Chapter II

Review of Literature
CHAPTER II

REVIEW OF LITERATURE

Review of literature is a significant part of any research study. An investigator acquires information about what has been done in a field of study, to arrive at a fruitful conclusion. A review presents a vivid picture about previous researches in the area in which he intends to study. Familiarity with available literature in the area of research is also required by the investigator for making new grounds besides the proper designing of the study. Review of related studies further avoids duplication of work that has already been done in that area. It also helps the investigator to study the different aspects of the problem. It enables the investigator to identify the research gaps or the unexplored areas if any, in order to create new grounds for research. By considering the efficiency of various dimensions of Bibliometric studies, the researcher has classified them into six categories namely

1. General Bibliometric studies
2. Literature growth oriented studies
3. Language of Contribution studies
4. Country of Publication
5. Pattern of Authorship and Literature output

A. Authorship Pattern  
B. Co-Authorship  
C. Collaborative Authorship

6. Bibliometric Analysis of Bibliometric Literature

A review of previous studies in each of the above categories was attempted and presented in this chapter.

2.1 Bibliography - A Knowledge Organisation Tool

Bibliography as a generic term is an exhaustive list of reading materials contributed by an author, or on a particular person or in a micro subject. It has come a long way in the history of libraries and knowledge organisation. Today, published and priced bibliographies have taken printed, electronic and online forms which provide a bibliographic control mechanism to library professionals as well as scholars. Bibliographic information apparatus is a wonderful tool which, when analysed, reveals a spectrum of results on the behaviour of literature. Bibliography is not a librarian’s tool exclusively. Information on scholarly publications in the form of bibliography is necessarily found appended at the end of each article and that gives the clue to a scholar for the furtherance of his/her information journey. Bibliography becomes a scholars instrument unfolding the treasure of literature to his community amidst the chaos and confusion arising out of information explosion.
The impact of Information explosion has been great and a few decades ago, it was widely felt among the librarians that the profession was at crossroads not knowing the direction to proceed to find solutions to the barriers to scientific communication. Any effort to provide a bibliography to a client became futile due to the exorbitant cost and widely scattered nature of information. Manually prepared bibliography provided the mechanism, to some extent, to have a control over the literature output with a smaller amount of effectiveness.

The advent of computers into libraries changed the scene. It did not take much for the librarian to realise that the solution was very much with the electronic publications rather than the printed hardcopies of reading materials. And also the bibliographic details found in the citations give a leading clue for the study of the personality in the subject field and problems of identification also. Librarians began to look into the bibliographic compilations for analysis in order to have an effective bibliographic control. It started with citations and later encompassed bibliographies in large measures.

Citation analysis retains the identity of individual article in the context of its citations included in its bibliography. While citation analysis is restricted to each article and the bibliography of citations appended to it, the analysis of bibliographic records in their cumulations is known as Bibliometrics. Bibliometric studies reveal a wide range of results leading to inferences regarding even the future Nobel laureates. Such
studies dealing in depth of bibliographic apparatus led to the evolution of the field called "Bibliometrics".

2.2 Bibliometrics: Scope

Bibliometrics is the technical name for a range of analytical methods using information—information found organized in bibliographic description of reading materials such as books, periodical articles, reports, patents, software, designs, prototypes, and blueprints to develop descriptive statistics, multidimensional analyses, and graphical representations of the output of science. Bibliometrics is often used to clarify and assist in the analysis and formulation of science policy by

- highlighting the networks of players or subjects that make up scientific research;
- providing strategic analysis of the relative position of research performers;
- sketching profiles of the activities and performance of individual centers;
- graphically presenting studies of strategic or innovative subjects.
- publication counts, involving counting of scientific publications published by a researcher or a research group;
- citation counts, involving identifying the number of times a specific article is cited in other scientific journal publications; these measures are considered to address questions of quality, influence, and the transfer of knowledge;
co-citation analysis, identifying pairs or groups of articles that are cited together in other articles or publications; from these pairs or groups of articles a 'cognitive structure' can be derived, providing information on the direction and flow of scientific thought;

c-word analysis, involving assigning keywords to a paper or article by a professional reader; papers which have the same keywords and sets of words are linked to each other via a clustering technique;

scientific mapping, involving developing a visual model or 'map' of the realm of scientific fields representing the structure of literature output of particular scientific fields.

Bibliometric studies fall mainly into two broad groups – those describing the characteristics or features of literature (descriptive studies), and those examining the relationships formed between components of literature (behavioural studies, sometimes referred to as citations studies, but not restricted to them).

The descriptive studies are probably the most easily understood, and can be thought of as resembling in many ways population census. Both have to provide comprehensive and accurate data about rapidly changing population in one case of people and in other of literature. Literature descriptions are assembled from the bibliographic study those that provide data on the condition or character of the literature as a
whole. There is, therefore, an important distinction between bibliographic and bibliometric data.

### 2.2 Literature Descriptions

Literature descriptions concern themselves with the following features or characteristics of documents.

1. Bodies responsible for the production and transmission of the information
2. Form of transmission (e.g. journal, monograph),
3. Medium of communication (e.g. article, letter),
4. Nature of information conveyed subject and language characteristics
5. Timing and frequency with which information is conveyed
6. Geographical origin'.

Descriptive studies are not only concerned with representing faithfully the features present in the current literature, but also they are interested in the condition of past literature, particularly in areas where past literature constitutes an important part of current research scholarship or interest.

### 2.4 Behavioural Studies

All documents rely to a greater or lesser extent on information contained within previously published documents. Inside the literature there exists, in fact, a web of relationships. By exploring the full
bibliographic ramifications of these relationships (by matching for instance, the bibliographic
description of the most publication with that of the cited publications), the nature and strength of
each relationship can be determined.

Information Scientists divide Bibliometrics into two as follows:

1. Productivity Count (Descriptive)
   a. Geographic (Countries)
   b. Time periods (Eras)
   c. Discipline (Subjects)

2. Literature usage counts (evaluative)
   a. Reference
   b. Citation

Fundamentally both categories of divisions are similar and do not contradict with one
another. Metrics are rapidly forming the methodology or a system of methodologies, which may
put them as candidates to be considered as forming a new scientific branch. Works of many
pioneers such as Price, Garfield, Small, and their contributions played vital role in formulating
the requirements of the methodology.

2.5 Bibliometric Laws

The three fundamental laws which laid the formation of bibliometrics are

a. Bradford's law of scattering of scientific papers
b. Zipf's law of word occurrence
c. Lotka's inverse square law of scientific productivity
2.5.1 Bradford's Law of Scattering

Bradford\(^5\) revealed a pattern of how literature in a subject is distributed in journals. "If scientific journals are arranged in order of decreasing productivity of articles on a given subject, they may be divided into a nucleus of periodicals more particularly devoted to the subject and several other groups of zones as the nucleus."

2.5.1.1 Bradford's formula.

\[ R(n) = N \log \frac{n}{s} \quad (1 \leq n \leq N) \]

Where \( R(n) = \) Cumulative total of items contributed by the sources of rank 1 to n.

\( N = \) the total number of contributing sources.

\( S = \) constant characteristic of the literature. We then can say that \( R(N) = N \log \frac{N}{S} \) is the total number of items contributed by N sources.

In other words Bradford\(^6\) describes that" If all the scientific journals relevant to a given subject are assigned ranks according to the number of articles they carry on the subject, are assigned ranks according to the number of the articles and the journals are grouped into m zones each containing the same number of articles, with the first zone containing the highest ranking journals and the last zone containing the lowest


\(^6\) Ibid.
ranking journals, then the number of journals in the succeeding zones form a geometric series with a common ratio \( P_m \) depends only on \( m \).

The statistical regularity pointed out by Bradford's law provides an objective means of determining zones of relative richness or value to a given kind of literature. This has implications to the various processes in a library. For example, a library can safely stock the journals which belong to the core or nuclear zone. It is advisable to extend the purchase list to the next zones till the budget limits permit. If at all the library budget is elastic, a point will be reached at which it would be desirable to obtain copies of articles in the journals on demand rather than subscribing to the journal.

Lancaster\(^7\) provides an excellent hypothetical example of applying Bradford's law in periodical collection building, while discussing the principle of diminishing returns.

Brookes\(^8\) is of the view that if the total expenditure on periodical provision is limited to the fraction \( f \) of the sum needed to cover the subject completely, the buying of periodicals may be supplemented by the buying of photocopies of the relatively few relevant papers published in the peripheral periodicals. While preparing bibliographies we are faced with the problem of coverage, the journals that are to be scanned etc. size of a bibliography and the periodicals that should necessarily be included in the list of items to be covered. On the application side of this Bradford's

\(^7\)Lancaster, F. W. *If you want to evaluate your library....”, 2\(^{nd}\) Ed London: Library Association, 1993, pp. 285-290.
distribution can be fruitfully used to estimate the total law, the studies of Gottman and Morris\(^9\) and I.K. Ravichandra Rao\(^{10}\) are significant.

Gottman and Morris found that the pattern of journal usage in the Allen Memorial Library follows a Bradford distribution. Ravichandra Rao, through his analysis of circulation data collected from six Canadian University Libraries, proved that the rank distribution of transactions follow a Bradford distribution. Bradford's law is very much akin to the Pareto's law relating to wealth distribution and the 80-20 principle used in warehouse management and the Mandelbrot distribution.

Naturally this law is applied to study not only the scattering of publications, but also in other spheres of activity also. A study conducted by Garg and Lalitha\(^{11}\) Sharma of R & D indicators in Indian Industry using Bradford's law bears testimony to this fact. By analysing the R & D expenditure of 452 in-house R & D units in different sectors of the Indian Industry, they had identified 19 in-house R & D houses as the core, 60 as the medium and the rest as small. As compared to medium and small-level in-house R & D units, there is a heavy concentration of manpower deployed, papers published, patents filed, processes/products developed in the core in-house R & D units. This shows the superiority

---

of the core not only in the R & D expenditure but also in other yardsticks too.

Performing a Bradford Analysis

1. should start with some subject area of interest (e.g. citation analysis).

2. locate all articles that cover the subject during some time period (e.g., 1990-2000). Using a database with comprehensive coverage like Library Literature, can speed things up.

3. record the number of relevant articles found as a function of the journals names where the articles originated.

4. summarize the data by counting how many journals have published x relevant articles (e.g. One journal PUBLISHED 5 articles; One journal published 4 articles; Two journals published 2 articles each; 12 journals published 1 article each).

5. rank the results in decreasing order with the most prolific journal given rank 1.

6. arbitrarily select a core of j journals. The number of articles found in the core j journals is denoted as a zone. Denote the first grouping (i.e., the core) as zone O. (e.g., select a core of 3 journals 0=3) which will have 12 articles associated with it (a=12)).

---

7. divide the remaining data into "Bradford Zones" such that each zone contains a articles (12 articles), dividing the data into Bradford Zones.

Bradford's law states that there is some constant k such that the zth zone containing articles consists of kzj journals where the initial, or core zone, is the 0th zone. K is also known as the "Bradford multiplier" and is determined by the data. The proportion k0:k1:k2:k3 is another view of the same phenomenon. To proceed further, one has to determine the value of k: consider the following:

Zone 0 contains 3 journals
Zone 1 contains 6 journals
Zone 2 contains 12 journals

Thus these data contain the proportion 3:6:12 = 1:2:4 = 20:21:22
Therefore, k = 2.

Apply the kzj formula to the data. Now the number of journals needed to provide multiples of the cores quickly grows to very large numbers. This is a classic example of the scattering effect i.e Bradford's law of scattering a measure of the rate of obsolescence by distinguishing the usage of the levels of items. Essentially, this is a method of clustering.

For example, 9 journals have 429 articles, the next 59 have 499, and the last 258 have 404. We roughly get three groupings (ranging
from 404 to 499) of articles. Three groupings, 9 titles, 9*5 titles, 9*5*5 titles.

2.5.1.2 Bibliographic Scattering

Taking a topic or a subject, arbitrarily chosen, identifying characteristics of an article, the scattering or distribution is shown to follow a clustering process of ranked groups, first noted by Bradford. Bradford's law is almost universal.

Quentin L. Burrel, Statistical laboratory, Dept. of Mathematics, University of Manchester in his article "On the growth of Bibliographies with the time: An exercise in Bibliometrics prediction" says,

Bradford sought ways of providing an efficient, service for abstracting and indexing scientific and technical literatures. As is surely well known Broadford's investigation concentrated on the productivity of journals more particularly he sought to identify those journals which were relevant in the sense of publishing articles in a particular subject area during the period of study. His empirical studies gave rise to the Bradford law of bibliometrics.

In Bradford's content we have a population of academic journals and in compiling a bibliography we wish to identify those journals which produce articles relevant to our field of interest and the number of articles so provided by each. While a no of authors have concentrated on the importance of having as nearly as possible a complete search of
the possible sources, an important practical point raised by Bradford is that 'even when the actual producer during a period of years had been ascertained, new sources would certainly appear during a further period'.

2.5.1.3 Models of Bradford's Law

Mathematical expressions or models have been suggested that reflect the graphical relationship observed by Bradford. These models can be classified as

1. empirical
2. causal and
3. formal

2.5.1.3.1 Empirical models

Empirical models are only justified a 'posteriori' by their ability to give a good fit to sufficient and varied data.

2.5.1.3.2 Causal models

The mathematical relationship is usually derived from some known underlying causative principles characterizing the process. Causal models are explanatory in nature, and may be either deterministic or stochastic.

2.5.1.3.3 Formal models

Formal models may be used in situations where the underlying process is too complex or dependent on too many factors to be
expressed as a simple causal relationship. Some examples of the three categories of models are the following

Egghe\textsuperscript{12} (1985) A generation of Leimkuhlers\textsuperscript{13} model, it hinges on a mathematical interpretation of Bradford's law.

Burrell: (1980), This is a model of library circulation based on a stochastic process and could be classed as a causal model.

Basu\textsuperscript{14} (1992,1993), A formal model of the most probable unequal distribution of a set of articles in a set of journals, called the Random Hierarchical model.

\textbf{2.5.2 Zipf's Law}

In the field of 'Language' studies statistics on word frequencies have been compiled for over a century. The first major work is by Dewey - Relative frequency of English words published in 1923. The major effort at statistical analysis of word frequencies was made by zipf's in the psychology of language published in 1935. Striking regularities in the frequencies distributions of words number of the different words $W(k)$ occurring $k$ times were noted in different language, authors and types of literature. In a very elaborate subsequent work Zipf extended the investigation beyond language in economics, geography etc established a' universal law of frequency distribution encompassing a wide

\footnotesize


diversity of statistical data. Zipf also propounded a principle of least effort as a rationale for the rank-frequency distribution.

Zipf had developed and extended an empirical law, governing a relation between the rank of a word and the frequency of its appearance in a long text. If "r" is a rank of a word and 'f' is its frequency then Zipf's law is stated as follows:

\[ rf = c \]

Where 'c' is a constant. Zipf derived his law from a general principle of least effort. Words whose cost of wage is small or whose transmitted demands the least effort are frequently used in a large text.

It states that in a fang textual matter, if the words are arranged in their decreasing order of frequency then the rank of any given word of the text will be inversely proportional to the frequency of occurrence of the word.

Mathematically Zipf's law can be written as

\[ r = \frac{1}{f} \text{ or } rf = c; \]

Where 'r', denotes rank word, 'f' stands for frequency of occurrence of the word and 'c' is Constant. Zipf's law can be effectively used in the generation of semi-automatic or automatic indexes useful for an information retrieval system. Its use has increased tremendously with the emergence of natural language indexing of textual matter especially in electronic form. Several studies aimed at finding out the pattern of
frequency distribution of descriptors of a thesaurus and the distribution of indexing terms are available.

A prominent one among them is that of Fedorowicz\textsuperscript{15}. Zunde and Siamecka\textsuperscript{16} have developed a function for the optimum distribution of indexing terms by the number of postings. This should make it possible to transmit information with maximum efficiency. Zipf's law provides a measure of the richness in vocabulary of an author. This technique can be used for deciding the correct authorship of disputed works.

For example, if there is difference of opinion as to the correct author of a work, the work predilections of the attributed authors can be analysed either manually or using a computer. Once the frequency of occurrence of favourite words are decided the disputed text can be analysed to see similarity and thereby decide the author conclusively.

The law is also used for identifying words more frequently used in different foreign languages. These words are taught first in the instructional programmes of foreign languages. Emilie C White\textsuperscript{17} observes that the super-imposition of the Bradford distribution over the linear Zipf distribution, which demonstrates the emergence of more used and popular items may yield a technique to describe the pattern of books used by library patrons. She feels that this law, when applied to

\begin{footnotesize}

\end{footnotesize}
circulation data, these formulations can support such policies as shortened loan periods for heavily used books and the identification of a core collection.

For analysis, this can be applied by counting all of the words in a document excluding the stop-words (a, an, the, therefore, etc.) with the most frequent occurrences representing the subject matter of the document. We could also use relative frequency instead of absolute frequency to determine when a new word is entering a vocabulary.

Naranan\textsuperscript{18} has shown some striking similarities to Zipf's Law In any democra tive election in which two or more candidate stand, the candidates are ranked in order of decreasing frequency of the votes they individually individually attract. In this example the candidates are the sources and the votes are the items.

2.5.3 Lotka's law:

Lotka's\textsuperscript{19} proposition led to a whole gamut of studies on scientific productivity. Such productivity studies have gained momentum in the post-second world war period. This in fact, has culminated in the rise of a new discipline called Scientometrics. It is defined as the' study of the measurement of scientific and technological progress'. It provides an understanding of the structure of scientific activity, the disciplines

\textsuperscript{17}White, Emilie C. "Bibliometrics: From Curiosity to Convention", Special Libraries, (1985), pp. 35-42.
being researched, the organisations involved, the strength and deficiency in the scientific groups and their communication channels and at different levels of aggregation. It follows a trajectory of econometrics in the use of quantitative data, concepts and models and extensive use of mathematical and statistical techniques of modelling and data analysis.

Scientific productivity studies have been made from different angles. Impact of social change on scientific productivity, relationship of publication output on scientific recognition, identification of elites in different disciplines, occurrence of discoveries in different cultures etc. are some of the approaches made in this line

An interesting study of scientific productivity made three decades back is that of Yuasa20. Yuasa's was a statistical study of the scientific achievements in various countries that showed the shifting of the world scientific dominance from one country to another. He found out that this dominance shifted from Italy to Britain, then to France, from France to Germany and finally to USA in the 20th Century. Price21 who had traced the development of science in Babylon and plotted the growth of big science from little science had observed that Latka's law applied equally well to the productivity of scientists in the 17th as well as in the 20th century. This meant that majority of publications emanated from a handful

of people. This is very much similar to the already existing Price's square root law. Similarly, the conclusion of an extensive review of early studies of scientific productivity made by Narain was that scientific talent was highly concentrated in a limited number of individuals.

Generally, Lotka's law is an inverse square law that for every 10 authors contributing one article, 25 will contribute two; 11 will contribute three, and 6 will contribute four each. There is general decrease in performance among a body of authors following $1:n^2$, this ratio shows that some produce much more than the average which seems agreeably true for all kinds of content creation.

Thus, we can say that Lotka's law may not be constant in value, but in following Inverse Square. Our challenge will then be to find the correct square in different mediums & fields. Other authors like the Spanish scientist A Lara, the Russian, Marschakova the German, D Schmidmayer, the Belorussian v Lazarev, the Lithuanian, O Voveriene to name a few proposed their own solutions, defining the limits of competence of these three-metrics.

2.6 The Logarithmic Law

Analysis of the data form many social contents over the past two years has led to the discovery that the general form of the fir distribution is logarithmic. Two special cases of this logarithmic law were first noted in1935 one by Zipf's in the statistics of natural language vocabularies
and one by Bradford in the distribution of scientific papers relating some specified topic over the contributing serials.

The law states that in a homogeneous activity the sources of the sources of rank 'r' is given by 
\[ g(r) = K \{ \log (a+r) - \log (a+r-1) \} \]

Where K is a constant dependent on the sample size. The cumulative activity of the first r sources is given by 
\[ G(r) = (q(r) = K \log (a+r) \]

So, the conformity with the law is most easily shown by plotting the successive values of G(r) for r = 1,2,3... in succession against \( \log (a+r) \) - which is easily done if semi-logarithmic graph paper is used. If the cumulative plot is linear then the data conform to the log law. The log law is simply a mixed Poisson distribution of a regular kind. It can be derived theoretically by considering an infinite' population of mixed abilities in a very natural way.

Because social activities are susceptible to social regulation within the group, the pure logarithmic form of the law is relatively rarely found in empirical data. But the variant forms are easily recognized and, of course, are of special interest because they reflect social interactions with the group.

2.7 The Social Log Law and the Individual

The social distribution of judgments is logarithmic. And so, by analogy with the Gaussian model, it is reasonable to consider the hypothesis that the individual distribution is also logarithmic.
A simple model to account for Bradford's law was proposed by Naranan\textsuperscript{22} in 1970. The most significant first step in the model was to recognize the Bradford's law is equivalent to a simple power law distribution of articles in journals. Specifically $J(p)$ the number of journals carrying exactly $p$ articles is of the form of $J(p) = Kp^\alpha$ (where $K$ being constant and $\alpha = 270$) to explain the power law relation. Naranan was inspired by a model proposed by Fermi in 1949 in cosmic ray astrophysics.

Naranan investigated statistical laws governing distribution of number of citations to articles, citation to journals etc. and found them also to be power laws similar to $J(p) = Kp$.

It has been found from the earlier studies that collaboration in research varies from discipline to discipline and for the same discipline from time to time and from one country to another collaborative research has also been found to be influenced by a number of factors such as funding use of large-scale equipments etc. However, the general conclusion of many a study has been that multiple authorship (i.e. collaborative research) is on the increase in science and technology.

Bibliometrics, as any other branch has been growing in quantum and quality over the years and has given birth to new branches such as Informetrics and scientometrics. The quantum of literature output in various disciplines is more from developed countries.

2.8 Economic Development and Literature Output

The nations of the world stand divided on the concept of 'Have and Have not' which decide the economic group among them. Gross National Product abbreviatedly known as GNP is considered as an indicator to the economic development of a nation. In the post modern era, countries are divided on the basis of 'information rich' and 'information poor' status and in general, GNP is the reflecting factor to a nation's economic growth.

The Gross National Product (GNP) is the total dollar value of all final goods and services produced for consumption in society during a particular time period. Its rise or fall measures economic activity based on the labor and production output within a country. The figures used to assemble data include the manufacture of tangible goods such as cars, furniture, and bread, and the provision of services used in daily living such as education, health care, and auto repair. Intermediate services used in the production of the final product are not separated since they are reflected in the final price of the goods or service. The GNP does include allowances for depreciation and indirect business taxes such as those on sales and property.

2.9 General Bibliometric Studies

Maclas-Chapula and Muanoos-Nolasco\textsuperscript{23} analysed AIDS documents as produced on Sub-Saharan Africa and found out that the main countries participating in AIDS research were Democratic Republic of
Congo and Cameron. The results indicated a high pattern of collaboration through multiple authorship. The subject content of the documents were found to be focused mainly on epidemiology, complications and prevention and control issues of AIDS.

Maclas-Chapula\textsuperscript{24} also studied the literature on Health system reform in Latin American and the Caribbean through the web as well as the databases on CD ROM. The results in his webometric and bibliometric investigations, there was no comprehensive databases in terms of time, document type and content coverage. The results indicated the need to organize and administrate the existing literature on healthcare reform so as to transfer it into the knowledge demanded by the user community.

Halkar et al.\textsuperscript{25} made a bibliometric study on the Journal of Family welfare drew the following conclusions:

1. Maximum number of articles published in 1993 was the same as in 1997 i.e, 37.(13.40 percent).
2. Authorship pattern showed that most of the papers were contributed by single author (52.18 percent) and went on decreasing by two and more authors.
3. India contributed the maximum number of articles i.e, 80.07 percent since the journal under coverage was from India.


4. The average length of the papers was between 4 and 9 pages constituting 60.05 percent

Kamlesh Goel\textsuperscript{26} studied the authorship pattern, areas of research, journals, institutions and regions covered in the social science citation index (SSGI) pertaining to the year 1998. The following are his findings

1. The number of highly productive research and academic institutions were almost the same. The highly productive institutions produced about 33 percent of the total Indian output in social sciences.

2. Research in sociology received maximum attention followed by psychology, economics, planning and development.

3. Most of the findings were published in journals from the West. However, most commonly used journals were from India.

Ramesh et al\textsuperscript{27} analysed the papers published in Oryza the Quarterly International Rice Journal from 1986-1995. The analysis showed that multiple author contributions constituted the maximum proposition (87.82 percent) and the degree of collaboration over this period varied from 0.90-0.95. The length of the articles with 1-5 pages were found to be at the maximum with 78.3 percent.

Subbiah Arunachalam and Subbiah Gunasekaran\textsuperscript{28} made a bibliometric study on tuberculosis research in India and China identified that there is a tremendous mismatch between the share of the burden of the disease and the share of research efforts.

Turkeli\textsuperscript{29} examined the post doctoral productivity of Turkish Physicists along with related social and environmental factors.

Dr. S. Ravi\textsuperscript{30} analysed the nuclear science research productivity of Indian scientist and found that

1. Nuclear science research papers were published mainly in journals.
2. Among the international sources of publications United States and United Kingdom predominated in publishing Indian Nuclear Science Research papers.
3. Two authored and three authored papers were more than the single authored and other multi authored papers.

\textbf{2.10 Language of Contribution Studies}

Loutti\textsuperscript{31} identified the language preference of writing research papers by Psychologists, Chemists and physicists. It was observed that


\textsuperscript{29}Turkeli, A. "Doctoral Training Environment and Post Doctorate Productivity Among Turkish Physicists", \textit{ScienceStudies},3 (1973): 311-318


\textsuperscript{31}Louttit C.M. "Use of Foreign Languages by Psychologists, Chemists and Physicists", \textit{American JournalofPsychology},70, (1957): 314-316
reference made by writers in English language journals were 92.5 percent in English, in German journals 91 percent German and in French journals 64.6 percent French. Further it was said that numerous studies in Social Sciences show reference in American sources having around 90 percent in English.

Simonton\textsuperscript{32} identified that in two English language source journals in the field of Fine Arts, more than half the references were the materials in foreign language references.

Heussman and Pulmer\textsuperscript{33} analysed the literature kept in the general library catalogue of Michigan University. It was noted that out of the total literature, 89.52 percent were in English.

2.11 Country of Publication

Rangarajan and Poonam Bhatnagar\textsuperscript{34} analysed the Bibliometric data compiled from Physics Abstracts on research papers published in the field of Mossbauer effect studies over a period of two decades from its discovery in respect of media choice. The findings indicated that there was a world-wide trend to publish in journals outside the country of origin of the research work.


Klaic\textsuperscript{35} examined the research activity of chemists from the Rugjer Boskovic, Yugoslavia during 1976-1985 covering 2018 research papers of scientific work. The papers were classified according to subfields used in the Journal Citation Reports.

Nederhof et al\textsuperscript{36} made a Bibliometric study to assess the performance of departments in the field of Natural and Life sciences, the social and Behavioral sciences and Humanities. The result explains that nearly one third of the departments publications were not covered in the Science Citation Index.

Braun\textsuperscript{37} classified all the scientific publications on the basis of the data obtained from the Institute for Scientific Information, (ISI) Philadelphia. All countries which published atleast 50 first authored papers in the field in question during the study period were included. The sources journals during the period 1980-84 and 1985-89 were considered as source items and citations to them were counted for the periods from 1980 to 89 respectively.

2.12 Pattern of Authorship and Literature Output

Collaborative contribution is the result of Team Research and Team Relay Research. Collaboration has existed in science since its beginning, however, the first collaborative publication was published in 1655. Since


then, the number of collaborative publications have increased, initially slowly, and then dramatically faster after the middle of 18th century. The extent of collaboration further increased towards the end of this century and the trend continues.

2.13 Co-Authorship Studies

Gupta and Karisiddippa\textsuperscript{38} have studied the collaborative pattern of scientific papers in theoretical population Genetics for a period ranging from 1956 to 1980 and concluded that

1. The United States contributed 41.66 percent to the world's total international co-authored publications. It was among the countries with a significant percentage of co-authorship (47.02 percent), but it ranked in the low end in terms of overall percentage of publications involving international collaboration (11.53 percent during 1956-60 to 15.81 percent during 1976-80). This occurred solely as a result of its large publication base.

2. The United Kingdom contributed 16.23 percent to the world's international co-authored publications. It was among the countries with 32.15 percent of its total publications appearing as co-authored publications, but ranks in lower end in terms of its overall percentage of


publications involving international collaboration (6.25 percent during 1956-60 to 24.13 percent during 1976-80).

3. Australia contributed 7.45 percent to the world's total international co-authored publications. Its percentage of co-authored publications were 23.41 percent, but its contribution in terms of publications involving international collaboration decreased (from 50 percent during 1956-60 to 33.33 percent during 1976-80).

4. Japan contributed to the world's total international co-authored publications and its percentage of co-authored publications was 39.58 percent.

5. Canada contributed 13.59 percent to the world's total co-authored publications and its percentage of co-authored publication was 43.05 percent

2.14 Collaborative Authorship

Many studies have been reported in the literature on the nature of collaboration among scientists in different disciplines. One of the earliest studies was undertaken by De Solla Price\textsuperscript{39} who in one of his investigations reports that the incidence of collaboration had increased steadily over time and has been rapidly growing since the beginning of twentieth century.

Since then, a series of studies have shown that an increase in collaborations is evident in most areas of science and technology and that

the extent of collaboration varies from one discipline to another. Clarke\textsuperscript{40} in his study on Bibliometric papers criticized the view of Price and concluded with a generalization as regards the increasing trend towards multiple authorship is not valid for science as a whole.

There has been consistent trend towards increased collaboration in all major branches of sciences over the years in the present (twentieth) century", reports Meadows\textsuperscript{41}

At the same time, the rate of increase in multiple authorship varies from one subject area to another. In Physics the proportion of the single author papers have fallen from 75 percent in the 1920s to 39 percent in the 1950s. The corresponding figures for psychology are from 84 percent to 55 percent, reports Merton\textsuperscript{42}.

It is evident that single authored articles termed as solo research or research in peril has been on the declining trend in the modern era, reports S.R.Ranganathan\textsuperscript{43}. What S.R.Ranganathan predicted for the modern era holds good for the post modern era also.

\textsuperscript{41}Meadows, A. I. Communication \textit{in Science}, London: Butterworths, 1974
\textsuperscript{42}Merton R. K. and Garfield, E. "Forward to Little Science, Big Science and Beyond by D.J. De Solla Price", New York: Columbia University Press, 1963
\textsuperscript{43}Ranganathan, S. R. \textit{Prolegomena to Library Classification}, Bombay: Asia Publishing House, 1963
Manten\textsuperscript{44} has studied the multiple authorship in the field of Earth Science. This study identified the increase in the frequency of multiple author papers.

Zuckerman\textsuperscript{45} examined the research output contribution of 41 Nobel laureates. The result indicated that there was high degree of collaboration and productivity among them.

Heffner\textsuperscript{46} identified the relationship between collaboration and financial support for research in four disciplines relating to Political Science, Psychology, Biological Science and Chemistry. In all these disciplines financial support for research was associated with an increase in the total number of persons involved in the production of knowledge. His inference is an indirect indication of the relationship between the GNP and research output.

Maheswarappa, et al.,\textsuperscript{47} analysed the collaborative research in Indian Science and Technology based on authorship data collected from the Indian Science Abstracts covering the periods 1965-70, 1975-80 and 1980-83. The findings revealed that two authored papers were maximum in Science and Technology as whole. The single authored papers constitute more than one fourth of publication.

\textsuperscript{44} Manten, A A. "Multiple Authorship in Earth Science", Atlas, 4A (1968): 149 A -152 A
\textsuperscript{46} Heffner AG. "Funded Research Authorship Pattern and Sub-authorship Collaboration in Four disciplines." Scientometrics, 31 (1) 1981: 5-12
Harsanyi\textsuperscript{48} examined the authorship pattern of publications in Library and Information Science and considered the methodological impact of various ways of allotting "credit" for multi-authored works and relationships between multiple authorship and other publication variables such as quality and impact. Given the complex relationship between collaboration and productivity, the concomitant use of non-bibliometric methods of studying collaboration, as well as the application of meta analysis is suggested.

Kannappanavar\textsuperscript{49} and Vijayakumar made a study on the authorship trend in international monetary fund literature for a period from 1991-1998 and concluded that collaborative research is in an increasing trend varying from 0.45-0.62. The average degree of collaboration was found to be 0.56.

Mahapatra and Bhagavan Dossso also have found out that collaborative research has a positive growth trend varying from 0.65-0.81 by studying five selective journals in geology covering a period from 1987-1996.


\textsuperscript{49}Kannappanavar, B.D., and Vijayakumar, M. "Authorship Trend and Solo vis Team Research in International Monetary Fund Literature: A Bibliometric Study." \textit{Annals of Library and Information studies} 48 (2001): 117-120

They have also concluded that the growth of publications in geology is gradually reducing towards the end of the period. Seglen and Aksnes\textsuperscript{51} analysed the relationship between research group size and scientific productivity within the highly co-operative research environment, characteristics of contemporary biomedical science. They made an investigation in Norwegian Microbiology. They found out that most of the Norwegian microbiological research (73 percent of the microbiology articles) appeared to be performed by specialist groups (with \( \geq 70 \) percent of their production as microbiology), the remainder being published by groups with a broader biomedical research profile (who were responsible for 95 percent of the non-microbiological articles). There was no correlation between group size and productivity.

R. Kundra\textsuperscript{52} investigated the collaborative research trends in Indian Medical Sciences 1900-1945 and drew general and broad conclusions.

1. There is a perceptible increase in the collaborative research and substantial decrease in the single-authored papers in medical science in India during the period 1925-45. In 1900, not a single collaborative papers was reported. But by 1925, 12.4 percent collaborative papers were reported and the single-authored papers were 87.57 percent. The figure rose to 33.6 percent co-authored papers in 1945 while the single-authored papers

---


declined to 66.36 percent. There was almost three times increase in collaborative research.

2. Of the total number of co-authored papers published in the 1900-45, 21.95 per cent collaborative papers were from basic research while 7.63 per cent were from applied research. This confirms the findings of Frame and Carpenter.

3. A rise is not only seen in the collaborative papers but also in the nature of collaboration. In 1925 only 22.9 per cent authors participated in collaborative research which rose to 49.7 per cent in 1945, a two fold increase.

4. The collaborative authorship pattern steadily increased from 1.00 per paper in 1900 to 1.39 per paper in 1945. But the average authorship remained 1.2 author per paper.

The growth pattern suggests that a large proportion of co-authored papers in a discipline or a journal depends to some extent on the type of research and the discipline involved. As a result, it is not impossible to have a relatively lower proportion of collaborative papers in a particular sample, even when collaborative research overall has become the normal practice.

The collaborative and authorship trend discussed in the study suggests that medical science was still in the developing stage in India in
the period 1900-1945, and there was possibility of its expansion in the near future.

Dhiruv Raina and others have studied the evolution of collaboration in four sub-disciplines of Physics for the period 1800-1950. They have reported that

1. The overall evolution of physics publications in India reveals a remarkable break with the past in the decade 1920s onwards. In fact the growth rate enters a new face after this time. Furthermore, these growth rates exhibit quasi-doubling.

2. The collaboration coefficients taken decade wise are strongly correlated with the total number of publications taken decade wise. This conforms with increasing collaboration as the number of publications increases.

Chemical Index (CI) and Current Contents (CC) had been used to investigate the nature of collaboration characteristics of the sub-disciplines. It was found that 1920 onwards the collaboration measures showed an increasing trend, and that trend is uniformly reflected across sub-disciplines.

Furthermore, the increased collaboration is also a reflection of the presence of strong leaders in the field, leaders who were subsequently responsible for the institutionalization of the discipline in the country.

---

Using a set of measures developed by Egghe, the collaboration profiles of the four Indian physicists were calculated and it was found that while that collaboration measures increased over time, as is normally expected, there were decades where the number of publications decreased but the collaboration measures continued to be significant. Further these decades corresponded with the decades in the life of these physicists where they were busy establishing institutes for physics research.

3 The study hoped to insinuate the manner in which the history of science can endow "scientometrics" with a modality of explanation.

4 In the breaks, evident is the evolution of the growth trajectories of the sub-disciplines of physics and collaboration profiles of physicists, historians of physics in India can converge upon fruitful problems for research.

Amit Kumar Bandyopadhyay\textsuperscript{54} analysed the references appended to 92 doctoral theses submitted to department of Mathematics, Physics, Mechanical Engineering, Philosophy and Political Science, University of Burdwan. The findings of his analysis are

1. Authorship collaboration is high in physics. Moderate collaboration exists in mathematics and mechanical engineering. The collaboration is very low in political science and philosophy

excluding psychology. It is highest in nuclear physics followed by optics.

2. Although the multiple authorship trend has increased steadily through decades (1950-1990) in all the branches of physics and mathematics and also in psychology: in mechanical engineering, philosophy excluding psychology and in all the branches of political science the multiple authorship trend has declined for certain periods.

Cano\textsuperscript{55} reviewed 17 years of research in Library and Information Science in Spain for a period from 1977-1994. He identified that the Spanish research in Library and Information Science had concentrated more on information retrieval, description of services and studies of scientific communication. Authorship pattern suggested a prevalence for individual authorship(68 percent).

\textbf{2.15 Literature Growth Studies}

Cole and Eales\textsuperscript{56} made a pioneering effort in the Bibliometric oriented growth analysis of literature. Their analysis was deemed to be statistical analysis of literature of comparative anatomy covering the period 1550-1860. They identified the fluctuations of interest and distribution of literature among various countries.

\footnotesize

Ozinonu\textsuperscript{57} made an early survey relating to growth of Basic Science in Turkey. The author identified the growth of manpower and frequency of publications in Mathematics, Physics, Astronomy, Chemistry and Bio-Science for the period 1933-1966.

Menard\textsuperscript{58} identified the changes in the mode and rate of growth in Earth Science form 1800 to 1970. From the period 1945-1970 the author observed that Earth Science began to grow rapidly with the doubling time of 5 ½ years.

Uzun\textsuperscript{59} analysed 860 Physics publications in Turkey for the period 1938-1981. The productivity and growth characteristics of the research in the experimental and the theoretical areas as well as in different sub-fields and institutions in the country were brought under the purview of the discussion. Sengupta and Lalita Kumari\textsuperscript{60} analysed the growth of AIDS literature between 1976 and 1986. They identified the international channel of communication, contributing countries, authorship trend and related facets.

\begin{itemize}
\item[\textsuperscript{57}] Ozinonu K. "Growth in Turkish Positive Basic Science 1933-1966." \textit{Publication of the Middle East Technical University}, no.17, Ankara, 1970
\end{itemize}
Parvathamma and Gunjal\textsuperscript{61} examined the growth of Indian Earth Science literature during the period 1978-88 was found to be 4.8 years.

Parameshwaran and Smitha\textsuperscript{62} made a bibliometric analysis of library and information science abstracts (1994-1998). Their findings are

1. Maximum number of publications fall under the broad fields information and communication technology with 13.41 percent coverage.
2. More people write individually i.e, single authorship amounts to 77.5 percent and double authorship is 15.83 percent indicating that solo research predominates in the field of LIS.
3. The portion of Indian contribution to LIS research is very meager (1.142 per cent).

Ramakrishna and pangannaya\textsuperscript{63} applied a key word search on four databases DERWENT BIOTECHNOLOGY abstract BIOSIS, INVESTEXT and CASEARCH and analysed the relative growth rate and doubling time of the journals covered in these databases with the key word "Biotechnology". The study indicated that within equal economic, intellectual and environmental conditions, literature on different subjects follow a similar growth trend.


Iles Rocamora, Maria De Los Angeles investigated about the Spanish Historical epidemiology. They have surveyed publications about the causes of cholera-morbo and measures to prevent it by collecting books, brochures and journals printed in Spain between the period of 1854 to 1885. The selection criteria were: publications related to the origin, transmission, infection, and/or description of the microorganism and the prevention of cholera. We have excluded those publications about Ferran's vaccination, following bibliometric criteria. They developed a prosopographic study about the main related sources of scientific information and a text analysis of a characteristic work of that time: Gonzalez de Samano, Memoria historico del colera-morbo asiatico en Espana. (1854-1885).

They concluded that the three epidemic infections during the periods 1854-1856, 1865-1866 and 1884-1885 and the discovery by Koch (1883) of *vibrio* cholera trigger an increase in the number of publications. One example was the huge number of reports derived from the International Sanitary Conference in Istanbul. The most important journal is: El Siglo Medico. The maximum editorial production of brochures (fast diffusion) was reached at the beginning of the epidemic infection by 1854. But it is after knowledge about the origin and prevention of the sickness that the number of books edited was higher (1884). Those books were

---

published in big cities and usually were not reedited. Tips about papers, together with news about congress and foreign institutions show the importance of Pasteur’s bacteriology school.

News about Spanish institutions showed more activity in Madrid and Barcelona in 1885 reflecting the controversy surrounding Ferran's vaccination. Authors published in a rather sporadic way, being generally rural physicians.

Nerur, Sridhar P.\textsuperscript{65} in order to verify the question of whether the field of Object Orientation (OO) is growing at a rate that is sufficiently high for it to be deemed to be an emerging technology used bibliometric analysis. His research also examined the ontological, epistemological, and methodological dimensions of both OO and the traditional approach (e.g., structured) to software development, with a view to establishing their paradigmatic differences.

Kuhn's model of paradigms was the conceptual basis for this research. Author Co-Citation Analysis (ACA), a form of bibliometrics, was used in this study. A total of 31 authors was taken up for study and co-citation counts for each author pair, during the period 1980-1993, were obtained. The frequency of co-citations between authors was used as a measure of the degree of consensus. Cluster analysis, factor analysis, and multidimensional scaling (MDS) were used to validate the cluster

compositions. The results suggest that OO is conceptually different from structured methods.

The proposition that OO is an emerging technology was tested independently on two sources of data. The growth rate of OO indicated in this study suggests that OO is no longer in the preparadigmatic phase, but is now normal science. The implications and challenges of a new paradigm were explored. Suggestions for extending the study and directions for future research were also presented.

Guillen-Salazar, Rederico⁶⁶ used bibliometric technique to assess recent progress in animal behavior research in Spain. The main part of the dissertation consists of an analysis of the contributions made to animal behavior science by Spanish researchers in the period between 1970 and 1980. To this end, we collected about 2000 scientific works (articles, abstracts of presentations delivered at professional meetings, doctoral dissertations, books, etc.) produced in Spanish research centers and analysed them by means of bibliometric techniques. They discussed the data concerning each author's productivity, journals and languages used in publication of results, the author's work place, topics and taxonomic groups studied, etc. This analysis was complemented with a description of the most productive groups of authors.

Suresh Kumar, Praveen Sharma and K.C. Garg\textsuperscript{67} have contributed a study which aims at determining the applicability of Lotka's Law, negative binomial distribution and lognormal distribution for Institutional productivity, in the same way as it is to authors and their productivity in the field of engineering sciences and the patterns filed by industrial firms in laser Science and Technology. The study indicates that none of the three distributions for institutional productivity in engineering sciences. However, Lotka's law holds good for full as well as truncated set of data for the patents filed by industrial firms.

\textbf{2.16 Bibliometric Analysis of Bibliometric Literature}

Humayoon Kabir\textsuperscript{68} made a Bibliometric study of bibliometric literature and reported that solo research predominates and degree of collaboration ranged from 0.20-0.35. Bibliometric literature is doubling in every 10 years.
