

BIBLIOGRAPHIC CITATIONS IN BIOCHEMISTRY: A Chronological Analysis

M A Gopinath, *Documentation Research and Training Centre, Indian Statistical Institute, 31 Church Street, Bangalore 560 003.*

Presents an analysis of references given at the end of the technical papers in *Journal of Physiology*, *Experimental Cell Research*, *Journal of Comparative Neurology*, and *Journal of Molecular Biology*. Sample articles were chosen for subject coterminous with each other for three decades 1950, 1960, and 1970. The analysis resulted in ranking of journals, increase or decrease of reference articles, variety of documents cited and age of references. It was found that journals have greater self-citations and journals were the most frequently cited source of information. It was also found that older documents have fairly good impact values in the field of Biochemistry. The data elements in the bibliographic citations remain same over the period 1950's to 1970's. The secondary information services have been identified as important sources for finding references and verifying them.

1. Introduction

Bibliographic citations or references provided at the end of a technical papers in periodical publication have attracted attention of information scientists from the point of assessment of use of already published literature. The frequency of use of earlier published documents in a subject-field and its currency can be studied by such citations. Further, one can also assess the source of citations and accessibility to bibliographic services in different fields of knowledge. In this paper, we shall discuss some of the features of the citations to technical papers in the related fields of Biochemistry. Papers published during 1950's, 1960's and 1970's were selected and the citations data were collected for these articles.

The study covered the articles published in the following periodical publications :

- 1 Journal of Physiology (for Cardiology)
- 2 Experimental Cell Research (for Cell)
- 3 Journal of Comparative Neurology (for Neurology)
- 4 Journal Molecular Biology (for Protein)

Sample articles on a given specific subject for example "Cardiology" in a particular journal for example in the *Journal of Physiology* for three decades 1950s, 1960s, and 1970s were analysed to find the following factors :

- 1 A ranked list of periodical publications cited and also the self-citations rate of that periodical publications along with percentage of use;
- 2 Increase or decrease in reference per article cited over the three decades;
- 3 PI value (percentage of 1 to 5 year old references). An average of percentage of 1 to 5 year old references over three decades to study the use of recent literature in the field;
- 4 Comparative study of data elements cited during 1950's, 1960's and 1970's to notice changes if any;
- 5 To know the various types of documents cited and also the average number of increase or decrease of each type of documents cited.

These studies will help in user studies, to study communication patterns, in particular in evaluative Bibliometrics. These studies can be used to the various interests of authors, their field of research and their work of interest etc.

Our study here is mainly concentrated on the source of information and the impact of these sources on the habit of citing by the biochemists. The data obtained on sample articles for three decades (1950's, 1960's and 1970's) for the four subjects in the field of Biochemistry, namely physiological aspects of Cardiology, experimental studies on the Cell Biochemistry, comparative studies of Neurobiochemistry and the Molecular Biochemistry of protein were

studied. The paragraphs of this report presents the details of the study.

2 Analysis of citations on Cardiology in the *Journal of Physiology*

A total of 610 citations on Cardiology in the *Journal of Physiology* from sample articles chosen over a period of three decades (1950's, 1960's and 1970's) were analysed. These were obtained from 20 sources articles. Various factors such as citation counts, PI values and elements cited were made.

A ranked list of periodical publications during the given period showed that self-citation rate was higher—that is out of 610 citations, 122 citations were selected from the same journal, a percentage of 20.3. The other citations were scattered among different periodicals shown in the table below :

TABLE 1: Ranked list of journals cited in *Journal of Physiology* on "Cardiology"

Rank	Title of the periodical	Total number of citations	Percentage of Total use
1	<i>Journal of Physiology</i>	122	20.3
2	<i>American Journal of Physiology</i>	49	8.1
3	<i>Circulation Research</i>	26	4.3
4	<i>Journal Biological Chemistry</i>	24	4.0
5	<i>Biochemistry Journal</i>	18	3.0
6	<i>Journal Pharmaceutical Experiments</i>	9	1.5
7	<i>Physiological Reviews</i>	8	1.3
8	<i>Nature</i>	8	1.3
9	<i>Experientia</i>	7	1.1
10	<i>American Heart Journal</i>	6	1.0

A study on the average number of citations per article showed that there was a steady increase in number during the given period. As shown in table below it was 13 numbers in 1950's, 24 in 1960's, and 31 during 1970's. This can be attributed to the two main factors :

(1) An exponential growth in scientific information.

(2) Access to information, that is primary literature with the help of secondary or tertiary sources like Abstracting and Indexing periodicals, and computer-based information systems and services.

These two factors have alarmed the scientists/researchers about the need for information and also the information scientists about the techniques of processing and dissemination of information.

(3) In the field of Science and Technology the vast information produced has to be transferred and utilised as early as possible. It is believed that the literature after five years of its publication becomes obsolete. Derek de Solla Price (7, 8) has devised PI values which indicate the immediacy of use of information. That is given as the percentage of 1 to 5 year old references in articles. As our study is concentrated on the use of latest information and its sources it was decided to calculate the average PI value of articles over the given period. The table 2 shows the variation in PI values in three decades.

TABLE 2: The R/A (References per article) and PI (% of 1 to 5 yrs. old ref) values of citations on cardiology in the *Journal of Physiology*

Values	1950's		1960's		1970's	
1 R/A (Reference per Article)	13		24		31	
2 PI (% of 1 to 5 year old references)	Younger	Older	Younger	Older	Younger	Older
	55.3	44.6	44.5	55.4	41.4	58.6

The data shows that there is an increase in the volume of citations. But the PI values indicate that the citation which are older than 5 years are on the increase in 1960's and 1970's. (55.4 in 1960's and 58.6 in 1970's). That means that citation impact value of the older documents is on the increase.

(4) A comparative study of data elements in citations showed no difference over the period and typical citation cited the following details—that is, author, name of the journal, volume number,

pagination and the year. All the citations were arranged by the names of the authors alphabetically.

(5) With this analysis, it was possible to locate various types of documents cited like periodicals, books, conference proceedings etc. The analysis on Cardiology showed that periodicals were cited more times than any other types of documents. The following table shows the results of analysis.

TABLE 3: Varieties of documents

Types of documents (Average number of per article)	1950's	1960's	1970's
1 Periodicals	12	21	41
2 Books	0	1	2
3 Proceedings	0	1	1

It is interesting to note that there was an increase in the number of journals cited that 12, 21 and 41 from 1950's to 1970's.

3 Analysis of Citations on Cell in the *Experimental Cell Research*

The total number of citations analysed is 400 for the field of Cell Biochemistry published in the periodical "Experimental Cell

TABLE 4: Rank list of Journals cited on 'Cell in the *Experimental Cell Research*

Rank	Title of the periodicals	Total Number of citations	Percentage of Total use
1	Experimental Cell Research	38	8.7
2	Nature	21	4.8
3	Cancer Research	20	4.5
4	Science	17	3.9
5	Journal of Biological Chemistry	8	1.8
6	Journal of Physiology	6	1.3
7	Biochemica et Biophysica Acta	4	0.9
8	British Journal of Cancer	4	0.9
9	Development Biology	4	0.9
10	Anatomical Record	3	0.65

Research" for the sample article chosen over the period 1950 to 1970. These were obtained from 20 source articles. The number of citations to the same periodical is 38. The percentage of total use of the same periodical is 8.7. A ranked list of periodicals during the given period shows that self-cited periodicals were higher than any other periodicals.

(2) A study on the average number of citations per article showed an increase in number during 1970's as shown in the following table.

(3) The PI values calculated for three decades are tabulated as follows:

TABLE 5: The R/A and PI values for citations on Cell Biochemistry in the Experimental Cell Research

Values	1950's		1960's		1970's	
R/A	19		19		32.6	
PI	Younger	Older	Younger	Older	Younger	Older
	42.9	57.01	74.5	25.5	44.4	55.5

The PI values indicate a growth in impact of older citations during 1950's and 1970's. The 1960's however showed high impact of references for younger documents than the older ones.

(4) The comparative study of data elements showed no differences.

(5) The various types of documents cited over the period in order of preferences is given in the following table:

TABLE 6: Variety of Documents

SN	Types of documents	1950's	1960's	1970's
1	Periodicals	14.8	1.5	21.6
2	Books	2.0	2.5	3.8
3	Conference Proceedings	4.2	2.6	2.0
4	Thesis	—	—	4.0

The data indicates that the periodical publications and books are on increase in citation counts. Thesis have been added to citations in 1970's. But the citations for conference proceedings is on decrease. It may be stated that the publications of periodicals and books are on the increase and easily accessible and hence they have been cited more often. It also appears that the dissemination of thesis literature is better today than earlier decades. The decline in conference proceedings may be attributed to the fact that many of them get published in more formal documents such as periodical publications and books.

4 Analysis of Citations on Neurology in the *Journal of Comparative Neurology*

The total number of citations analysed in the field of Neurological Biochemistry is 544. Of these 94 citations were for the same journal, which amounts to 17.28 percent of the total.

A ranked list of journals in Neurological Biochemistry showed that the self-citation rate was higher and concentrated. The other citations were diffused in nine other journals in the field.

TABLE 7: Rank List of Neurological Biochemistry

Rank	Title	No. of citations	% of total use
1	Journal of Comparative Neurology	94	17.28
2	Brain Research	29	5.33
3	Experimental Brain Research	23	4.23
4	Journal of Cell Biology	20	3.68
5	Journal of Neurological Physiology	16	2.94
6	Journal of Physiology	15	2.96
7	Brain	9	1.65
8	Archives of Neurology	9	1.65
9	Journal of Anatomy	8	1.47
10	Review of Neurology	7	1.29

Citation counting—Average number of citations per article over the period showed an increase in the number from 1950's to 1970's as shown in the following table:

TABLE 8: R/A and PI values for Neurological Biochemistry

Values	1950's		1960's		1970's	
R/A	21		37		39	
PI	Younger 37.8%	Older 51%	Younger 38.7%	Older 58%	Younger 39.9%	Older 60%

The data indicates that the impact values for the older documents is fairly high in Neurological Biochemistry. The comparison of data elements showed no difference at all and the data elements included in bibliographical citation is as follows: Serial number, name(s) of author(s), name of the periodical, volume number, pagination and the year of publication. The count of the various types of documents cited in the sample is shown in the following table:

TABLE 9: Variety of Documents cited in Neurological Biochemistry

Sl No	Types of documents	(Average Number per Article)		
		1950's	1960's	1970's
1	Periodicals	18	26	29
2	Books	2	5	6
3	Conference Proceedings	1	1.5	1.75
4	Letters	1	-	1

The table indicates a steady increase in the number of documents cited per article in relation to periodical publications, books and conference proceedings. It is interesting to note that the citation count indicates the presence of letters in 1950's and 1970's and its absence in 1960's.

5 Analysis of Citations on Protein Biochemistry in the *Journal of Molecular Biology*

The total number of citations analysed for the sample articles chosen for Protein Biochemistry from the *Journal of Molecular Biology* for the decades 1950's, 1960's and 1970's is 761. Out of these, the

citations for the same periodical that is Journal of Molecular Biology is 185. This amounts 24.31 percent.

A ranked list of the periodicals is given in the following table:

TABLE 10: Ranked list of periodicals in Protein Biochemistry

Rank	Title of periodicals	Number of citations	% of the total uses
1	Journal of Molecular Biology	185	24.31
2	Virology	71	9.33
3	Nature	49	6.44
4	Biochemica and Biophysica Acta	35	4.60
5	Journal of Biological Chemistry	34	4.47
6	Biochemical and Biophysical Research	24	3.15
7	Journal of Bacteriology	23	3.02
8	Science	14	1.84
9	Annales of Institute of Pasteur	13	1.71
10	European Journal of Biochemistry	13	1.71

The data indicates that the citations of articles published in the periodicals in the field of Biochemistry and Microbiology is fairly significant. The Journal of Molecular Biology shows a high rate of citations.

Citation Counting: The citation counting indicating reference per article (R/A) and percentage of impact (PI) over 1 to 5 years in the field of Protein Biochemistry is given in the following table:

TABLE 11: R/A and PI values for Citation in Protein Biochemistry

Values	1950's		1960's		1970's	
R/A	32		26.8		40.9	
	Younger	Older	Younger	Older	Younger	Older
PI	73.8	26.92	76.81	23.18	53.41	46.59

The references per article has shown an increase in the 1970's while it was on decline in the 1960's. The PI values indicates that the

impact values for newer citations is higher than that for older ones in all the three decades.

Data elements: A comparative study of the data elements included in a bibliographic citation indicated no difference over the three decades.

Variety of documents: The various types of documents cited per article in a periodical provided the following data:

TABLE 12: Variety of Documents in Protein Biochemistry

Sl No	Types of documents	(Average Number per Article)		
		1950's	1960's	1970's
1	Periodicals	25	32	37
2	Books	3	8	9
3	Conference proceedings	2	3	5
4	Letters	1	1	4

The data shows the increase in the numbers citations for all the varieties of documents.

6 Overall Conclusions on Citation Analysis

It may be seen that there is a steady increase of reference per article in all the four subjects during 1950's to 1970's. It may also be observed that the citations for articles for periodicals in the same periodical is higher than the other periodicals. It is further found that periodical publications have greater number of citations among the variety of documents cited with books, conference proceedings and letters coming next in order. The P1 value (that is the percentage of 1 to 5 year old reference in articles) indicate that the older documents have fairly good impact values in the field of Biochemistry. The data element in the bibliographic citations over the period 1950's to 1970's appears to be same.

61 Study on the Source of Citations

The Ortega hypotheses (1) states that "Citations generally represent an authentic indicator of influence" and the norms of science

require scientists to cite the work that they have found useful in pursuing their own research and for the most part they abide by these norms." Therefore, a survey was made among a group of biochemists who were attending the international Biochemistry conference. A prepared questionnaire was used for the interview. About 50 biochemists were interviewed. An analysis of the results showed that

- (1) Out of 50 biochemists 47 scanned the periodical publications for study and consequent citations.
- (2) Out of 50 biochemists 43 used Abstracting and Indexing periodicals as important source for the citations.
- (3) The following table indicates the number of biochemists using various types of citation sources in decreasing order.

TABLE 13: Variety of Documentary Sources for Citation

Sl No	Name of the source	Number of Biochemists out of 50	Percentage of use
1	Primary periodicals	49	24.5
2	Abstracting and Indexing periodicals	40	20.0
3	Correspondence and Preprints	39	19.5
4	Seminars and Conferences	15	7.5
5	Machine Readable data base	1	.5

- (4) Biochemists were given a list of important bibliographic services and were asked to rate each one of them so as to know their importance. The rating showed the following results.

TABLE 14: Variety of Secondary Sources for Citation

Sl No	Name of the Source	Cannot live without it	Useful	Consult occasionally	Do not consult at all
1	Biological abstracts	15	8	3	3
2	Chemical abstracts	10	10	4	2
3	Commonwealth Agricultural Bureau Abstracts	—	4	1	13
4	Current contents	19	6	4	—
5	Physics abstracts	1	—	1	13
6	International Research Abstracts	1	—	1	13

- (5) The population of Biochemistry authors were grouped into Professors, teachers (Readers and Lecturers) and Research fellows:

	Professors	Teachers	Research Fellows
Periodicals	20 scan out of 21	17 scan out of 18	12 scan out of 12
Secondary	21 scan out of 21	18 scan out of 18	12 scan out of 12

- (6) Apart from books and periodicals the other used by the biochemists as source for citations are as follows:

TABLE 15: Non-Book Sources by Biochemists

Sl No.	Non-Book Sources	Cannot live without it	Useful	Useful occasionally
1	Preprints and correspondence	20	10	2
2	Discussions	22	11	6
3	Libraries	23	7	4
4	On-line services	8	6	1
5	Conferences	3	3	1

It may be seen that the correspondence, discussions and the libraries are still the dominant media for citations selection. On-line bibliographic services are yet to pick up in rating.

- (7) Age and experience has a great effect on the selection citation sources. To study this the population under observation has been divided into two groups. Those with less than 10 years experience are grouped in Group I and above 10 years experience in Group II.

These were analysed into two sets. The first was for the rating of secondary periodicals.

TABLE 16: Secondary Sources used by Biochemists

Sl No	Source	Group I: Below 10 years Experience			Group II: Above 10 years Experience		
		(1)	(2)	(3)	(1)	(2)	(3)
		1	Biological Abstracts	12	4	1	10
2	Chemical Abstracts	12	4	2	9	5	1
3	Commonwealth Agricultural Bureau	—	2	—	—	—	—
4	Current Contents	19	4	3	18	4	3
5	Physics Abstracts	—	1	1	1	1	1
6	Information Research Abstracts	1	—	1	1	1	1
7	Science Citation Index	3	3	3	4	3	3

The age and experience does not indicate much difference in rating the different abstracting and indexing periodicals used as source of citations.

The use of different sources of information apart from books and periodicals in these two groups were analysed.

TABLE 17: Non-Documentary Sources by Biochemists

Sl No	Source	Group I: Below 10 years Experience			Group II: Above 10 years Experience		
		(1)	(2)	(3)	(1)	(2)	(3)
		1	Preprints and correspondence	6	4	—	8
2	Discussions	10	2	8	11	5	4
3	Libraries	10	1	1	8	3	3
4	On-line services	6	6	2	8	7	2
5	Conferences	6	2	2	6	3	3

Older scientists get citation information from discussions while younger persons appear to use libraries more in addition to discussions. On-line services appear to have fairly good response from both the groups.

- (8) As part of the survey the responding biochemists were asked to comment on the installation of computer-based information services in their respective libraries. Some of the responses are as follows:
- 1 Computerisation is not needed until the libraries are well-balanced in stock and better organised.
 - 2 Twenty persons out of 50 think that computer-based bibliographic sources provides better facility for accessing information and hence stimulate more volume of citations.
 - 3 Fifteen scientists out of 50 definitely suggests the immediate need for computer-based bibliographical sources in case the software and hardware are easily available freely to all users and if the database concerned is a comprehensive era.
 - 4 Ten scientists out of 50 suggest that on-line bibliographic data base are very useful for citation purposes and they must be established immediately.
 - 5 Five scientists out of 50 stated that MEDLARS on-line as a very useful source of information for citation purposes.

In another survey, it was found that biochemists primarily use the abstracting and indexing periodicals for identifying documents for citation purposes. The following table presents details in this regard.

The biochemists listed the following secondary information services percentages indicate the order of preference.

TABLE 18: Secondary Source of Information for Biochemists

Biological abstracts	100%
Chemical abstracts	22.22%
Physics abstracts	—
Current contents	33.33%
Analytical abstracts	—
Biochemical abstracts	63.33%
Review of Plant pathology	44.44%
Bio-research indexes	11.11%

Majority of biochemists use *Biochemical abstracts* and *Biological abstracts*. The latter has a maximum percentage of use because biochemists search for application areas in a broad spectrum of subjects. The same argument holds for a fairly good amount of incidence of use in relation to review of plant pathology by the biochemists.

The following table presents a priority of purposes for scanning of a secondary bibliographic services.

TABLE 19: Purpose of the Use of Bibliographic Sources

Purpose	Percentage of Incidence
To keep up-to-date with latest development	77.77
To make retrospective search	66.66
To get references	88.88
To know literature published in Non-English Language	22.22

Majority of biochemists refer the abstracting periodicals to select references to documents. Abstracting periodicals can help to overcome the language barrier to certain extent, by providing abstracts in English of documents in different international languages. In such cases, the informative abstracts are preferred. The approaches to a secondary information service by the biochemists is indicated by the following table:

TABLE 20: Purpose of the Use of Indexes

Author Index	33.33%
Subject Index	88.88%
Generic Index	22.22%
Concept Index	11.11%
Keyword Index	23.5%
Other Indexes	9.83%

Subject index has the highest preference, because it provides for browsing facility. The author index is usually known as author search. Probably, the impact is of invisible colleges or other media of familiarity. The fact that the biochemists peruse author index sug-

gests also they use them for getting references and in particular verification of references for citation purposes.

In summary, the preferred sources of citations by biochemists are almost even between the documentary and non-documentary sources. In documentary, secondary information services have better impact than the primary sources. Here again, the conventional sources in print on paper have better impact than the machine-readable databases.

7 Overall Summary

It may be stated that the report identifies that there is increasing number of references per article in the field of Biochemistry from 1950's to 1970's. The references per article indicate that there are greater variety in the type of documents being cited. The periodical articles have the largest number of references per article. The next in order is books, conference proceedings and letters. The citations to articles published in the periodical hosting the source document (a variety of self-citation to host periodical) are predominant in this survey. The impact value (PI) for older papers is fairly large in the sample studied, although in some areas the newer publications have higher impact value. The secondary information services have been identified as important sources for finding references and verifying them. While the machine-readable databases are yet to find preference over print on paper databases, one can find that there is an urge among biochemical researchers to have on-line access to databases. More user-training and subsidising the cost of on-line search may provide better facility for researchers. The increasing volume and variety of literature published in sciences calls for a better access than the conventional sources.

8 References

- 1 COLE (JR) and COLE (S). The Ortega hypothesis. (*Science*. 178; 1972; 368-78).
- 2 GARFIELD (E). Citation index for studying science. (*Nature*. 227; 669-71).
- 3 _____. Trends in biochemical literature. (*Trends in Biochemical Science*. 4; 1959; N 290-N 295).

- 4 MARTON (J). Changes in the time distribution of biochemical article references from 1962 to 1977. (*Scientometrics*. 3; 1981; 397-400).
- 5 MEADOWS (AJ). The citation characteristics of astronomical literature. (*Journal of documentation*. 23; 1967; 28-33).
- 6 MORAVCSIK (MJ) and MURUGESAN (P). Some results on the function and quality of citations. (*Social Studies of Science*. 5; 1975; 86-91).
- 7 PRICE (DJ de Solla). Citation measures of hard science, soft science, technology, non-science. (*In Nelson (CE) and Pollock COK, Eds Communication among scientists and engineers*. 1970. Lexington. Mass. Heath. p. 9-10).
- 8 PRICE (DJ de Solla). Networks of scientific papers. (*Science*. 149; 1965; 510-515).