458	2888	2069	1.395
41	BIOCHEM BIOPHYS RES	3417	5108	1190	4.292
42	FED P	3372	4036	7374	0.547
43	J PHYS	3308	3256	2379	1.368
44	T FARADAY SOC	2922	1808	879	2.056
45	ACTA CRYSTALLOGR	2917	2164	1803	1.200
46	DOKL AKAD NAUK SSSR	2869	2456	5385	0.456
47	J PHARMACOL EXP THER	2781	2020	566	3.568
48	ANGEW CHEM	2728	3660	1251	2.925
49	J IMMUNOL	2627	2992	726	4.121
50	INORG CHEM	2620	3976	1247	3.188
51	SOV PHYS SOLID STATE	2620	2984	1561	1.911
52	CIRCULATION	2601	2624	2160	1.214
53	ENDOCRINOLOGY	2548	2276	783	2.906
54	ACTA CHEM SCAND	2444	1984	943	2.103
55	NUOVO CIMENTO	2431	3436	1938	1.772
56	B SOC CHIM FRANCE	2416	2664	2704	0.985
57	VIROLOGY	2376	2620	584	4.486
58	CANCER RES	2349	2344	814	2.879
59	CAN J CHEM	2280	2392	1182	2.023
60	HELV CHIM ACTA	2249	1524	539	2.827
61	Z NATURFORSCHUNG	2200	2172	1650	1.316
62	AM J MED	2191	1784	395	4.516
63	J LAB CLIN MED	2120	1284	754	1.702
64	TETRAHEDRON	2071	3220	1313	2.452
65	EXP CELL RES	1958	1464	653	2.241
66	LIEBIGS ANN CHEM	1952	768	492	1.560
67	ANN INT MED	1946	1844	1098	1.679
68	PHIL MAG	1943	1180	547	2.157
69	J CLIN ENDOCR METAB	1903	1888	488	3.868
70	J APPL PHYSIOL	1836	1460	643	2.270
71	ACTA PHYSIOL SCAND	1816	1024	413	2.479
72	J PHYS SOC JAP	1786	1768	2074	0.852
73	Z PHYS	1764	1228	844	1.454
74	CIRC RES	1750	1820	432	4.212
75	PHYTOPATHOLOGY	1713	1632	1597	1.021
76	J NAT CANCER I	1668	1672	417	4.009

Fig. 4. The 152 most frequently cited journals ranked by frequency of citation in journals covered by the SCI. Column 1 gives rank, and column 2 gives abbreviations of the titles of cited journals. Column 3 shows the total number of times each journal was cited during the last quarter of 1969. Column 4 gives an estimate of the total number of citations in 1969 of items published in 1967 and

Item No. (1)	Cited Journal (2)	Times Cited Last Quarter 1969 (3)	1969 Citations to 1967 and 1968 Articles (4)	Articles Published in 1967 and 1968 (5)	Impact Factor (6)
77	AM J OBSTET GYNECOL	1657	1440	1193	1 207
78	PLANT PHYSIOL	1646	1808	1149	1 573
79	IND ENG CHEM	1644	928	856	1 084
80	ANN SURG	1641	1036	642	1 613
81	B CHEM SOC JAP	1639	2004	1567	1 278
82	EUR J BIOCHEM	1635	1992	501	3 976
83	GENETICS	1618	1340	738	1 815
84	BLOOD	1614	1256	566	2 219
85	P IEEE	1610	1856	756	2 455
86	J OPT SOC AM	1587	1196	1322	0 904
87	ANALYT BIOCHEM	1519	1672	502	3 330
88	J GEN PHYSIOL	1507	1208	407	2 968
89	ARCH INTERN MED	1501	860	486	1 769
90	AM HEART J	1453	1036	539	1 922
91	J EXP PSYCHOL	1449	1152	644	1 788
92	J GEN MICROBIOL	1445	1136	534	2 127
93	J COMP PHYSIOL PSYCH	1444	888	476	1 865
94	J PHYS CHEM SOLIDS	1430	1572	801	1 962
95	CANCER	1416	1224	593	2 064
96	AM J PATHOL	1401	960	529	1 814
97	RUSS J PHYS CHEM	1400	1116	1545	0 722
98	METHODS ENZYMOL	1391	1456	482	3 020
99	J INORG NUCL CHEM	1391	1356	908	1 493
100	PEDIATRICS	1382	1060	709	1 495
101	SURG GYNECOL OBSTET	1374	868	535	1 622
102	ANAT REC	1365	752	1836	0 409
103	REV MOD PHYS	1364	816	189	4 317
104	T MET SOC AIME	1359	1196	901	1 327
105	CAN J PHYS	1352	2156	1019	2 115
106	BRIT J PHARMACOL	1348	1348	507	2 658
107	APPL PHYS LETT	1337	2556	721	3 545
108	PHYS STAT SOLIDI	1329	2192	1485	1 476
109	J ELECTROCHEM SOC	1308	1208	1538	0 785
110	ACTA METALLURG	1304	964	452	2 132
111	PHYS FLUIDS	1304	1548	1050	1 474
112	EXPERIENTIA	1297	1592	1565	1 017
113	GASTROENTEROLOGY	1286	1428	1244	1 147
114	Z ZELLF MIKR ANAT	1286	1800	653	2 756
115	SURGERY	1274	996	790	1 260
116	REV SCI INSTR	1273	968	1148	0 843
117	AM J ROENTGENOL	1272	1044	860	1 213
118	AIAA J	1269	1456	1231	1 182
119	T ASME	1246	800	1332	0 600
120	AM J CARDIOL	1238	1600	737	2 170
121	J HISTOCHEM CYTOCHEM	1229	828	362	2 287
122	J PEDIAT	1229	1076	783	1 374
123	J ACOUST SOC AM	1219	1016	2196	0 462
124	NATURWISSENSCHAFTEN	1218	944	1091	0 865
125	J NUTR	1209	952	489	1 946
126	SPECTROCHIM ACTA	1201	1248	679	1 837
127	Z ANORG ALLG CHEM	1188	580	549	1 056
128	J PERSON SOC PSYCHOL	1186	676	581	1 163
129	RADIOLOGY	1175	1244	835	1 489
130	AM J BOT	1171	644	726	0 887
131	Z PHYS CHEM LEIPZIG	1170	332	252	1 317
132	J CHROMATOGR	1161	1708	1343	1 271
133	HOPPE SEYLERS Z	1145	1412	863	1 636
134	J UROL	1142	656	712	0 921
135	ARCH PATHOL	1138	556	409	1 359
136	ARCH SURG	1134	748	867	0 862
137	AM J DIS CHILD	1127	748	610	1 226
138	ACTA MED SCAND	1112	680	472	1 440
139	ANN PHYSICS	1105	692	224	3 089
140	COLD SPR HARB SYMP	1091	1060	194	5 463
141	J ORGANOMET CHEM	1089	2784	796	3 497
142	PFLUGERS ARCH	1083	896	732	1 224
143	OPT SPECTROSC USSR	1076	1100	814	1 351
144	KLIN WSCHR	1057	800	1198	0 667
145	CHEM IND LOND	1049	648	1703	0 380
146	BER BUNSEN PHYS CHEM	1044	688	771	0 892
147	BIOCHEM PHARMACOL	1030	1292	684	1 888
148	PHYSIOL REV	1022	572	33	17 333
149	J BONE JOINT SURG	1021	500	745	0 671
150	J NEUROPHYSIOL	1015	692	156	4 435
151	CR SOC BIOL	1010	596	1316	0 452
152	REC TRAV CHIM	1010	728	337	2 160

1968 (the estimate was made by quadrupling the 1969 citations of 1967 and 1968 items in the 3-month sample). Column 5 shows the total number of items processed from each journal by the SCI during 1967 and 1968. Column 6 indicates the impact factor (average *citations per published item)* derived by dividing the numbers in column 4 by those in column 5.

nal collection need be (or how small it can be) if it is to provide effective coverage of the literature most used by research scientists.

It is also immediately apparent that the majority of all references cite relatively few journals. Figure 5, which plots the distribution of citations among cited journals, shows that only 25 journals (little more than 1 percent of SCI coverage) are cited in 24 percent of all references; that only 152 journals (those listed in Fig. 4) are cited in 50 percent of all references; that only 767 journals are cited in 75 percent of all references; and that only 2000 or so journals are cited in 85 percent of all references. In addition, the data from which Fig. 5 was plotted show that only 540 journals are cited 1000 or more times a year, and that only 968 journals are cited even 400 times a year. When one considers that only 165 or so journals publish 400 or more papers a year, the impact of the average paper must be recognized as relatively slight. In fact, the average paper is cited only 1.7 times a year (*19*).

This analysis gives good reason for concern about any increase in the number of scientific and technical journals. It is not merely that increased numbers of journals make coverage of the literature more difficult, but that so many journals now being published seem to play only a marginal role, if any, in the effective transfer of scientific information. If one accepts the contention (highly debatable, in my opinion) that there are between 50,000 and 100,000 scientific and technical "journals," the data presented here indicate that only 5 to 6 percent of them are being cited. If such percentages seem unrealistically low, it may be because the estimate of 50,000 to 100,000 scientific and technical journals (requiring indexing and abstracting for "total coverage" of the literature) is itself as unrealistic as information scientists have for some time suspected it must be (*20*). Meaningful discussion of this point—the best of current serials catalogs notwithstanding —is probably impossible in the absence of agreement on what, quantitatively as well as qualitatively, constitutes a "scientific journal." At the very least, it may be advisable to distinguish as journals only those periodicals that publish, for example, 20 or more articles a year.

The predominance of a small group of journals in the citation network has been confirmed by a weekly scanning of SCI input to a system for selective dissemination of information (SDI) (*21*). In this SDI system, a newly published article can be retrieved on the basis of journals cited in the article's bibliography or footnotes. This retrieval criterion is known in an SDI profile as a "cited-journal question." A retrieval profile consisting of only 25 different cited-journal questions will retrieve 50 percent of all articles processed for the SCI every week. In other words, half of all articles published cite at least one of the 25 most frequently cited journals at least once.

It is also interesting that a small group of journals is found to be predominant when the literature is analyzed in other ways. Figure 6, for example, plots numbers of articles published against numbers of journals. It shows that, of the 2200 journals covered by the SCI in 1969, about 500 published about 70 percent of all articles published. As shown in Fig. 7, a small group of 250 journals provided almost half of the 3.85 million references processed for the SCI in 1969. The predominance of cores of journals is ubiquitous. An analysis of what

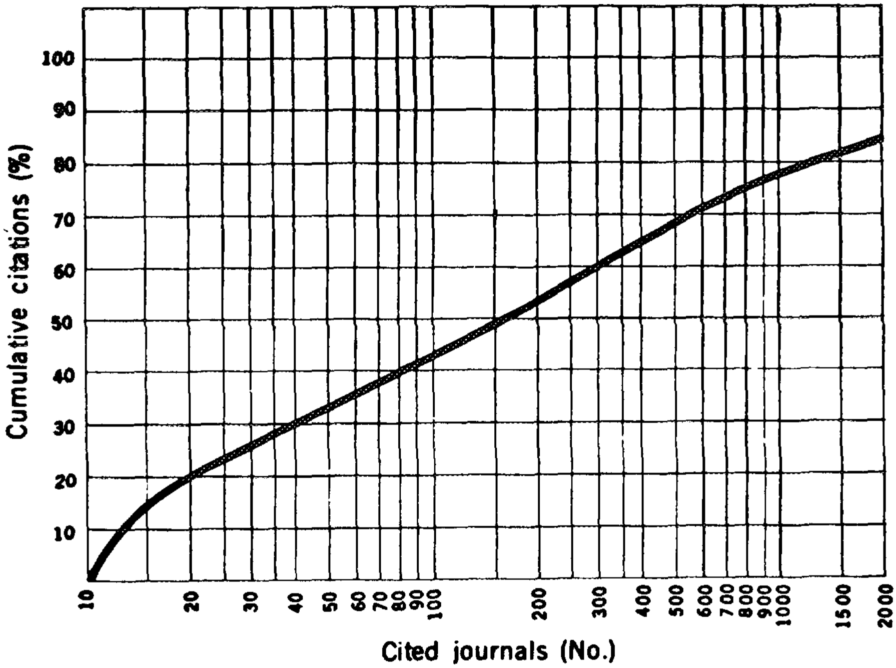

Fig. 5. Distribution of citations among cited journals. The curve shows that a relatively small core of 152 journals accounts for about half of all citations and that only 2000 or so journals account for about 84 percent of all citations.

journals first published reports of newly synthesized compounds indexed for *Current Abstracts of Chemistry and Index Chemicus* gives a similar result: of the 183 journals indexed by this publication in 1969, 11 percent of the journals accounted for 60 percent of the new compounds, 22 percent of the journals for 79 percent of the compounds, 32 percent of the journals for 89 percent of the compounds, and so on (*22*). *Chemical Abstracts* presents an even more striking example of this predominance: about 8 percent of the journals it covers publish more than 75 percent of the items it abstracts (*23*).

The data reported here demonstrate the predominance of a small group of journals in the citation network. Indeed, the evidence seems so conclusive that I can with confidence generalize Bradford's bibliographical law concerning the concentration and dispersion of the literature of individual disciplines and specialties (*2*). Going beyond Bradford's studies, I can say that a combination of the literature of individual disciplines and specialties produces a multidisciplinary core for all of science comprising no more than 1000 journals. The essential multidisciplinary core could, indeed, be made up of as few as 500 journals if, for example, one is attempting to satisfy the needs of libraries in developing countries.

Other Considerations

Citation frequency reflects a journal's value and the use made of it, but there are undoubtedly highly useful journals

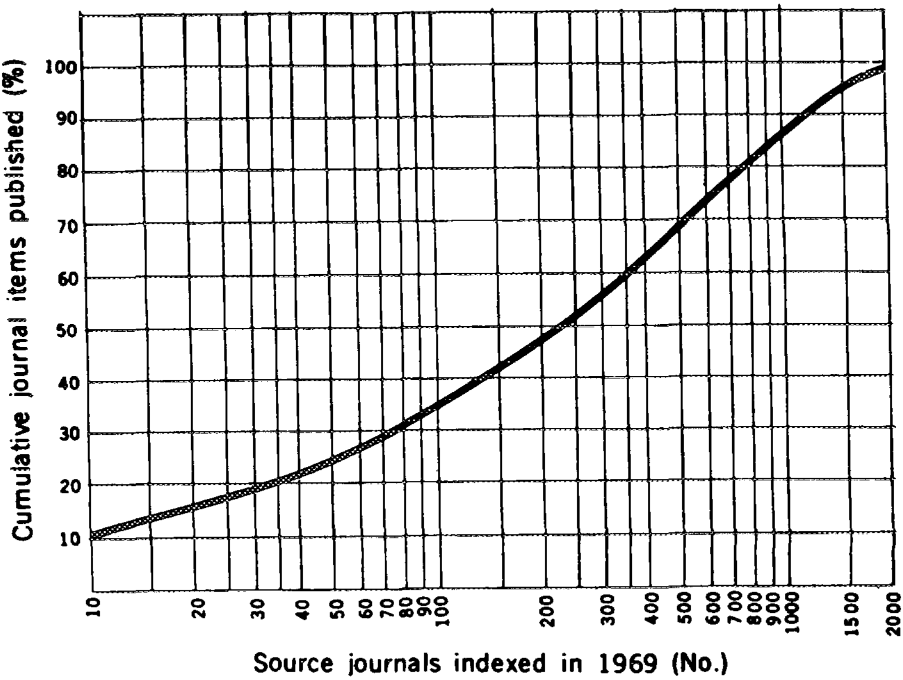

Fig. 6. Distribution of the number of published items among the approximately 2200 journals covered by the SCI in 1969. The curve shows that a relatively small core of journals carried the majority of items published.

that are not cited frequently. Scientists read some journals for the same reason people read newspapers—to keep up with what's going on generally—and they may rarely or never cite such journals in their published work (*24*). A popular review journal such as *Scientific American* or a news-oriented journal such as *New Scientist* may rank relatively low on a times-cited list (in fact, *Scientific American* is 449th, while *New Scientist* ranks well below 1000), but that does not mean that they are therefore less important or less widely used than journals that are cited more frequently. It merely means that they are written and read primarily for some purpose other than the communication of original research findings.

Citation frequency is, of course, a function of many variables besides scientific merit. Some of them are known or can reasonably be assumed: an author's reputation, controversiality of subject matter, circulation, availability and extent of library holdings, reprint dissemination, coverage by secondary services, priority in allocation of research funds, and others. It is extremely difficult, even when possible, to clarify the relations among such variables and their relative impact on citation frequency. One such variable is, however, fairly obvious. If every article has an equal likelihood of being cited, it should follow that the more articles a journal publishes, the more frequently the journal will be cited. For the most part, the data show that such is indeed the case. Although many articles are never cited (*25, 26*), I have very rarely found among the 1000 most frequently cited journals one that is not also among the 1000 journals that are most productive in terms of articles published. Citation frequency of a journal is thus a function not only of the scientific significance of the material it publishes (as reflected by citation), but also of the amount of material it publishes.

In view of the relation between size and citation frequency, it would seem desirable to discount the effect of size when using citation data to assess a journal's importance. We have attempted to do this by calculating a relative impact factor—that is, by dividing the number of times a journal has been cited by the number of articles it has published during some specific period of time. The journal impact factor will thus reflect an average citation rate per published article (*27*). However, the development of impact factors that fairly relate the size of a journal during the cited years to its current citation rate is a formidable challenge to statistical analysis. With the SCI data base, it is easy to determine how frequently a journal has been cited within a given period of time, but it is much more difficult to agree on a total-items-published base to which such citation counts can properly be related because the items may have been published at any point in the journal's history. In selecting an items-published base (*28*) for each journal, I have been guided by the chronological distribution of cited items in each annual edition of the SCI (*19*, p. 15; *29*). An analysis of

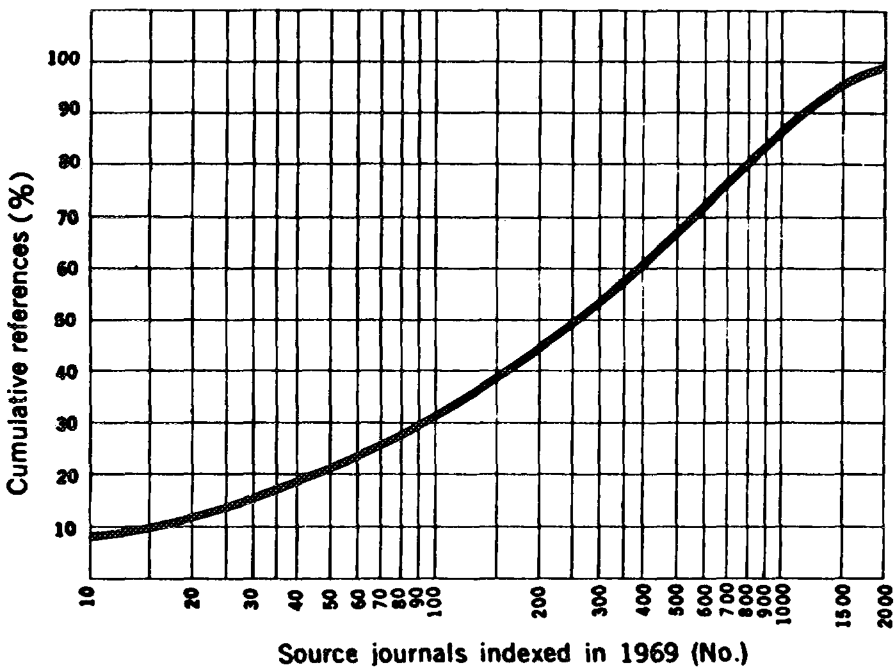

Fig. 7. Distribution of references among journals covered by the SCI in 1969. The curve shows that fewer than 300 journals provided more than half of the references processed.

Item No. (1)	Cited Journal (2)	Times Cited Last Quarter 1969 (3)	1969 Citations to 1967 and 1968 Articles (4)	Articles Published in 1967 and 1968 (5)	Impact Factor (6)
1	ADV PROTEIN CHEM	373	184	8	23 000
2	PHARMACOL REV	725	448	20	22 400
3	BACTERIOL REV	646	804	39	20 615
4	ANNU REV BIOCHEM	468	932	53	17 584
5	PHYSIOL REV	1022	572	33	17 333
6	ACCOUNTS CHEM RES	247	820	48	17 083 *
7	SOLID STATE PHYS	384	228	14	16 285
8	ADV ENZYMOL	291	192	20	9 600
9	INT REV CYTOL	230	144	16	9 000
10	J MOL BIOL	4982	7340	833	8 811
11	REC PROG HORMONE RES	417	232	27	8 592
12	P NAT ACAD SCI USA	8260	11548	1348	8 566
13	J EXP MED	3871	2700	325	8 307
14	Q REV	488	452	55	8 218
15	CHEM REV	1003	408	50	8 160
16	ANNU REV PL PHYSIOL	314	296	42	7 047
17	J CRYST GROWTH	232	820	125	6 560
18	ANNU REV MICROBIOL	254	288	44	6 545
19	J BIOL CHEM	17112	10768	1777	6 059
20	METHODS BIOCHEM ANAL	285	80	14	5 714
21	BIOCHEMISTRY	4076	6344	1114	5 694
22	J AM CHEM SOC	26323	22156	3946	5 614
23	SOV PHYS USP	586	612	109	5 614
24	COLD SPR HARB SYMP	1091	1060	194	5 463
25	BIOL REV	358	176	34	5 176
26	J VIROL	560	1860	360	5 166
27	MEDICINE	410	240	48	5 000
28	J CELL SCI	552	600	122	4 918
29	PHYS REV LETT	6581	11380	2317	4 911
30	ASTROPHYS J	4271	5440	1167	4 661
31	AM J MED	2191	1784	395	4 516
32	SOV PHYS JETP	4295	3400	754	4 509
33	VIROLOGY	2376	2620	584	4 486
34	J NEUROPHYSIOL	1015	692	156	4 435
35	PSYCHOL REV	593	368	83	4 433
36	REV MOD PHYS	1364	816	189	4 317
37	BIOCHEM BIOPHYS RES	3417	5108	1190	4 292
38	MON NOT ROY ASTR SOC	868	1008	238	4 235
39	CIRC RES	1750	1820	432	4 212
40	J IMMUNOL	2627	2992	726	4 121
41	Q J MED	437	284	70	4 057
42	J NAT CANCER I	1668	1672	417	4 009
43	EUR J BIOCHEM	1635	1992	501	3 976
44	MOL PHARMACOL	300	564	144	3 916
45	DEVELOP BIOL	435	552	142	3 887
46	J CLIN ENDOCR METAB	1903	1888	488	3 868
47	CHEM ENG LONDON	268	392	104	3 769
48	J LIPID RES	929	876	235	3 727
49	ADV PHYS	318	284	77	3 688
50	PSYCHOL B	610	564	154	3 662
51	IMMUNOLOGY	801	1208	335	3 605
52	PHYS REV	20674	20740	5767	3 596
53	J PHARMACOL EXP THER	2781	2020	566	3 568
54	APPL PHYS LETT	1337	2556	721	3 545
55	J ORGANOMET CHEM	1089	2784	796	3 497
56	J CELL PHYSIOL	860	628	180	3 488
57	BRAIN RES	420	1140	327	3 486
58	BRIT MED B	426	432	127	3 401
59	J CELL BIOL	4813	4596	1357	3 386
60	J GEOPHYS RES	3537	5312	1569	3 385
61	J CLIN INVEST	4785	3652	1086	3 362
62	J BACTERIOL	4147	4712	1410	3 341
63	ANALYT BIOCHEM	1519	1672	502	3 330
64	IMMUNOCHEMISTRY	271	404	125	3 232
65	ARCH BIOCHEM BIOPHYS	3689	3776	1169	3 230
66	INORG CHEM	2620	3976	1247	3 188
67	J CHEM PHYS	13690	11696	3738	3 128
68	J ULTRASTRUCT RES	979	1392	445	3 128
69	AM J PHYSIOL	5420	3156	1013	3 115
70	TRANSPLANTATION	513	1000	321	3 115
71	BIOCHIM BIOPHYS ACTA	9550	10956	3531	3 102
72	ANN PHYSICS	1105	692	224	3 089
73	P ROY SOC LOND	4864	1916	621	3 085
74	BIOCHEM J	7638	6348	2074	3 060
75	J CHEM SOC	14028	17764	5827	3 048
76	METHODS ENZYMOL	1391	1456	482	3 020

Fig. 8. The 152 most frequently cited journals ranked by impact factor (average number of citations per item published). The column headings are explained in the legend of Fig. 4.

Item No. (1)	Cited Journal (2)	Times Cited Last Quarter 1969 (3)	1969 Citations to 967 and 1968 Articles (4)	Articles Published in 1967 and 1968 (5)	Impact Factor (6)
77	SCIENCE	9752	11880	3968	2 993
78	GENET RES	371	464	155	2 993
79	J GEN PHYSIOL	1507	1208	407	2 968
80	ANGEW CHEM	2728	3660	1251	2 925
81	ENDOCRINOLOGY	2548	2276	783	2 906
82	CANCER RES	2349	2344	814	2 879
83	EXP PARASITOL	437	492	171	2 877
84	NUCL PHYS	4034	6716	2345	2 863
85	TETRAHEDRON LETT	3937	8252	2902	2 843
86	PLANTA	707	1172	414	2 830
87	HELV CHIM ACTA	2249	1524	539	2 827
88	J COMP NEUROL	969	376	133	2 827
89	BIOPOLYMERS	452	656	235	2 791
90	CHROMOSOMA	458	440	159	2 767
91	Z ZELLF MIKR ANAT	1286	1800	653	2 756
92	CLIN SCI	680	552	205	2 692
93	BRIT J PHARMACOL	1348	1348	507	2 658
94	SURFACE SCI	399	844	321	2 629
95	AM J HUM GENET	405	332	128	2 593
96	PLANET SPACE SCI	508	892	348	2 563
97	DISCUSS FARADAY SOC	702	292	114	2 561
98	J NEUROCHEM	801	900	357	2 521
99	SOV J NUCL PHYS	742	1588	630	2 520
100	MUTAT RES	274	532	213	2 497
101	J CATAL	431	764	308	2 480
102	ACTA PHYSIOL SCAND	1816	1024	413	2 479
103	CHEM PHYS LETT	294	996	402	2 477
104	GEOCHIM COSMOCH ACTA	814	744	301	2 471
105	P IEEE	1610	1856	756	2 455
106	STEROIDS	473	680	277	2 454
107	TETRAHEDRON	2071	3220	1313	2 452
108	J PHYSIOL LOND	4966	3036	1248	2 432
109	INT J CANCER	275	452	189	2 391
110	PSYCHOPHARMACOLOGIA	277	388	163	2 380
111	NEW ENGL J MED	4512	5252	2226	2 359
112	PHYS LETT	3943	7160	3034	2 359
113	EARTH PLANET SC LETT	269	672	286	2 349
114	NATURE LONDON	15325	15956	6811	2 342
115	J PHYS CHEM	4703	4516	1939	2 329
116	J ORG CHEM	5401	5756	2475	2 325
117	J EXP ANALYSIS BEHAV	509	424	184	2 304
118	J HISTOCHEM CYTOCHEM	1229	828	362	2 287
119	J APPL PHYSIOL	1836	1460	643	2 270
120	AM J ANAT	637	256	113	2 265
121	EXP CELL RES	1958	1464	653	2 241
122	BLOOD	1614	1256	566	2 219
123	J FLUID MECH	998	1036	472	2 194
124	HISTOCHEMIE	323	668	305	2 190
125	AM J CARDIOL	1238	1600	737	2 170
126	REC TRAV CHIM	1010	728	337	2 160
127	PHIL MAG	1943	1180	547	2 157
128	J BIOCHEM	966	1064	498	2 136
129	ACTA METALLURG	1304	964	452	2 132
130	J GEN MICROBIOL	1445	1136	534	2 127
131	CAN J PHYS	1352	2156	1019	2 115
132	ANN MATH	702	184	87	2 114
133	ACTA CHEM SCAND	2444	1984	943	2 103
134	BRIT J HAEMATOL	581	608	290	2 096
135	METABOLISM	550	564	270	2 088
136	RADIO SCI	385	760	365	2 082
137	CANCER	1416	1224	593	2 064
138	PHOTOCHEM PHOTOBIOL	343	472	229	2 061
139	AM J SCI	602	284	138	2 057
140	T FARADAY SOC	2922	1808	879	2 056
141	CHEM BER	4541	2128	1037	2 052
142	MOL PHYS	698	652	319	2 043
143	CAN J CHEM	2280	2392	1182	2 023
144	J EXP BOT	352	336	167	2 011
145	B SEISMOL SOC AM	344	416	208	2 000
146	J SEDIMENT PETROLOGY	423	480	240	2 000
147	ARCH MIKROBIOL	438	600	305	1 967
148	DIABETES	785	936	477	1 962
149	J PHYS CHEM SOLIDS	1430	1572	801	1 962
150	INORG NUCL CHEM LETT	247	620	316	1 962
151	J ENDOCRINOL	983	1104	566	1 950
152	PROTOPLASMA	301	380	195	1 948

* There was an error in the numbers given in cols. 5 and 6 of this figure for *Accounts Chem Res* when this article was originally published in *Science*. The error has been corrected in this reprint.

this distribution has shown that the typical cited article is most heavily cited during the 2 years after its year of publication. (In any given year, 21 to 25 percent of all references cite articles that are 3 or fewer years old.) Therefore, since my sample consists of references made in 1969, I have taken as the items-published base for each journal the number of items it published during 1967 and 1968. To calculate an impact factor for each journal, I divided the number of times 1967 and 1968 articles were cited in 1969 by the number of articles published in 1967 and 1968. Martyn and Gilchrist used a similar method in ranking British journals in an analysis of 1965 SCI data (*30*).

Figure 8 shows the top 152 of the 565 most frequently cited journals ranked by impact factor. Many of the 152 journals do not appear on this high-impact list; in fact, only 75 journals are common to both lists. It will be interesting to observe further changes in the ranking of the most frequently cited journals as calculations of impact factor are extended to 1000 journals and eventually to the approximately 2400 journals now covered by the SCI.

Some Applications

The results of this type of citation analysis would appear to be of great potential value in the management of library journal collections. Measures of citation frequency and impact factor should be helpful in determining the optimum makeup of both special and general collections. Analysis of the chronological distribution of items cited can serve as a guide in determining the optimum size of back files, and, since the data give a detailed view of each journal's citation history, binding and retention schedules can be rationally established journal by journal, rather than for groups of journals (*31*). Another application, which harried librarians may welcome, is the correlation of data on citation frequency and impact with subscription costs. Such a correlation can provide a solid basis for cost-benefit analysis in the management of subscription budgets.

Individual scientists also face the problem of selecting journals to read and keep, as well as compiling reference and reading lists for themselves and their students. Although each of the relatively few journals that are very useful in a given discipline or specialty may be well known, it can be difficult to gauge the merits of the other journals in that discipline or specialty and to decide what journals to get and how long to keep them. It should be noted, in this connection, that analyses of citation frequency and impact factor can be tailored to the specific interests and requirements of individuals by restricting the number of citing journals to a small group of familiar titles. Thus, with a list of the ten or so most frequently cited chemical journals (or merely ten journals favored by a particular chemist), one can, by constructing lists of citation frequency and impact factor, gradually augment the small group and the citation data base with journals of demonstrable relevance. I am using this technique to establish a list of journals for the projected *Social Sciences Citation Index*.

Editors and editorial boards of scientific and technical journals may also find citation analysis helpful. As it is, those who formulate editorial policies have few objective and timely measures of their success. A wrong policy, or a policy wrongly implemented, may have serious effects on revenue and prestige, and the work of regaining readers and reputation can be difficult and expensive. Editors can find useful indicators of a journal's performance in the extent of self-citation, the number of times cited per year, and the distribution of citations among citing journals within and outside the specialty literature.

Perhaps the most important application of citation analysis is in studies of science policy and research evaluation. Price has shown how citation data can be used to identify research fronts (*25*). Soviet information scientists are using citation data to evaluate the implementation of science policy in the U.S.S.R. (*12*, *14*, *32*), and the sociological studies of Hagstrom and others (*33*) give convincing evidence of the utility of this approach.

Unanswered Questions

The data reported here suggest many avenues for further study. What, for example, is the significance of an abnormally high self-citation rate? Is it characteristic of parochialism, eccentricity, mediocrity? Does it indicate that a particular field of study has as yet no basis for interaction with other fields? Which is true in a particular case and how does one go about finding out?

What is the significance of a wide, multidisciplinary spread of titles cited in the references of a given journal or group of journals? Is it a measure of multidisciplinary activity? If so, is it a valid enough measure to warrant applying the Weinberg criterion (*34*) of multidisciplinary impact in order to determine the amount of government support merited by particular areas of research? Does the fact that *Ecology*, for example, cites more than 500 different journals in a total of 1000 references make it more multidisciplinary than the *Journal of the American Chemical Society*, which cites only twice as many journals in ten times the number of references, or than *Physical Review*, which cites only 600 or so journals in 15,000 references? Does the nonlinearity of publication and citation distributions among citing and cited journals confirm beyond doubt only that relatively few journals are primary nodes in the communications network, or does it have some other significance?

And what is the significance of a wide disparity between the number of journals cited by a given journal and the number that cite it? *Ecology* cites 500 or so journals but, in turn, is cited by only 115. What does this say, other than the obvious, about ecology and *Ecology*? Is the applicability of work done in ecology and being reported in *Ecology* much narrower than the interests of ecologists?

Several investigators of problems in science policy and science management are already using the SCI data base to explore such questions and many others. Each is trying, for different reasons, to build a model of the journal communication network that will provide more functional definitions of disciplines and specialties, that will make it possible to define in detail how different fields of knowledge interact, that will provide methods of predicting interdisciplinary impact, and that will perhaps provide more effective ways of monitoring research performance. Using the SCI data base to map the journal communications network may contribute to more efficient science.

References and Notes

1. P. L. K. Gross and E. M. Gross, College libraries and chemical education. *Science* **66**, 385 (1927).
2. S. C. Bradford, *Documentation* (Lockwood, London, 1953), 2nd ed.
3. E. S. Allen, Periodicals for mathematicians. *Science* **70**, 552 (1929).
4. P. L. K. Gross and A. O. Woodford, Serial literature used by American geologists. *Science* **73**, 660 (1931).
5. R. H. Hooker, A study of scientific periodicals. *Rev. Sci. Instrum.* **6**, 333 (1935).
6. H. H. Henkle, The periodical literature of biochemistry. *Bull. Med. Libr. Ass.* **27**, 139 (1938).
7. H. H. Fussler, Characteristics of the research literature used by chemists and physicists in the United States. *Libr. Quart.* **19**, 19 (1949); *ibid.*, p. 119.
8. C. H. Brown, *Scientific Serials: Characteristics and List of Most Cited Publications in Mathematics, Physics, Chemistry, Geology, Physiology, Botany, Zoology, and Entomology*, ACRL monograph No. 16 (Association of College and Reference Libraries, Chicago, 1956).
9. (a) R. C. Coile, Periodical literature for electrical engineers. *J. Doc.* **8**, 209 (1952).
(b) T. Barrett and M. A. Barrett, Journals most cited by chemists and chemical engineers. *J. Chem. Educ.* **34**, 35 (1957).
(c) R. E. Burton, Citations in American engineering journals. I. Chemical engineering. *Amer. Doc.* **10**, 70 (1959); II. Mechanical engineering. *ibid.* **10**, 135 (1959); III. Metallurgical engineering. *ibid.* **10**, 209 (1959).
(d) R. F. Trimble, The journals of inorganic chemistry. *J. Chem. Doc.* **3**, 79 (1963).
(e) P. K. Anderson, The periodical literature of ecology. *BioScience* **16**, 794 (1966).
(f) A. J. Meadows, The citation characteristics of astronomical research literature. *J. Doc.* **23**, 28 (1967).
(g) H. East and A. Weyman, A study in the source literature of plasma physics. *Aslib Proc.* **21**, 160 (1969).
(h) A. M. Hall, *The Use and Value of Citations; a State-of-the-Art Report.* Report No. R 70/4 (The Institution of Electrical Engineers, London, 1970).
(i) R. A. Davis and C. A. Bailey, *Bibliography of Use Studies*, project No. 195, prepared under NSF Grant No. GN-170 for the Office of Science Information Service, (Graduate School of Library Science, Drexel Institute of Technology, Philadelphia, March 1964).
10. (a) E. Garfield, Citation indexes for science. *Science* **122**, 108 (1955).
(b) ———, *Science Citation Index*, a new dimension in indexing. *Science* **144**, 649 (1964).
11. Citation frequency can be distributed according to years of publication only in the case of serial publications, periodicals, monographs with more than one edition, and so on.
12. (a) G. M. Dobrov, "Scientific potential as an object of investigation and control in the Soviet Union." In: *Ciba Foundation and Science of Science Foundation Symposium on Decision Making in National Science Policy.* A. de Reuck, M. Goldsmith, and J. Knight, Eds. (Churchill, London, 1968), p. 189-201.
(b) ——— "Ispol'zovanie matematicheskikh metodov i vychislitel'noi tekhniki dlia analiza tendentsii i prognozirovaniia nauchnotekhnicheskogo progressa" (Use of mathematical methods and electronic technology in analysis of trends and in forecast of scientific and technical progress). In: *Analiz Tendentsii i Prognozirovanie Nauchno-Tekhnicheskogo Progressa; Primenenie Matematicheskikh Metodov EVM v Issledovaniiakh po Istorii Nauki i Tekhniki (Trend Analysis and Forecast of Scientific and Technical Progress: EDP Application of Mathematical Methods in Investigations of the History of Science)*, G. M. Dobrov, Ed. (Izdatel'stvo Naukova Dumka, Kiev, 1967), p. 5-25.
13. R. S. Giliarevskii, Z. M. Mul'chenko, A. T. Terekhin, A. I. Cheryni, "Opyt izycheniia *Science Citation Index*," (Experience in study of the *Science Citation Index*). In: *Prikladnaia Dokumentalistika (Applied Documentation)*, V. V. Nalimov, Ed. (Akademiia Nauk SSSR, Nauchnyi Sovet po Kibernetike, Seriia Organizatsiia i Upravlenie, Izdatel'stvo Nauka, Moscow, 1968), p. 32–53.
14. A. A. Korennoi, Indeksy nauchnykh ssylok; ikh primenenie v naukovendenii i informatike (Indexes of scientific references; their use in the science of science and in information science). *Nauch-Tekh. Inform. Ser. II* No. 10, 3 (1968).
15. A list of articles that have been cited very frequently has been published: E. Garfield, "Citation indexing, historio-bibliography, and the sociology of science." In: *Proceedings of the Third International Congress of Medical Librarianship, Amsterdam, 5-7 May 1969*, (Excerpta Medica, Amsterdam, 1970), p. 187-204.
16. For example, in 1951 the *Journal of Biological Chemistry (JBC)* published an article by O. H. Lowry *et al.* on protein analysis: O. H. Lowry, N. J. Rosebrough, A. L. Farr, and R. J. Randall, Protein measurement with the folin phenol reagent. *J. Biol. Chem.* **193**, 265 (1951). That article represents only 0.7% of the *JBC*'s average quarterly output, but it accounts for 5% of all citations of the *JBC* in our quarter-year sample. Lowry's article is not as atypical of articles from *JBC* as it is atypical of all articles. Of the 50 articles most cited in 1967, 13 (26%) had appeared in *JBC*.
17. The largest number of variants of the same journal title and its abbreviations was 42, the total reached in the case of *Comptes Rendus Hebdomadaires des Seances de l'Academie des Sciences.*
18. For example, the *South African Journal of Obstetrics and Gynaecology*, the *South African Journal of Nutrition*, the *South African Journal of Radiology*, and others are composed of specially titled (and color-coded) issues from the regularly numbered series of the *South African Medical Journal.*
19. *Science Citation Index 1970 Guide & Journal Lists* (Institute for Scientific Information, Philadelphia, 1971), p. 14. For the years 1964 to 1970, the number of citations per year per cited item has been, respectively, 1.60, 1.65, 1.65, 1.66, 1.67, 1.67, and 1.73.
20. E. Garfield, The mystery of the transposed journal lists; wherein Bradford's law of scattering is generalized according to Garfield's law of concentration. *Current Contents* No. 31, p. 5 (August 4, 1971).
21. ——— and I. H. Sher, ISI's experiences with *ASCA* (Automatic Subject Citation Alert)—a

selective dissemination system. *J. Chem. Doc.* **7**, 147 (1967).

22. E. Garfield and G. S. Revesz, Ten years of synthetic chemical literature; a statistical analysis of *Index Chemicus* from 1960–1969. *Nature*, in press.
23. J. L. Wood, "The parameters of document acquisition at Chemical Abstracts Service," paper presented at the American University Eighth Annual Institute of Information Storage and Retrieval, Washington, D.C., 14–17 February 1966.
24. C. Scott, The science of science; what scientists read and why. *Discovery* (March 1969), p. 110.
25. D. J. D. Price, Networks of scientific papers. *Science* **149**, 510 (1965).
26. ———, Citation practice. *Aslib Proc.* **21**, 328 (1969).
27. The impact factor discussed in this article (average citation rate per published item) gives some idea of the frequency with which the "average" paper in a particular journal is cited. The factor is, of course, adversely affected by papers in the journal which are not cited at all, and, as noted above, favorably affected by papers cited with unusual frequency. The influence of uncited and very frequently cited papers can be discounted by considering the total number of citations in relation to cited items only (rather than in relation to all published items), or by considering only number of cited items (rather than total citations) in relation to all published items. For example, if the items-published base is 100 articles and one article has been cited 50 times while 99 articles have not been cited at all, the impact factor would be 0.50. The same impact factor of 0.50 could be derived for the journal that had half of its 100 articles cited only once and the other half not cited at all. If one considers only cited items, a per-cited-item impact factor can be derived to distinguish between the two journals. In the first case, the per-cited-item impact factor would be 50.0; in the second, 1.0. If one considers only the number of citable items cited, a citable-items-cited impact factor can be derived to further distinguish between the two journals. In the first case, the citable-items-cited impact factor would be 0.01; in the second, 0.5. Enormous amounts of computer time would be required to derive these different impact factors, although one must acknowledge their potential usefulness. It should be noted also that either of these impact factors can be derived from the other by dividing one of them into the impact factor used in this article.
28. The problem of selecting an items-published base is further complicated by the variety in the kinds of items published in scientific journals. Many journals publish only full-length reports of original research. Many others publish, in addition, editorials, technical communications, letters, notes, general correspondence, scientific news surveys and notes, book reviews, and so on; all of these are potentially citable items. I have not attempted in this article to limit the definition of items-published to lead articles, original communications, or the like. Even assuming it were possible to construct an acceptable classification that would accommodate all of the different kinds of published material, it would have been impossible for me, within the resources available for this article, to have examined individually each of the approximately 600,000 items that I use for the items-published base. If such a differentiation among kinds of material were included in an analysis such as this one, it is reasonable to assume that lead articles in such journals as *Science*, *Nature*, *Lancet*, and *Journal of the American Medical Association* would, as a group, have higher impact factors than those that are shown for these journals in Fig. 8.
29. The percentage (in terms of total citations) of citations of items that are 3 or fewer years old has been, for the years 1964 to 1970, 31.09, 30.24, 26.60, 25.91, 25.32, 25.18, and 23.95, respectively. It is interesting to note that the yearly percentage of such items has gradually decreased as SCI coverage has increased, while the citation rate per cited item has gradually increased (*19*). The significance of these trends is an interesting matter for future investigation.
30. J. Martyn and A. Gilchrist, *An Evaluation of British Scientific Periodicals*. Aslib Occasional Publication No. 1 (Aslib, London, 1968).
31. (a) D. MacRae, Jr., Growth and decay curves in scientific citations. *Amer. Sociol. Rev.* **34**, 361 (1969).
(b) R. W. Chadwick and K. W. Deutsch, Doubling time and half-life: two suggested conventions for describing rates of change in social science data. *Amer. Behav. Scientist* **11**, NS9 (1968).
32. (a) V. V. Nalimov and Z. M. Mul'chenko. *Naukometriia (Scientometrics)*. (Izdatel'stvo Nauka, Moscow, 1969).
(b) L. N. Beck, Key issues in Soviet information. *J. Amer. Soc. Inform. Sci.* **22**, 85 (1971).
33. (a) W. O. Hagstrom, Inputs, outputs, and the prestige of university science departments. *Sociol. Educ.* **44**, 375 (1971).
(b) A. E. Bayer and J. Folger, Some correlates of a citation measure of productivity in sciences. *Sociol. Educ.* **39**, 382 (1966).
(c) J. A. Creager and L. R. Harmon, *On-the-Job Validation of Selection Variables*. Tech. Rep. No. 26 (National Academy of Sciences—National Research Council, Washington, D.C., 1966).
(d) S. Cole and J. R. Cole, Scientific output and recognition; a study in the operation of the reward system in science. *Amer. Sociol. Rev.* **32**, 377 (1967).
(e) E. B. Parker, W. J. Paisley and R. Garrett, *Bibliographical Citations as Unobtrusive Measures of Scientific Communication*. (Institute for Communications Research, Stanford University, Stanford, Calif., 1967).
(f) J. Margolis, Citation indexing and evaluation of scientific papers. *Science* **155**, 1213 (1967).
(g) J. Larabi, Mesure de l'efficacite des laboratoires de recherche fondamentale selectionnes par le Centre National d'Etudes Spatiales. *Rev. Fr. Inform. Rech. Operat.* **3**, (No. V-3), 103 (1969).
(h) D. J. D. Price, Measuring the size of science. *Proc. Israel Acad. Sci. Humanities* **4**, 98 (1969).

(i) R. D. Whitley, Communication nets in science; status and citation patterns in animal physiology. *Sociol. Rev.* **17**, 219 (1969).
(j) S. Cole, Professional standing and the reception of scientific discoveries, *Amer. J. Sociol.* **76**, 286 (1970).
(k) C. R. Myers, Journal citations and scientific eminence in contemporary psychology, *Amer. Psychol.* **25**, 1041 (1970).
(l) H. Rothman and M. Woodhead, The use of citation counting to identify research trends, *J. Doc.* **27**, 287 (1971).
(m) J. Cole and S. Cole, Measuring the quality of scientific research; problems in the use of the *Science Citation Index*. *Amer. Sociol.* **6**, 23 (1971).
(n) J. P. Martino, Citation indexing for research and development management. *IEEE Transactions Eng. Management* **EM-18**, 146 (1971).
(o) Anyone interested in citation analysis in study of the history and sociology of science will find the following bibliography invaluable: A. Pritchard, *Statistical Bibliography—An Interim Bibliography* (North-Western Polytechnic School of Librarianship, London, May 1969). Available from the U.S. Department of Commerce, National Bureau of Standards, Clearinghouse for Federal Scientific and Technical Information.
34. A. M. Weinberg, Criteria for scientific choice. *Minerva* **1**, 159 (1963).

Reprinted from *Science* 178: 471-79, 1972.

"current comments"

Citation Frequency and Citation Impact; and the Role They Play in Journal Selection for *Current Contents* and Other ISI Services

February 7, 1973

For more than a year I've been citing what I consider to be one of the most significant papers I've ever published. At long last it has appeared.[1]

The paper deals primarily with the use of our *Journal Citation Index* data bank to determine the frequency with which scientific and technical journals are cited in the journal literature. It shows that a 'large' journal that publishes many articles is, as a rule, more frequently cited than a journal that publishes fewer articles. In addition, however, through development of 'impact factors', it shows that articles in about half of these most-cited journals are cited less frequently than articles in smaller, less-cited journals.

Regretably, the editors of *Science*, where the paper has been published, could not give space to include the originally submitted list of the 565 most-cited journals that are mentioned in 75% of all references published. But we intend to make this list available, and will update it sometime in 1973. The list of 152 journals that appears in the article is now three years old.

It should be significant to our subscribers that all of the 565 journals mentioned above—indeed almost all of the 1000 most-cited journals—are covered in either the *Science Citation Index* or in *Current Contents*.

Since completion of the work reported in the article, we have been able to measure the 'impact' of a much longer list of journals. As a result, we've determined that certain 'small' journals that publish relatively few but very frequently cited articles deserve to be included in the coverage of *SCI* and *CC*.

For many years, we've included review journals in *SCI* and *CC*, and our studies show we were right to do so. Although few are among the most-cited, several are at the top of a list ranked by impact. However, review journals are very expensive to process for *SCI*. The average review article contains from 3 to 10 times the number of references as the typical research article—some contain as many as 2000 references! This must be taken into account when selecting journals for the *SCI*. It is less important, of course, for *CC*.

I hope it is obvious that we take the question of journal coverage very seriously. ISI has devoted enormous time and energy to finding objective criteria for journal selection. Unfortunately, the objective criteria alone don't solve the problem. There is more that we feel should be considered. Readers frequently recommend that we cover journals that, for one reason or another, show up poorly in citation studies. It would be courting disaster to ignore all such recommendations on the basis of citation data alone. In the process we might denigrate the just motives of those who strongly support *CC*, whether from professional self-interest, national pride, and so on. All such decisions are acceptable provided that we diligently make certain that the *best* is always included.

Every year we go through the painful budgeting process that determines, among other things, how many journals we can add to each ISI service. It's plain that some readers imagine we can and should cover any journal that comes to our or their attention. Some seem actually shocked to learn that ISI is not some sort of inexhaustibly funded quasi-governmental agency. To expect that we cover any and every journal is almost as unreasonable as our own fervent wish that every scientist in the world subscribe to *Current Contents*. This might make it *economically* feasible for us to cover almost any journal, even if space did not call a halt at some point.

Once we have established our annual budget for journal coverage, we are invariably approached by some journal publisher or editor who wants some new journal covered immediately. What are we to do if the journal meets all important editorial requirements for selection? We can either tell the journal's sponsor to wait until next year, or we can ask him to underwrite coverage of his journal in order to help us live within the budget as we must.

It is never easy to drop a journal once it has been added to *Current Contents*. Librarians especially expect a continuity of coverage, but sometimes the facts of life require that we weed out journals that no longer 'cut the mustard.' Retaining a mediocre journal means omission of another, better journal. We are a long way from zero population growth of journals. In fact, older journals not only survive, but the hard core grow in size. Their growth in articles published almost equals the number in new journals. These journal dynamics are probably a healthy phenomenon in scientific communication.

If, in spite of all the hocus-pocus about journal selection discussed here, you feel that a particular journal is improperly omitted from *CC* or from *SCI*, please let us know. But don't fail to tell the editor or publisher of the journal as well. We'll always be responsive to readers' suggestions, though we must hope that you and he will appreciate the economics as well as the 'art' of journal coverage.

1. **Garfield E.** Citation analysis as a tool in journal evaluation. *Science* **178**:471–79, 1972. — In early citations of the article, I used the title on the MS submitted to *Science*: "Citation analysis as a sociometric tool for journal evaluation and science policy studies." Reprints are available.

The article cited in the reference above is reprinted in this volume beginning on page 527.

Reprinted in Essays of an Information Scientist, Vol:6, p.468-471, 1983

Reprinted from SCIENCE, July 15, 1955, Vol. 122, No. 3159, pages 108–111.

Citation Indexes for Science

A New Dimension in Documentation through Association of Ideas

Eugene Garfield, Ph. D.

"The uncritical citation of disputed data by a writer, whether it be deliberate or not, is a serious matter. Of course, knowingly propagandizing unsubstantiated claims is particularly abhorrent, but just as many naive students may be swayed by unfounded assertions presented by a writer who is unaware of the criticisms. Buried in scholarly journals, critical notes are increasingly likely to be overlooked with the passage of time, while the studies to which they pertain, having been reported more widely, are apt to be rediscovered." (*1*)

In this paper I propose a bibliographic system for science literature that can eliminate the uncritical citation of fraudulent, incomplete, or obsolete data by making it possible for the conscientious scholar to be aware of criticisms of earlier papers. It is too much to expect a research worker to spend an inordinate amount of time searching for the bibliographic descendants of antecedent papers. It would not be excessive to demand that the thorough scholar check all papers that have cited or criticized such papers, if they could be located quickly. The citation index makes this check practicable. Even if there were no other use for a citation index than that of minimizing the citation of poor data, the index would be well worth the effort required to compile it.

This paper considers the possible utility of a citation index that offers a new approach to subject control of the literature of science. By virtue of its different construction, it tends to bring together material that would never be collated by the usual subject indexing. It is best described as an association-of-ideas index, and it gives the reader as much leeway as he requires. Suggestiveness through association-of-ideas is offered by conventional subject indexes but only within the limits of a particular subject heading.

If one considers the book as the macro unit of thought and the periodical article the micro unit of thought, then the citation index in some respects deals in the submicro or molecular unit of thought. It is here that most indexes are inadequate, because the scientist is quite often concerned with a particular idea rather than with a complete concept. "Thought" indexes can be extremely useful if they are properly conceived and developed.

In the literature-searching process, indexes play only a small, although significant, part. Those who seek comprehensive indexes to the literature of science fail to point out that such indexes, although they may be desirable, will provide only a better *starting point* than the one provided in the selective indexes at present available. One of the basic difficulties is to build subject indexes that can anticipate the infinite number of possible approaches the scientist may require. Proponents of classified indexes may suggest that classification is the solution to this problem, but this is by no means the case. Classified indexes are also dependent upon a subject analysis of individual articles and, at best, offer us better consistency of indexing rather than greater specificity or multiplicity in the subject approach. Similarly, terminology is important, but even an ideal standardization of terminology and nomenclature will not solve the problem of subject analysis.

What seems to be needed, then, in addition to better and more comprehensive indexes, alphabetical and classified, are new types of bibliographic tools that can help to span the gap between the subject approach of those who create documents —that is, authors—and the subject approach of the scientist who seeks information.

Since 1873 the legal profession has been provided with an invaluable research tool known as *Shepard's Citations*, published by Shepard's Citations, Inc., Colorado Springs, Colo. (*2*). A citation index is published for court cases in the 48 states as well as for cases in Federal courts. Briefly, the Shepard citation system is a listing of individual American court cases, each case being followed by a complete history, written in a simple code. Under each case is given a record of the publications that have *referred* to the case, the other court decisions that have affected the case, and any other references that may be of value to the lawyer. This type of listing is particularly important to the lawyer, because, in law, much is based on precedent.

Citation indexes depend on a simple system of coding entries, one that requires minimum space and facilitates the gathering together of a great volume of material. However, a code is not absolutely necessary if one chooses to compile a systematic listing of individual cases or reports, with a complete bibliographic history of each of them. Thus, it would be possible to list all pertinent references under each case with sufficient com-

Dr. Garfield is Director of the Institute for Scientific Information, 33 South 17th Street, Philadelphia 3, Pa.

pleteness to give the index more of the appearance of a bibliography. However, this would result in an extremely bulky volume.

There are analogies in bibliographic operations. For example, in cataloging books for booksellers' or library catalogs, an attempt is made to find references to each book in one or more authoritative bibliographic sources, such as the catalogs of the British Museum (BM), Bibliothèque Nationale (BN), or the Library of Congress (LC). The "authority" card used in cataloging sometimes looks like a Shepard entry.

Another example is a book-review digest, in which one finds for each book title a series of references and selections from published reviews, critical and otherwise. Certain indexing publications perform a similar function.

Some time ago I became concerned with the problem of developing a citation code for science. This was necessary for the efficient manipulation by mechanical devices of entries to scientific indexes. In the course of this research I developed a very simple system for identifying an individual scientific article that had appeared in the periodical press. The resulting numerical code consisted of two parts. The first part was a serial number, used instead of an abbreviation, to identify each periodical; it was similar to the serial numbers employed in the *World List of Scientific Periodicals*, by no means a new idea. For example, *Die Bibliographie der fremdsprachigen Zeitschriften Literatur* has for many years used such a system to save space.

The second part of the code number was also a serial number, assigned to each article in a particular publication, starting with 1 and continuing throughout all volumes. The code thus gives no indication of year or volume number, a serious shortcoming. The article number is also not unique, having been used by the *Proceedings of the Society for Experimental Biology and Medicine* since its inception. These two serial numbers taken together, it can be seen, can identify any published periodical article. It soon became apparent, after such codes had been utilized on an experimental basis, that the use of the codes would facilitate the compilation of a citation index. (Other coding systems would be equally applicable.)

A citation index to science would have the following main characteristics. First there would be a complete alphabetic listing of all periodicals covered, in addition to the code numbers for each periodical. The list would be similar to the *World List*, but without the library holdings information. The main portion of the citation index would list in straight numerical order the code numbers for all the articles covered. Under each code number, for example, 3001-6789, there would be listed other code numbers representing articles that had *referred* to the article in question, together with an indication of whether the citing source was an original article, review, abstract, review article, patent, or translation, and so forth. In effect, the system would provide a complete listing, for the publications covered, of all the original articles that had *referred* to the article in question. This would clearly be particularly useful in historical research, when one is trying to evaluate the significance of a particular work and its impact on the literature and thinking of the period. Such an "impact factor" may be much more indicative than an absolute count of the number of a scientist's publications, which was used by Lehman (*3*) and Dennis (*4*). The "impact factor" is similar to the quantitative measure obtained by Gross (*5*), in evaluating the relative importance of scientific journals, a method later criticized by Brodman (*6*) but used again by Fussler (*7*).

Other advantages would also obtain. In a way such listings would provide each scientist with an individual clipping service. By referring to the listings for his article, an author could readily determine which other scientists were making reference to his work, thus increasing communication possibilities between scientists. It is also possible that the individual scientist thus might become aware of implications in his studies that he was not aware of before.

Most authors like to see how their works are received. Bringing together all book reviews and abstracts is very important, for it is not possible for an author to keep up with the thousands of publications in which his contribution might be reviewed. This applies equally to publishers. It would not be impossible to include books in the citation index. Indeed, as a first suggestion, the use of Library of Congress card numbers as the identifying code for books would seem appropriate.

It is necessary next to discuss some realistic questions concerned with the realization of such an index. Bitner (*8*) has estimated that 30,000 cases are covered by *Shepard's Citations* in 1 year, the cases and articles appearing in not more than a few hundred publications. In 1953 about 1 million citations were added—close to 40 citations per case.

What is the prospect in scientific literature? The last published edition of the *World List of Scientific Periodicals* contained more than 50,000 titles in science and technology. It is variously estimated that between 1 and 3 million new scientific articles are published each year. The *Journal of the American Chemical Society* alone publishes more than 3000 per year, including approximately 2000 original articles. The order of magnitude is therefore potentially from 50 to 100 times as great as it is for *Shepard's Citations*.

However, not all of these 50,000 publications are being covered in our present indexing activities, and yet this has not prevented us from continuing indexes of standard type or from starting new ones. Lack of complete coverage is not necessarily an argument against a citation index. It is in fact an argument in its favor. Coverage could perhaps be limited to the list of periodicals covered by one of the leading indexing services. This approach would, of course, have an immediate disadvantage. Such a *subject* selection would mean that less directly related subjects of interest would be excluded, and these are the publications that the individual is least likely to cover in his own research. It would be necessary to consider all the pros and cons in a selective approach and then to determine the possible utility of such a tool. For example, would a citation index to the 1500 periodicals covered by the *Current List of Medical Literature* be of real value, or, similarly, a citation index to the 5000 periodicals covered by *Chemical Abstracts?* The *Current List* would, in fact, offer a good starting point, since it already provides a unique code for the 100,000 items indexed by it each year. Presumably these are the most significant contributions in the covered fields for the year. If 10 is the number of references in the average article, then about 1 million citations would be involved. The preparation of that number annually is not unreasonable. Shepard's has already used well over 50 million citations in its publishing activities.

The ultimate success of a citation index would depend on many factors. For example, if each periodical would assign unique code numbers to the articles published, it would be possible for authors to list these numbers in their bibliographies and, thus, to save the work of coding on the part of the citation index staff. It is unlikely that such a development could take place in less than 5 or 10 years, but it is comparable to the problem of getting publishers to include Library of Congress card numbers in their publications.

When such a large volume of data is to be handled, mechanical devices of high speed and versatility could be used to great advantage and would probably determine success or failure. Once the coding is done, compilation itself is quite mechanical. This could be done by means of conventional filing slips; the Shepard organization itself has used them successfully for 80 years. However, it would be facilitated by a mechanical approach using punched cards.

The utility of a citation index in any field must also be considered from the point of view of the transmission of ideas. A thorough scientist cannot be satisfied merely with searching the literature through indexes and bibliographies if he is going to establish the history of an idea. He must obviously do a great deal of organized, as well as eclectic, reading. The latter is necessary because it is impossible for any one person (the indexer) to anticipate all the thought processes of a user. Conventional subject indexes are thereby limited in their attempt to provide an ideal key to the literature. The same may be said of classification schemes. In tracking down the origins of an idea, the citation index can be of real help. This is well illustrated by an example from my own experience.

Many years ago the Radio Corporation of America developed a reading-aid for the blind (*9*). This device had an electronic system for converting printed letters into recognizable sound patterns. Using the device, a blind man could scan a printed page; in a set of headphones he could hear a series of sound patterns, each letter having its own recognizable sound pattern. In effect, the words were spelled out, letter by letter, in code. I was particularly interested in this device because I had been independently working on a device that would *copy* print, letter by letter, and reproduce it for bibliographic and other purposes. The two devices had something in common in that they both employed scanning devices. I then wanted to learn whether anyone had ever suggested that the RCA reading-aid could be used for this purpose. It will be apparent that if anyone had known of the RCA device and had thought of adapting it for copying purposes, a reference to the article might have been made. This reference could easily have been included in an article or patent that was not at all related to the problem of reading devices. A citation index would have given me just what I was after. Nothing could substitute for extensive reading, but a great deal of time could have been saved by bringing the appropriate works to my attention.

In the course of my reading I did find a few references to this device, one in a book (*10*), and several others in periodical articles, one of which was a German article on the mechanization of philological analyses and concordance building. The latter article (*11*) did not discuss my own special interest in copying devices, but it did show the similarity between the author's and my own thinking from the point of view of letter-recognition devices, which is what the RCA device attempts to be. In other words, both of us were interested in this device as a letter-recognition device for the analysis of text.

In another instance the RCA article was unexpectedly cited in the journal *Electronic Engineering* in an article on information theory (*12*) that I was reading because of an entirely different interest. No subject indexer could have anticipated this crossbreeding of interests. Perhaps there are many other articles and books unknown to me that have made similar references to this device. How can they be located when the main subject matter of the article is, on the surface, so unrelated in nature?

One might say that it would be possible to index articles more thoroughly to achieve the same results. For example, the article on information theory, if thoroughly indexed, might have included an entry under reading devices for the blind. Yet if this were done, our periodical indexing services would clearly become hopelessly overloaded with material that is not necessary to lead us to the micro unit—the entire article or one of its major sections. Although it might be said that no scientist interested in the greater comprehensiveness to be found in a citation index would object to having such a great mass of references in a subject index, this is impracticable. It would require an army of indexers to read the articles and identify the exact subject matter of every paragraph or sentence. Yet this would be necessary. To illustrate, it is only in the very last paragraph of the article on information theory that one would find a reference to reading devices for the blind.

Were an army of indexers available, it is still doubtful that the proper subject indexing could be made. Over the years changes in terminology take place, that vitiate the usefulness of a standard subject index. To a certain extent, this is overcome through the citation approach, for the author who has made reference to a paper 40 or 50 years old has interpreted the terminology for us. By using authors' references in compiling the citation index, we are in reality utilizing an army of indexers, for every time an author makes a reference he is in effect indexing that work from his point of view. This is especially true of review articles where each statement, with the following reference, resembles an index entry, superimposed upon which is the function of critical appraisal and interpretation. To the indexer this has its advantages as well as its disadvantages (*13*).

To determine in a practical way what the citation index could offer, it was decided to track down the citations made in one journal to a single significant article, in order to compile a sample entry for the citation index. At the suggestion of Erich Meyerhoff, I selected Hans Selye's famous article on the general adaptation syndrome (*14*). A systematic search was then made of all papers that were published in the *Journal of Clinical Endocrinology* subsequent to Selye's paper up to 1951—a period of 5 years, including well over 500 articles. Every bibliography in each of the 500 articles was checked for a reference to Selye's article. Twenty-three articles were found to make such reference; each of them was then checked for the character of the information provided.

Examination of the citation list (Table 1) shows the great variety of subject matter included. One thing became quite clear, even to the uninitiated—that is, the influence of Selye's article has been quite pronounced. Such evidence is extremely valuable to the historian.

It is interesting to note that, although all the articles cited were indexed in *Quarterly Cumulative Index Medicus*, not one is to be found there under the heading "Adaptation." In fact, it is surprising not to find any articles from this journal under this subject heading.

It also becomes quite obvious that many references to Selye's paper were general and contribute little or nothing

Table 1. Index sample based on article by Hans Selye, "General adaptation syndrome" [*J. Clin. Endocrinol.* 6, 117 (1946)]. The code number for this journal in the *World List* is 11,123a; the article number is arbitrarily taken as 687; and the code number for the article is 11123a-687. The 23 articles that cited Selye's article are listed, followed by *A hypothetical citation index entry for Selye's article: R*, review article; *A*, abstract; *O*, original article.

1. Williams, R. H.: Thyroid & Adrenal Interrelations, 7: 52–57 (1947).
2. Venning, E. H.: Glycogenic Corticoids, 7: 79–101 (1947).
3. Forbes, *et al.*: 17-Ketosteroids in Trauma and Disease, 7: 264–288 (1947).
4. Talbot, *et al.*: Excretion of 11-Oxycorticosteroids, 7: 331–350 (1947).
5. Castillo, E. B. del, *et al.*: Syndrome of Rudimentary Ovaries, 7: 385–422 (1947).
6. Forsham, P. H., *et al.*: Pituitary Adrenocorticotropin, 8: 15–66 (1948).
7. Pincus, G., *et al.*: Rhythm in Biped Excretion, 8: 221–226 (1948).
8. LeCompte, P. M.: Width of Adrenal Cortex in Lymphatic Leukemia, 9: 158–162 (1949).
9. Wolfson, W. Q.: 17-Ketosteroids in Gout, 9: 497–513 (1949).
10. Stein, H. J., *et al.*: Hormonal Response to Heat and Cold, 9: 529–547 (1949).
11. Davis, M. E.: Eosinophils in Pregnancy and Labor, 9: 714–724 (1949).
12. Conn, J. W.: Na and Cl of Sweat as Cortical Index, 10: 12–23 (1950).

13. Recant, L., *et al.*: Effect of Epinephrine on Eosinophils, 10: 187–229 (1950).
14. McArthur, J. W., *et al.*: Urinary Excretion of Corticosteroids in Diabetic Acidosis, 10: 307–312 (1950).
15. Bors, E.: Fertility in Paraplegic Males, 10: 381–398 (1950).
16. Grossman, S., *et al.*: Idiopathic Lactation following thoracoplasty, 10: 729–734 (1950).
17. Cooper, J. S., *et al.*: Metabolic Consequences of Spinal Cord Injury, 10: 858–870 (1950).
18. Hioco, D.: Adrenal Metabolites in Bronchial Asthma, 10: 1570–1578 (1950).
19. Jailer, J. W.: Pituitary-Adrenal System in Infants, 11: 186–192 (1951).
20. Deane, H. W.: The Adrenals in Experimental Hypertension, 11: 193–208 (1951).
21. Hioco, D., *et al.*: Epinephrine and ACTH in Bronchial Asthma, 11: 395–407 (1951).
22. Schaffenberg, C. A., *et al.*: p-Hydroxypropiophenone (PHP) and other so-called pituitary inhibitors, 11: 1215–1223 (1951).
23. Talbot, N. B., *et al.*: Urinary Water-Soluble Corticosteroids, 11: 1223–1236 (1951).

Citation Index Entry

11123s-687

464-9789(R)
869-3366(R)
1105-9876(A)
1123-4432(R)
a11,123-0752(O)
-0779(O)
-7264(O)
-7331(O)
-7385(O)
-0866(O)
-8221(O)
-9158(O)
-9497(O)
-9529(O)

to the readers' enlightenment, since exact page references are not provided. In several cases the Selye article is even cited but not referred to in the text. Selye's influence on all of these authors is quite apparent. In particular instances the citations are of value in locating confirmatory evidence of some of Selye's claims.

Thus, in the case of a highly significant article, the citation index has a quantitative value, for it may help the historian to measure the influence of the article—that is, its "impact factor." With regard to a less significant work, one would suspect that the bibliographic advantages might be increased, because the scientist or librarian would be provided with references not to be found in conventional indexes. The preliminary evidence presented indicates that the citation index offers interesting possibilities for another approach to bibliographic control.

The next step in compiling the index for the Selye article would be to seek out additional references to it in more peripheral journals, but obviously the farther away you get from the immediate subject area of the main article, the fewer the references to it you will locate. Yet these may well be the most useful references of all, for the cross-fertilization of subject fields is one of our most important problems in science literature.

It will be well to close with a brief description of how the citation index might be compiled. The first step would be the selection of the particular group of periodicals to be covered; next, the period to be covered, say, only that since 1900.

The problem actually has two facets: the selection of periodicals to be covered in order to obtain citations, and the selection of those articles for which we want a citation record. For example, all *articles* in journals in the *Current List of Medical Literature* that have remained in continuous publication since 1900 might be coded, in which case the *Journal of Clinical Endocrinology* would not be included. However, we might include as citation sources all *journals* covered by the *Current List*. Thus, the bibliographies appearing in articles in the *Journal of Clinical Endocrinology* would supply references to the basic group of articles.

Each coder would be assigned a group of articles in a particular journal. The first step would be to number each article in the journal in ascending order, by utilizing a complete table of contents of that journal from its inception.

Once a code number has been assigned to each article, the proper codes may then be assigned to each periodical. This might be the number given in the *World List*, with new numbers for any periodicals not to be found there.

Actual coding starts with the first article in a particular periodical. The coder prepares a 3- by 5-in. card for each citation made in the article. Each card should give (i) the code number for the citing article, (ii) the code number for the article cited, and (iii) a classification of the citing article as an original contribution, review article, abstract, and so forth.

Many references will be excluded by the limits of coverage set up. Thus all references to articles not in the prescribed list of journals would be excluded.

All books would be excluded unless otherwise specified, in which case the reference card would carry the code for the citing article and the code for the book (its LC card number).

After all the articles had been coded, it would next be necessary to sort the cards by the code numbers for the items cited. This would yield a group of cards for each cited article. These would then be sorted by code numbers for the citing articles. This completes the coding and sorting. The next step would be preparation for the printer.

From this description it will be apparent that, although a great volume of material is to be covered, relatively unskilled persons can perform the necessary coding and filing. Professional supervision would still be required, because certain decisions require skilled judgment, for example, when *ibid.* or *loc. cit.* must be carefully interpreted. Footnotes tend to make coding somewhat cumbersome. The code I have described is merely an example used to illustrate the method in principle. If the system were adopted, then in the future every author ought to be required to include the serial number of each item he referred to, so as to facilitate not only the compilation of citation indexes but also other operations such as requests for reprints (*15, 16*).

In a certain sense a citation index is not very different from a compendium like *Beilstein*, which gives a rather complete record of a compound, compiled by a similar method. A citation index for the literature of chemistry would undoubtedly make the preparation of such works as *Beilstein* much easier than it is at present. The new bibliographic tool, like others that already exist, is just a starting point in literature research. It will help in many ways, but one should not expect it to solve all our problems.

References and Notes

1. P. Thomasson and J. C. Stanley, *Science* 121, 610 (1955). Thomasson and Stanley were commenting on C. Zirkle's discussion of the use of fraudulent data [*Science* 120, 189 (1954)].
2. W. C. Adair, *Am. Documentation* 6, 31 (1955).
3. H. C. Lehman, *Sci. Monthly* 78, 321 (1954).
4. W. Dennis, *ibid.* 79, 180 (1954).
5. P. L. K. Gross and E. M. Gross, *Science* 66, 385 (1927).
6. E. Brodman, *Med. Library Assoc. Bull.* 32, 479 (1944).
7. H. H. Fussler, *Library Quart.* 19, 19 (1949).
8. H. Bitner, personal communication, April 1954.
9. V. K. Zworykin and L. E. Flory, *Proc. Am. Phil. Soc.* 91, 139 (1947).
10. R. R. Shaw, *Machines and the Bibliographical Problems of the Twentieth Century* (Univ. of Illinois Press, Urbana, 1951) p. 19. (Reprinted from *Bibliography in an Age of Science* (Univ. of Illinois Press, Urbana, 1951).
11. R. Busa, *Nachr. Dokumentation* 3, 14 (1952).
12. A. M. Andrew, *Electronic Eng.* 25, 471 (1953).
13. E. Garfield, "Review literature as a source of critical entries for scientific indexes, unpublished paper, December 1952.
14. H. Selye, *J. Clin. Endocrinol.* 6, 117 (1946).
15. E. Garfield, *Science* 120, 1039 (1954).
16. J. A. Behnke, *Science* 120, 1055 (1954).