JISSI The International Journal of Scientometrics and Informetrics

Volume 2 Number 2-3 June - September 1996

*Editorial Correspondence to :*
Subir K Sen
Honorary Executive Editor
85 (Old) Devi Nivas Road
P O - Motijhil
Calcutta 700 074, India
Fax : +91 33 551 2180

*Business Correspondence to :*
Brzark Information Systems (P) Ltd
112 Humayun Pur
(Near N.C.C.Office)
Safdarjung Enclave,
New Delhi - 110029, India
Fax : +91 11 617 7148
+91 11 618 2366

*Subscription Information :*
Subscription rates include air mail surcharge.
For detailed subscription information see inside.
Reduced price (two-third of the net price) for advanced subscription for first three years together.
Subscriptions may start from any time of the year.

*Payments may be made through :*
Demand Draft/Cheque/Money Order/ *Credit Card*
payable to
Brzark Information System (P) Ltd.
For further enquiry and for proforma invoice write to the Publisher.

*Advertisements :*
JISSI offers publication of selected advertisements from interested parties.
For rates and other conditions please contact the Publisher.

*Publishers :*
Prabhat Pandey
Bina Pandey

*Cover Design :*
Bibhas Dutta

*Price : Single Copy*
US $ 55.00 / Rs. 700.00

# Introductory Note

## *Relevance and Future of Sciento-Informetrics*

In the early eighties I started doing some research in different aspects of what was then called bibliometrics. The work in bibliometrics moved around three aspects - the scattering laws, the growth of literature or documents and citation and reference analyses. Soon I found - (i) there was no foundational integrity in the field; (ii) there was little predictive power in the so-called laws or theories of bibliometrics; (iii) both scientometricians and bibliometricans were trying to do almost the same thing without caring about each other; (iv) except for citation analyses (or reference analyses) there was little scope of practical application; (v) there was no standard norms or procedures for working with bibliographic populations and bibliographic data; (vi) Bradford's (and others, like Lotka's) laws could be made to fit almost all types of bibliographic datasets (it was rather disappointing for me as I realised Popper's criterion of falsifiability was not going to work); (vii) many persons and groups were working in the field without having a forum or without knowing each other; (viii) there was no way of replicating research result or comparing any two results.

I also did not know that others were having similar feelings. I was no mathematician. I therefore concentrated on the conceptual and empirical sides of the problems. I wrote a few articles. One of them - 'Philosophy of bibliometrics' was presented in a seminar in 1985 and was never published. But some peers like Prof Leo Egghe (whom I sent mimeographed copies) were responding to some of the ideas put in it. The other articles were also in some obscure places and could not be much visible. I had also some correspondence with the late Prof B C Brookes. Only a few months before his death he wrote me a long letter saying that till at that time nobody except a very few like him was interested in such conceptual issues. Within a decade however, things changed radically. Dr Egghe

and Rousseau organized the first international conference in Belgium. Scientometricians and informetricians came much closer and admitted that their research fields were not separated and independent. The term bibliometrics was considered as a branch of informetrics. Many persons were thinking about standardising terminology and sampling population of data.Yet I still have to encounter such questions as what relevance and practical uses this field has and what are its future dimensions. Everytime I face such a question I recall the following anecdotes. When Michael Faraday demonstrated his experiment of producing electric current by a moving magnet, a lady asked why he was doing all this. Faraday replied that as she would be nurturing a baby expecting it to become an adult, so he would do with his experiments. The Prime Minister of England asked him what benefit may come out of it. A smiling Faraday replied that in some future the products of the device may fetch the government revenue. On the other hand, when Galvani was asked why he was doing his experiments with electrical shocks to frog legs, he replied that he enjoyed the experiments. Are we nurturing our field only for art's sake or for a utilitarian goal or for both ? I am extremely hopeful for both.

My answer may sound childlike. In sciento-informetrics we have a possible means of quantitatively following man's socio cultural evolution and assessing future trends. Informetrics can ultimately lead to an understanding of information dynamics and the nature of information. [We should remember it took about a hundred years to clarify the nature of energy in physics]. I strongly believe that we need to develop an information physics for this and informetric laws should be integrated to it.

Haitun and others have shown that informational processes follow a distinct natural course which is called Zipfian in contrast to the other major course of natural events which is Gaussian. Brookes once discussed about something called *statistics of individuality*. But he never could develop the concept to its maturity. None else has been found interested in it.

My speculation is that both Gaussian and Zipfian are fundamental properties of nature - one represents entropy and the other represents information, the order out of chaos. Looking from a different perspective one is accounting for the future when no ranking is possible, because individuality is not asertained; the other is accounting for the past when a set of results is already at hand, individuality can be asertained and ranking is possible.

So, to my point of view, sciento-informetrics has the practical value in writing quantitative socio-historiography of events in socio-cultural evolution, and it has the pedagogic future of the highest nature in supplying the foundation for a future information physics.

***Subir K Sen***

# *Analysis of Social Relations In Coauthorship Networks*

**Hildrun Kretschmer**
Borgsdorfer Str. 5
D-16540 Hohen Neuendorf, Germany

**In literature a number of empirical findings report on the pattern of social relations which in general arise more frequently between similar persons (*Birds of a feather flock together*). But with increasing distance the preference between the persons is diminishing. Here similarity can be predicated on a variety of characteristics. This study is to document that the number of publications of authors that appeared in the journal *Zeitschrift fuer Sozialpsychologie* within the period from 1970 to 1993 has qualified as such a characteristic. In this case the distance between the authors is commensurate with the logarithm of the number of publications. The relative frequency of observed coauthorships between authors with *i* and *j* publications per author serves as a non-linear function of the difference between the logarithms of *i* and *j*. This non-linear function is also validly applicable to other coauthorship networks, such as to physicists.**

***Key Words : Social interaction; graduated structural parameter; coauthorship network; non-linearity; bibliometrics.***

## 1. Preference Given to Social Relations Among Similar Persons

Several sociopsychological and sociological studies have revealed that social relations are established more frequently than by chance among similar persons. Similarity was predicated here on an array of most diversified characteristics. For instance, there is a number of studies that refer to the idea of friendship between groups of persons, who were classified for their level of education, Marsden [5], Wolf [7], or there are studies on marriages that were entered into between persons of different religious affiliation [5]. However, age, sex, general approach to life, etc. can also serve as additional characteristics.

Just as characteristics can vary, relations can equally be varying and changing, e.g. long-term relations such as marriage and friendship, or short-term relations, such as the relationship established by a 13-year-old boy in a five-week summer camp, Feger [2].

In many studies the subject of investigations was *only one* definite relation that was singled out together with *only one* characteristic used to classify persons. However, there are considerations suggesting that several characteristics should be integrated into the study and that statistical correlations between these properties should be taken account of, Wolf [7].

In literature, Hofstätter [3] referred already to the fact that in groups with hierarchies the distances between members of equal rank were smaller than between member of different rank, or differently put : relations between members of equal rank are established more frequently than between members of unequal rank.

In 1974 Blau [1] developed a *general concept on structural parameters of groups.* Here the question is to identify the characteristics of individuals in a population that serves the organization due to relations. If the individuals in a population are grouped, or classified respectively according to these characteristics, a definite pattern of these relations among these groups will be

obtainable, which is commensurate with the patterning described by Hofstätter. However, these characteristics cannot be classified as structural parameters in those cases when such a pattern is not obtainable, i.e. when relations among groups were established according to a random principle. In the event of the existence of a structural parameter in line with Blau's conception the specific aspect of the patterning of relations is that insiders will be given preference compared to outsiders, cf. Hofstätter (insiders are members of equal rank). If the existing hierarchy should fail to give preference to 'outsider groups' (e.g. religious affiliation), then it would be a nominal parameter.

On the other hand, according to Blau, a graduated structural parameter is a variable that is to categorize persons of a population in terms of their preference.

When going beyond Blau's ideas and proceeding from an 'Unidimensional Social Distance Model' it is assumed by Marsden [5, p.21] : '*...that the likelihood of social intercourse between persons in groups is an inverse function of the distance between those groups along a single dimension*'. In this context Marsden described the above-mentioned friendships as if being based on classification of persons according to their level of education. Preference was given to insiders and not to outsiders and, additionally, there was also an inverse relation between status distance and the preference of contacts between persons.

This paper is aimed to find out to what extent Marsden's assumption is reflected in the *Zeitschrift fuer Sozialpsychologie*. It is to be verified whether or not the number of publications by authors in this journal is a graduated structural parameter in Blau's sense, which results in a definite pattern of coauthorship network, and it is to be examined whether, according to Marsden, the relative frequency of coauthorships Y (by analogy to the 'likelihood of social intercourse') is an inverse function of the distance X along the dimension of the 'number of publications'

$$Y = f(X) \qquad (1)$$

The number of publications per author is coupled with different criteria :

(i) Ability
(ii) Professional recognition
(iii) Endurance
(iv) Social rank
(v) Communication
(vi) Integration into a team

and with a great many other criteria. Who will collaborate with whom can thus be influenced to a varying degree by these criteria or also by still other mechanisms. But here, unlike in laboratory experiments, it is not intended to single out a definite mechanism from all the others, but rather the totality of phenomena as an entity, i.e. the dependence of the relative frequency of coauthorships from the graduated structural variable 'number of publications' as such.

## 2. The importance of the Variables X and Y Within the Co-authorship Network and Hypothesis

The relative frequency of the coauthorships Y is the ratio of the observed coauthorships between scientists to the statistically expected ones on the condition that coauthorships happen to be established by chance independently of the number of their publications. A more detailed discussion on it is found in chapter 3 on the methods.

What is hidden behind the distance X ? Is the distance between a scientist with i = 50 publications and a scientist with j = 20 publications equal to i - j = 30 ?

Psychophysics reports on the transformation of external variables into perception or awareness either in the form of a logarithmic function (Fechner) or in the form of a power function (Stevens). This applies to both physical (e.g. brightness) and social variables (e.g. rank). Which of the two transformations becomes evident will be dependent upon the methods in the experiment. While Fechner in his experiment used the indirect method, Stevens made use of the direct one.

Moreover, the relative frequency of coauthorships appears to be a reaction to the subjective distance in terms of the dimension 'number of publications'. In contract to laboratory experiments there is a process happening in reality

instead of a method stipulated by the head of the laboratory.

The physicist and science historian Derek John de Solla Price presumed in 1963 that the logarithm of the number of publications was of importance, p.49-50 :

'*...the number of publications is not a linear additive measure of productivity in the way required for Gaussian distributions ... One feels intuitively that the step from three papers to six is similar to hat from 30 to 60 rather than from 30 to 33. Because of all this it is reasonable to suggest that we have there something like the approximate law of Fechner or Weber in experimental psychology, wherein the true measure of the response is taken not by the magnitude of the stimulus but by its logarithm; we must have equal intervals of effort corresponding to equal ratios of number of publications*', with the independent variable not being the external, but the internal : 'effort'.

As a result, the abilities of a scientist, e.g. (as the internal variable) or his rank (as the external variable) would correlate with the logarithm of the number of his publications rather than with the mere number of publications.

*As to the question 'Who with whom' the difference between the logarithms of the number of publications by scientists would be of interest.*

Thus the distance X between the two above-mentioned scientists would be

$$X = \ln i - \ln j = \ln 50 - \ln 20 \qquad (2)$$

Derived from Marsden's conjecture (cf. quotation) and from the laws of psychophysics it is examined whether the following hypothesis, initially generally formulated, is valid in the bibliography of the *Zeitschrift fuer Sozialpsychologie* :

$$Y = f(X) = f(\ln i - \ln j) \qquad (3)$$

With i, j - number of publications

For a better understanding this hyothesis shall be specified in greater detail after the chapter on methods. Empirical evidence on the significance of the logarithm of the number of publications in coauthorship networks was adduced by Kretschmer [4].

## 3. Methods

The relative frequency of coauthorships is expressed by the ratio between the observed and the statistically expected coauthorships between scientists. The statistically expected value is the product of the total sum of coauthorships in the bibliography and of the probability, suggesting that coauthorships between scientists occur independently of the distribution of their publications.

As an example, independence would have to be understandable in such a way that any contact to be established by a highly prolific scientist with any other one would be arbitrary, i.e. a less productive colleague would be equally welcome to him as an average or highly productive one. In the event of independence there is no social stratification, i.e. the number of publications would be no structural parameter.

The method for the solution of this problem is to produce two matrices independently of each other and to correlate the values with each other. The first matrix would be the matrix of probabilities, suggesting that coauthorships emerge independently of the distribution of publications, matrix $p_{ij} = f_i \bullet f_j$, whereas the second matrix would be that of the actually observed coauthorships, matrix $C_{ij}$.

### *3.1 Distribution of Publications In A Bibliography, Matrix $p_{ij} = f_i \bullet f_j$*

Given that there is a bibliography (partly presented, names of authors A, B, C...) :

1. A, B
2. C
3. A
4. D, A, F
5. D, E
6. G, H

...

By applying the *normal count procedure* the number of publications is determined for each author : Every appearance of an author's name is counted (e.g. A three times : in the first article, in the third and the fourth one times only).

It should be noted here that the designation *article* is used with regard to a paper which was

jointly written by one or several authors, cf. 1., 2., 3., etc. articles in the biblography. On the other hand, the term 'publication' is referred to persons. Thus, in a bibliography without coauthors, the number of articles is equal to the number of publications. But in bibliographies with coauthors, the number of publications is greater than the number of articles. From the bibliography quoted in above example, table 1 can be drawn up.

**Table 1. Distribution of Authors and Publications**

| Number of Publications per Author | Number of Authors | Number of Publications of all Authors $A_i$ | Relative frequency of Publications of all Authors $A_i$ |
|---|---|---|---|
| i | $A_i$ | $i \bullet A_i$ | $f_i = i \bullet A_i / \sum_j j \bullet A_j$ |
| 1 | 167 | 167 | 0.30926 |
| 2 | 64 | 128 | 0.23704 |
| 3 | 39 | 117 | 0.21667 |
| 4 | 27 | 108 | 0.20000 |
| ... | ... | ... | ... |
| 20 | 1 | 20 | 0.03704 |

$\sum_j j \bullet A_j = 540$

On assumption of independence the probability $p_{ij}$ that one publication of the authors' group with i-publications per author coincides with one publication of the authors' group with j-publications per author, is equal to $f_i \bullet f_j$, see table 2.

$$p_{ij} = f_i \bullet f_j \qquad (4)$$

**Table 2. Matrix 1 : Matrix of $p_{ij} = f_i \bullet f_j$**

| i/j | 1 | 2 | ... | 1 | ... |
|---|---|---|---|---|---|
| 1 | $f_1 \bullet f_1$ | $f_1 \bullet f_2$ | ... | $f_1 \bullet f_1$ | ... |
| 2 | $f_2 \bullet f_1$ | $f_2 \bullet f_2$ | .. | ... | ... |
| ... | ... | ... | ... | ... | ... |
| k | $f_k \bullet f_1$ | $f_k \bullet f_2$ | ... | $f_k \bullet f_1$ | ... |
| ... | ... | ... | ... | ... | .. |

### *3. 2 Matrix of Relations by Coauthorships Viewed From Each Individual Author to All the Other Authors, Matrix $C_{ij}$*

If the relations are recorded from the point of view of every individual author to all the other authors, then a symmetrical matrix is obtained. In the example, table 3, the number in brackets corresponds to the number of publications of the author that stands ahead (cf. partly presented bibliography above, articles with more than one author).

**Table 3. Matrix 2 : Example, see Bibliography Above.**

| | | | | |
|---|---|---|---|---|
| 1. | A(3), B(1) | i/j | 1 | 2 | 3 |
| 4. | D(2), A(3), F(1) | 1 | \|\| | \|\| | \|\| |
| 5. | D(2), E(1) | 2 | \|\| | | \| |
| 6. | G(1), H(1) | 3 | \|\| | \| | |

As an example, in the fourth article viewed by author D with i = 2, a relation exists with author A with j = 3 and with F with j = 1. From the point of view of A there is in the fourth article a relation with D(j = 2) and with F(j = 1). From the point of view of F(i = 1) there is another relation with D(j = 2) and with A(j = 3). Generally, it is the matrix of the observed coauthorship relations of each other to all the authors : $C_{ij}$, see table 4.

**Table 4. Matrix 3 : Matrix of $C_{ij}$**

| i/j | 1 | 2 | ... | l | ... |
|---|---|---|---|---|---|
| 1 | $C_{11}$ | $C_{12}$ | ... | $C_{1l}$ | ... |
| 2 | $C_{21}$ | $C_{22}$ | ... | ... | ... |
| ... | ... | ... | ... | ... | ... |
| k | $C_{k1}$ | $C_{k2}$ | ... | $C_{kl}$ | ... |
| ... | ... | ... | ... | ... | ... |

Total Sum of Coauthorships $T = \sum_i \sum_j C_{ij}$

### *3.3 Matrix of Coauthorship Relations On Condition of Independence of the Recorded Coauthorships Among Each Other, Matrix $C'_{ij}$*

Should articles be written by more than two authors, then apart from symmetry, all the relations $C_{ij}$ are no longer independent of each other.

In the fourth article of bibliography, for instance, after designation the author's D relation with i = 2 publications with author A with j = 3 publications and after designation the author's A relation (i = 3) with author F (j = 1), the relation of author D(i = 2) with author F(j = 1) is at the same time implied in it. As a prerequisite for the use of $Chi^2$ only those relations are recorded in matrix $C'_{ij}$ that are independent of each other, i.e. *only those relations between the always neighbouring authors;* that would mean in the upper example only between D and A and between A and F. Due to symmetry and addition to the use of inference statistics only half the relations is used in the matrix, i.e. *the data below* (or above) *ie main diagonal line and half of the data per each cell of the main diagonal line.* The analogy of the matrices' results with $C_{ij}$ and $C'_{ij}$ is shown : the hypotheses apply to both types of matrices, see table 5.

**Table 5. Matrix 4**

| | Matrix of $C'_{ij}$ from the Example Above. | | | | Date Used for Inference Statistics | | |
|---|---|---|---|---|---|---|---|
| i/j | 1 | 2 | 3 | i/j | 1 | 2 | 3 |
| 1 | \|\| | \| | \|\| | 1 | \| | | |
| 2 | \| | | \| | 2 | \| | | |
| 3 | \|\| | \| | | 3 | \|\| | \| | |

### *3.4. Classification of Data*

Very comprehensive bibliographies are available, for example with more than 50 publications per authors. In an attempt to avoid statistical fluctuations, the data are classified. For analyzing the pattern of coauthorship networks the variable *Difference between the logarithms of the number of publications of scientists*(Scientists with *i* and scientists with *j* publications per author).

$$X = \ln i - \ln j \qquad (5)$$

is decisive.

Therefore, the cells of a matrix are grouped according to this principle, i.e. all cells with the same difference X = ln i - ln j are put together.

*Classification*

$c_l$ - class width

$X_m$ - class centre

Data of the size X are assigned to $X_m$, if the following is valid :

$X_m - c_l/2 \leq X < X_m + c_l/2$

$X_m$ (m = -s,...-2,-1,0,1,2,...s)

2s + 1 - number of classes in the matrix.

In this as well as in other studies the class width was fixed to $c_l$ = 0.6931, since the distance between an author with the smallest number of publications, i.e. with j = 1 and an author with the following higher number i = 2 publications is X = ln 2 - ln 1 = 0.6931, see fig. 1.

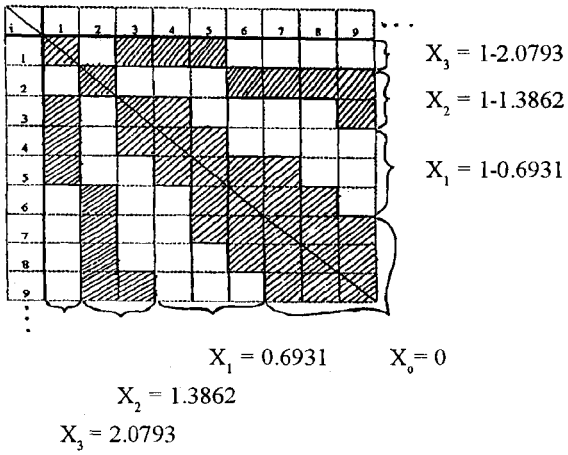



**Fig. 1. Matrix 5 : Grouping of the Cells of a Matrix According to X = ln i - ln j with $c_1 = 0.6931$**

Both for the matrix 1 with $f_i \cdot f_j$, see table 2, and for the matrix 3 with $C_{ij}$, see table 4, or the corresponding matrix with $C'_{ij}$ it is possible for each class m to assertion the adequate sum of data, i.e., $\sum_{Xm} f_i \cdot f_j$ or $\sum_{Xm} C_{ij}$ or $\sum_{Xm} C'_{ij}$ see table 6.

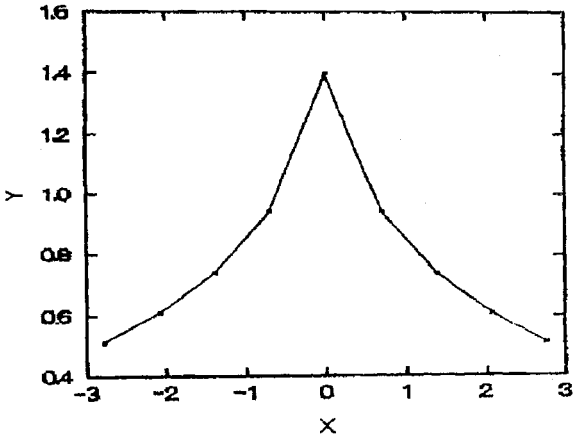
**Table 6. Allocations of Data From Different Matrices**

| m | Xm | $\sum_{Xm} f_i \cdot f_j$ | $\sum_{Xm} C_{ij}$ | $\sum_{Xm} C'_{ij}$ |
|---|---|---|---|---|
| - s | ... | ... | ... | ... |
| ... | ... | ... | ... | ... |
| - 2 | ... | ... | ... | ... |
| - 1 | ... | ... | ... | ... |
| 0 | ... | ... | ... | ... |
| 1 | ... | ... | ... | ... |
| 2 | ... | ... | ... | ... |
| ... | ... | ... | ... | ... |
| s | ... | ... | ... | ... |

## 4. Prediction of Coauthorships on the Basis of Distribution of Publications and the Difference of Logarithms of Publication Number

### *4.1 Precisely Specified Hypotheses*

Fig. 2 shows one of the possible graphs that correspond to Marsden's assumption, quotation see above. The relative frequency of contacts (Y) reaches its highest value with the difference X = 0 (insider get preference, or 'Birds of a feather flock together'). Apart from it, the relative frequency decreases additionally with increasing difference |X|, i. e. there is an inverse relationship between the preference of contacts between persons and their distance to each other.




$\sum_{Xm} f_i \cdot f_j$ or $\sum_{Xm} C_{ij}$ or $\sum_{Xm} C'_{ij}$ see table 6.

**Fig. 2. Function Analogously to Marsden's Assumption**

A simple non-linear function is shown in hypothesis 1.

$$Y = \sum_{X} C_{ij} / \sum_{X} W_{ij} \quad (6)$$

$$W_{ij} = f_i \cdot f_j \cdot T \quad (7)$$

with

- $C_{ij}$ - Coauthorships observed (= Coincidence of a publication from the group of authors with i publications per author with one publication from the group of authors with j publications per author)
- $\sum_{x} C_{ij}$ - Sum of observed co-authorships between all pairs of groups with the same distance X.
- $W_{ij}$ - Statistically expected value.
- $f_i \cdot f_j$ - Probability concerning the chance coincidence of above mentioned two publications
- T - Total sum of coauthorships ($T = \sum_i \sum_j C_{ij}$)

$\sum_x W_{ij}$ - Sum of expected values with regard to the equal distance X.

$$X = \ln i - \ln j \quad (5)$$

*Hypothesis 1*

$$Y = \text{const} \cdot (|X| + c_1)^{\alpha} \cdot (c_2 c_1 - |X|)^{\beta} \quad (8)$$

with

$c_1$ - Class width

$c_2$ - (Number of classes + 1)/2

$\alpha, \beta$ - Parameters

const. - constant

*Hypothesis 2*

Hypothesis 2 is a transposition of the equationof hypothesis 1. Here predictability of coauthorships stands in the centre which is obtained from the distribution of publications and from the difference of logarithms of the number of publications

$$\sum_x C_{ij} = \text{const} \cdot (\sum_x W_{ij}) \cdot (|X| + c_1)^{\alpha} \cdot (c_2 c_1 - |X|)^{\beta} \quad (9)$$

$\sum_x C'_{ij}$ can be used instead $\sum_x C_{ij}$ of in both hypothesis.

## 5. Results of the Analysis Into Coauthorship Networks taken from the Zeitschrift Fuer Sozialpsychologie

An analysis was conducted into the bibliography of the *Zeitschrift fuer Sozialpsychologie* of the period 1970-1993. Altogether 674 articles were covered by an evaluative analysis. These articles were graduatedly sorted according to the number of authors (or coauthors resp.), see table 7 and figure 3.

**Table 7. Distribution of Articles of the *Zeitschrift fuer Sozialpsychologie* from 1970-1993 According to their Number of Authors per Article**

| Number of Authors or Coauthors resp. per Article | Number of Articles |
|---|---|
| 1 | 427 |
| 2 | 183 |
| 3 | 43 |
| 4 | 13 |
| 5 | 5 |
| 6 | 0 |
| 7 | 2 |
| 11 | 1 |

SUM: 674

By analogy to table 1 the distribution of authors and publications was ascertained from the bibliography, cf. table 8 and figure 4.

By multiplying the relative frequencies with each other, by analogy to matrix 1, table 2, the matrix of probabilities $p_{ij} = f_i \cdot f_j$ was established, namely that a publication of the group of authors with *i* publications per author coincides by change with the publication of a group of authors with *j* publications per author. For example, $p_{12}$ = 0.45898·0.17578 = 0.08068. The complete portrayal of the matrix can be dispensed with due to its

scope. If nevertheless somebody should be interested in the data, they can, as described, determined easily from table 8. The matrix of coauthorships $C_{ij}$ is shown in table 9 and the matrix of $C'_{ij}$ in table 10.

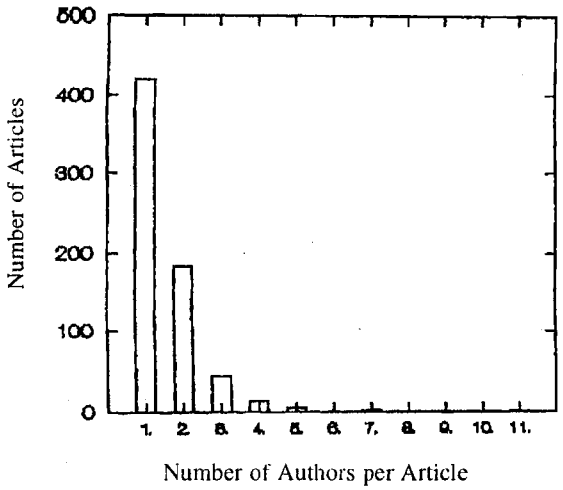



**Fig. 3. Distribution of Articles According to the Number of their Authors or Coauthors respectively**

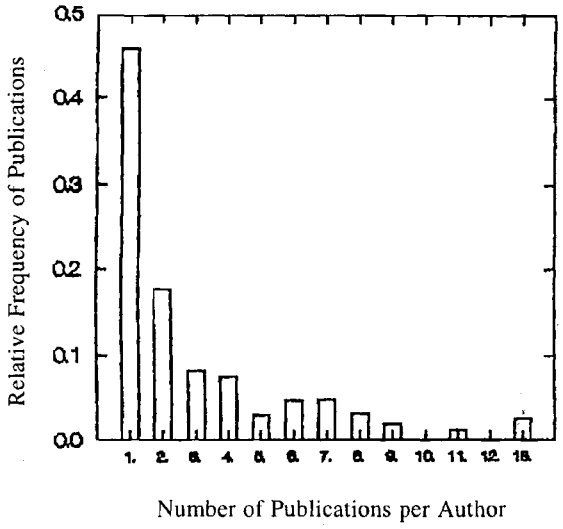



**Fig. 4. Relative Frequency of Publications of All Authors $A_i$ with $i$ Publications per Author**

**Table 8. Distribution of Authors and Publications of the Zeitschrift fuer Sozialpsychologie from 1970-1993**

| Number of Publications per Author | Number of Authors | Number of Publications of all Authors $A_i$ | Relative Frequency of Publications of all Authors $A_i$ |
|---|---|---|---|
| i | $A_i$ | $i \cdot A_i$ | $f_i = i \cdot A_i / \sum_j j \cdot A_j$ |
| 1 | 470 | 470 | 0.45898 |
| 2 | 90 | 180 | 0.17578 |
| 3 | 28 | 84 | 0.08203 |
| 4 | 19 | 76 | 0.07422 |
| 5 | 6 | 30 | 0.0293 |
| 6 | 8 | 48 | 0.04688 |
| 7 | 7 | 49 | 0.04785 |
| 8 | 4 | 32 | 0.03125 |
| 9 | 2 | 18 | 0.01758 |
| 10 | 0 | 0 | 0 |
| 11 | 1 | 11 | 0.01074 |
| 12 | 0 | 0 | 0 |
| 13 | 2 | 26 | 0.02539 |

$$\sum_j j \cdot A_j = 1024$$

**Table 9. Matrix 6 : Matrix of $C_{ij}$ of the *Zeitschrift fuer Sozialpsychologie***

| i/j | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 416 | 72 | 56 | 20 | 7 | 18 | 10 | 9 | 1 | 0 | 2 | 0 | 8 |
| 2 | 72 | 36 | 22 | 15 | 1 | 2 | 6 | 0 | 1 | 0 | 4 | 0 | 2 |
| 3 | 56 | 22 | 10 | 8 | 0 | 3 | 5 | 0 | 1 | 0 | 0 | 0 | 1 |
| 4 | 20 | 15 | 8 | 0 | 0 | 0 | 5 | 1 | 2 | 0 | 0 | 0 | 2 |
| 5 | 7 | 1 | 0 | 0 | 0 | 5 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 6 | 18 | 2 | 3 | 0 | 5 | 0 | 2 | 6 | 0 | 0 | 0 | 0 | 0 |
| 7 | 10 | 6 | 5 | 5 | 0 | 2 | 0 | 2 | 0 | 0 | 0 | 0 | 1 |
| 8 | 9 | 0 | 0 | 1 | 0 | 6 | 2 | 0 | 2 | 0 | 0 | 0 | 4 |
| 9 | 1 | 1 | 1 | 2 | 0 | 0 | 0 | 2 | 0 | 0 | 0 | 0 | 0 |
| 10 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 11 | 2 | 4 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 12 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 13 | 8 | 2 | 1 | 2 | 0 | 0 | 1 | 4 | 0 | 0 | 0 | 0 | 0 |

T = 1074

### *5.1 Review of Hypothesis 1*

A multivariate regression analysis is conducted for :

$$\ln(\Sigma C_{ij}/\Sigma W_{ij}) = \text{const} + \alpha . \ln(|X| + c_1) + \beta \bullet \ln(c_1 c_2 - |X|) \qquad (10)$$

with $c_1 = 0.6931$, as discussed in section 3.4 and following from it with $c_2 = 5$ and with $W_{ij} = f_i \cdot f_j \cdot T$, or $W'_{ij} = f_i \cdot f_j \cdot T'$, cf. table 11. The statistical procedures are only applied to the classes 0 to s(s = 4) due to symmetry, and they are applied only to half the observed values (*) and to half the expected values (**) for m = 0, as indicated in table 11, i.e. to $\Sigma C^*_{ij}/2 = 256$, $\Sigma W^{**}_{ij}/2 = 160.1$, $\Sigma C'^*_{ij}/2 = 159$, $\Sigma W'^{**}_{ij}/2 = 104.02$.

**Table 10. Matrix 7 : Matrix of $C'_{ij}$ of the *Zeitschrift fuer Sozialpsychologie***

| i/j | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 234 | 43 | 28 | 13 | 5 | 17 | 10 | 8 | 0 | 0 | 2 | 0 | 7 |
| 2 | 43 | 34 | 13 | 12 | 1 | 2 | 5 | 0 | 1 | 0 | 4 | 0 | 2 |
| 3 | 28 | 13 | 10 | 5 | 0 | 3 | 1 | 0 | 1 | 0 | 0 | 0 | 1 |
| 4 | 13 | 12 | 5 | 0 | 0 | 0 | 3 | 1 | 1 | 0 | 0 | 0 | 2 |
| 5 | 5 | 1 | 0 | 0 | 0 | 5 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 6 | 17 | 2 | 3 | 0 | 5 | 0 | 1 | 6 | 0 | 0 | 0 | 0 | 0 |
| 7 | 10 | 5 | 1 | 3 | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 1 |
| 8 | 8 | 0 | 0 | 1 | 0 | 6 | 1 | 0 | 2 | 0 | 0 | 0 | 3 |
| 9 | 0 | 1 | 1 | 1 | 0 | 0 | 0 | 2 | 0 | 0 | 0 | 0 | 0 |
| 10 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 11 | 2 | 4 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 12 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 13 | 7 | 2 | 1 | 2 | 0 | 0 | 1 | 3 | 0 | 0 | 0 | 0 | 0 |

T' = 698

**Table 11. Results of the Analysis of the Coauthorship Network of *Zeitschrift fuer Sozialpsychologie***

| m | $X_m$ | $\Sigma f_i \cdot f_j$ Xm | $\Sigma C'_{ij}$ Xm | $\Sigma W_{ij}$ Xm | $\Sigma C'_{ij}$ Xm | $\Sigma W'_{ij}$ Xm |
|---|---|---|---|---|---|---|
| 0 | 0 | 0.29805 | 512* | 320.11** | 318* | 208.04** |
| 1 | 0.6931 | 0.14546 | 131 | 156.23 | 82 | 101.53 |
| 2 | 1.3862 | 0.11858 | 100 | 127.35 | 62 | 82.77 |
| 3 | 2.0793 | 0.07528 | 42 | 80.85 | 39 | 52.55 |
| 4 | 2.7724 | 0.01165 | 8 | 12.52 | 7 | 8.13 |

The multiple correlation coefficient $R$ is given. As for the figures of the empirical and theoretical curves (estimated values) in Figure 5 and all classes from *-s* to *s* are included. The left pattern of coauthorship is taken from matrix $C'_{ij}$ and the right one from matrix $C_{ij}$.

### *5.2 Review of Hypothesis 2 :*

The non-linear regression with loss function to minimize the sum :

$$V_{min} = \Sigma \frac{(\text{Empirical value} - \text{Estimated value})^2}{\text{Estimated value}} \qquad (11)$$

is applied, see table 12. As for the (half-page) matrix $C'_{ij}$ i.e. the values in table 12, the use of Chi² is possible.

$V_{min}(C'_{ij}) = \chi^2 = 0.7781 > \chi^2_{(50\%)}$ with df 2.

This case is an average adaptation of empirical values to the theoretical distribution.

**Table 12. Comparison of theEmpirical Data with the Estimated Values**

| | Matrix $C'_{ij}$ | | Matrix $C_{ij}$ | |
|---|---|---|---|---|
| m | $\Sigma C'_{ij}$ Xm | Estimation of $\Sigma C'_{ij}$ Xm | $\Sigma C_{ij}$ Xm | Estimation of $\Sigma C_{ij}$ Xm |
| 0 | 159 | 157.124 | 256 | 253.290 |
| 1 | 82 | 87.762 | 131 | 142.201 |
| 2 | 62 | 58.545 | 100 | 87.615 |
| 3 | 39 | 37.892 | 42 | 48.227 |
| 4 | 7 | 8.063 | 8 | 7.423 |

**Minimum of the Loss Function $V_{min}$:**

$V_{min}(C'_{ij}) = 0.7781$ $\quad$ $V_{min}(C_{ij}) = 3.5107$

**Parameter**

$\alpha(C'_{ij}) = -1.0635$ $\quad$ $\alpha(C_{ij}) = -0.8873$

$\beta(C'_{ij}) = -0.8021$ $\quad$ $\beta(C_{ij}) = -0.2775$

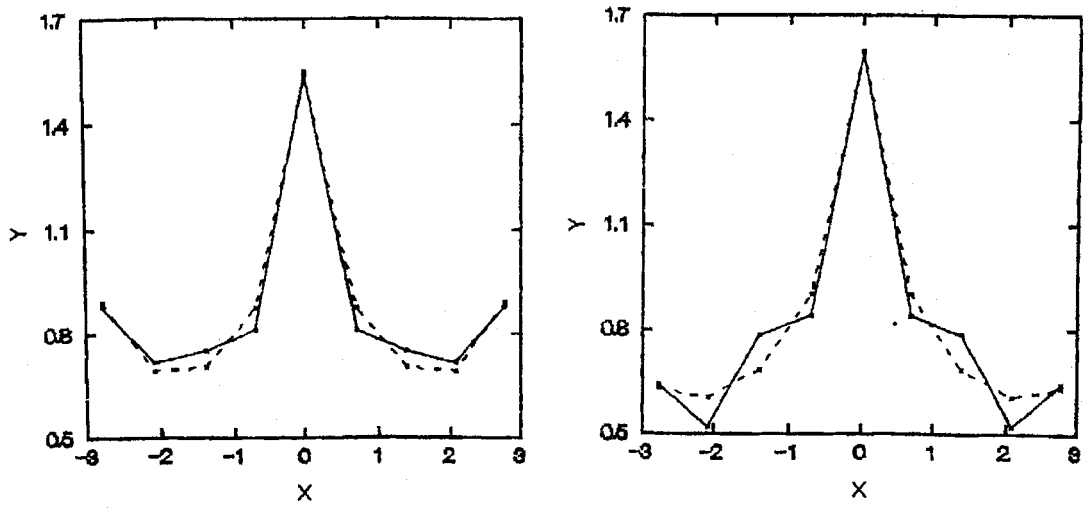
**Constant**

$const(C'_{ij}) = 2.77166$ $\quad$ $const(C_{ij}) = 1.6139$

For the half relations of the matrix in Table 10 the use of Chi² is possible.

$V_{min}(C'_{ij}) = \chi^2(2, N = 349) = 0.7781, p > .50$

It is an average adaption of empirical values to the theoretical distribution.




$R' = .981$

$p' = .0371$

$R = .967$

**Fig. 5. Coauthorship Pattern from the Bibliography of the *Zeitschrift fuer Sozialpsychologie* from 1970-1993 T = 1074. [Empirical Data (—) In Comparison with Estimated Values (- - -), the Left Pattern is Based on the Matrix $C'_{ij}$ and the Right One on the Matrix $C_{ij}$]**

**Table 13. Results of the Analyses of Coauthorship Networks of the *Zeitschrift fuer Sozialpsychologie* (A), *Quantized-Hall-Effect* (B), *and Monte-Carlo-Simulations...* (C), $\sum_{Xm} C_{ij}$, $\sum_{Xm} W_{ij}$ and Sums $\sum^{C}_{A\,Xm} \sum C_{ij}$, $\sum^{C}_{A\,Xm} \sum W_{ij}$**

| m | $\sum_{Xm} C_{ij}$ (A) | $\sum_{Xm} W_{ij}$ (A) | $\sum_{Xm} C_{ij}$ (B) | $\sum_{Xm} W_{ij}$ (B) | $\sum_{Xm} C_{ij}$ (C) | $\sum_{Xm} W_{ij}$ (C) | $\sum^{C}_{A\,Xm} \sum C_{ij}$ | $\sum^{C}_{A\,Xm} \sum W_{ij}$ |
|---|---|---|---|---|---|---|---|---|
| 0 | 256 | 160.1 | 426 | 224.2 | 229 | 152.9 | 911 | 537.2 |
| 1 | 131 | 156.2 | 360 | 384.9 | 295 | 250.0 | 786 | 791.1 |
| 2 | 100 | 127.4 | 206 | 294.4 | 181 | 210.2 | 487 | 632.0 |
| 3 | 42 | 80.8 | 109 | 172.8 | 89 | 133.3 | 240 | 386.9 |
| 4 | 8 | 12.5 | 37 | 61.6 | 38 | 85.6 | 83 | 159.7 |

For m = 0 both the observed and the estimated values are half

## 6. Discussion

It was possible to confirm the hypotheses, i.e. the lawful regularities of social psychology are mirrored in the scientific community of authors contained in the *Zeitschrift fuer Sozialpsychologie.*

The question arises whether this reflection of social psychology is only discernible in adequate journal, i.e. in the *Social Sciences*, or whether it is also observable in quite different disciplines of science, e.g., in physics. It might be assumed that in *Hard Sciences* the scientific process of work runs according to strictly logical rather than social points of view, which is also generally believed by the representatives of these disciplines of science.

Therefore, for reasons of comparison, the patterns of coauthorship of two bibliographies from physics are shown, cf. figures 6 and 7, together

with the tables 13 and 14. The hypothses 1 and 2 with $c_1 = 0.6931$ and $c_2 = 5$ are equally verified, as the bibliography of the *Zeitschrift fuer Sozialpsychologie.*

Since the same non-linear function is valid in different scientific communities the currently used formulation of question can even be extended : *Is the given non-linear function - or even a modified version of it - principally valid in* ***Scientific Communities*** *and does this function, in consequence, also apply to a summary of data drawn from several different bibliographies?* In addition to the analysis of individual bibliographies a summary was made per each class of the observed and the statistical values of expectation covering all three bibliographies A, B and C, cf. table 13 :

and table 14 : $\sum_{A\ Xm}^{C} \Sigma C'_{ij}$, $\sum_{A\ Xm}^{C} \Sigma W'_{ij}$

This new *summarized* bibliography was also evaluated, cf. figure 8. It was again possible to confirm the hypotheses, and this time even with a higher statistical significance than in case of the individual bibliographies.

**Table 14. Results of the Analyses of Coauthorship Networks *Zeitschrift fuer Sozialpsychologie* (A), *Quantized-Hall-Effect* (B) and *Monte-Carlo-Simulations* ...(C)**

| m | $\Sigma C'_{ij}$, (A) Xm | $\Sigma W'_{ij}$ (A) Xm | $\Sigma C'_{ij}$, (B) Xm | $\Sigma W'_{ij}$ (B) Xm | $\Sigma C'_{ij}$, (C) Xm | $\Sigma W'_{ij}$ (C) Xm | $\sum_{A\ Xm}^{C} \Sigma C'_{ij}$, | $\sum_{A\ Xm}^{C} \Sigma W'_{ij}$, |
|---|---|---|---|---|---|---|---|---|
| 0 | 159 | 104.02 | 222 | 104.82 | 68 | 33.45 | 449 | 242.29 |
| 1 | 82 | 101.53 | 159 | 179.95 | 60 | 54.68 | 301 | 336.16 |
| 2 | 62 | 82.77 | 88 | 137.62 | 26 | 45.98 | 176 | 26.37 |
| 3 | 39 | 52.55 | 45 | 80.8 | 14 | 29.16 | 98 | 162.51 |
| 4 | 7 | 8.13 | 18 | 28.81 | 14 | 18.73 | 39 | 55.67 |

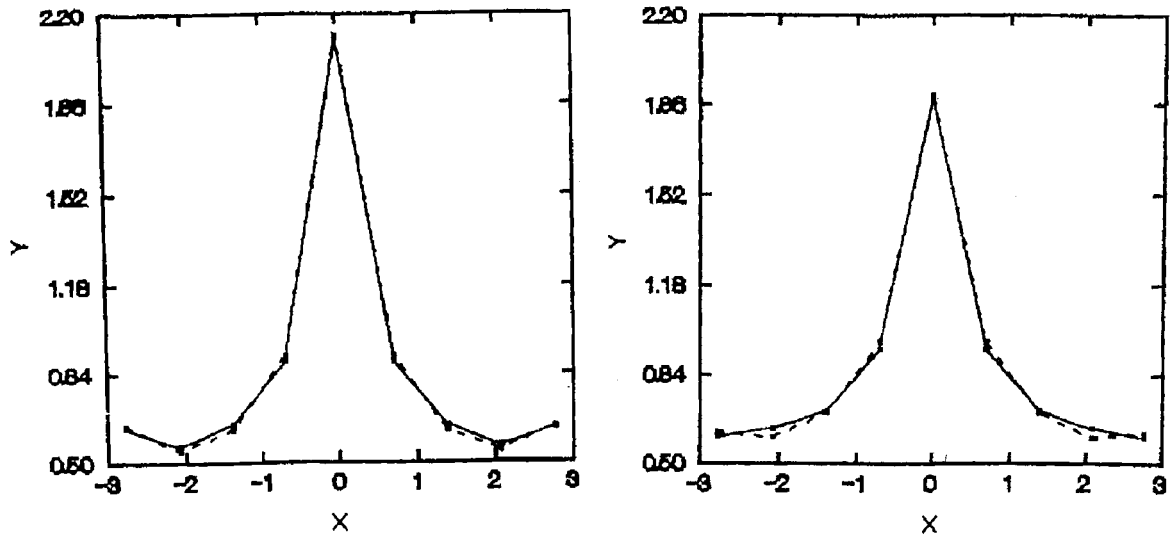
For m = 0 both the observed and the estimated values are half

In sum, it can be stated that the scientists, obviously guided by a scientific process of work, follow an interrelated pattern of relations, which appers to have gained validity in terms of diversified types of relations and characteristics (partly even in an animal kingdom), as expressed in the proverb *Birds of a feather flock together.* Similarity may be related to a whole bundle of characteristics. Therefore, the structural variable in the sense of Blau : *Number of publications* is only a kind of *Tip of the ice - berg* rather than the only variable per se that sets off the merging pattern of relations.

It seems to be imaginable that young scientists, for example, prior to the onset of their first publication, build up patterns of relations in line with their skills, features and in terms of their similarities, with these features of personality possibly qualifying as the potential for their future number of publications. Thus, the described pattern of relations might later become empirically provable.

As indicated in a former section, the number of publications of a scientist is related to a variety of criteria which can influence the contacts between persons.

On the other hand, following the appearance of publications, the status of the scientist thatis obliquely related to the number of publications could exert influence on the pattern of relations. But by and large, the number of publications seems to be the appropriate and suitable focal point to reflect the bundle of characteristics that account for similarity.




$R' = .999$

$p' = .0019$

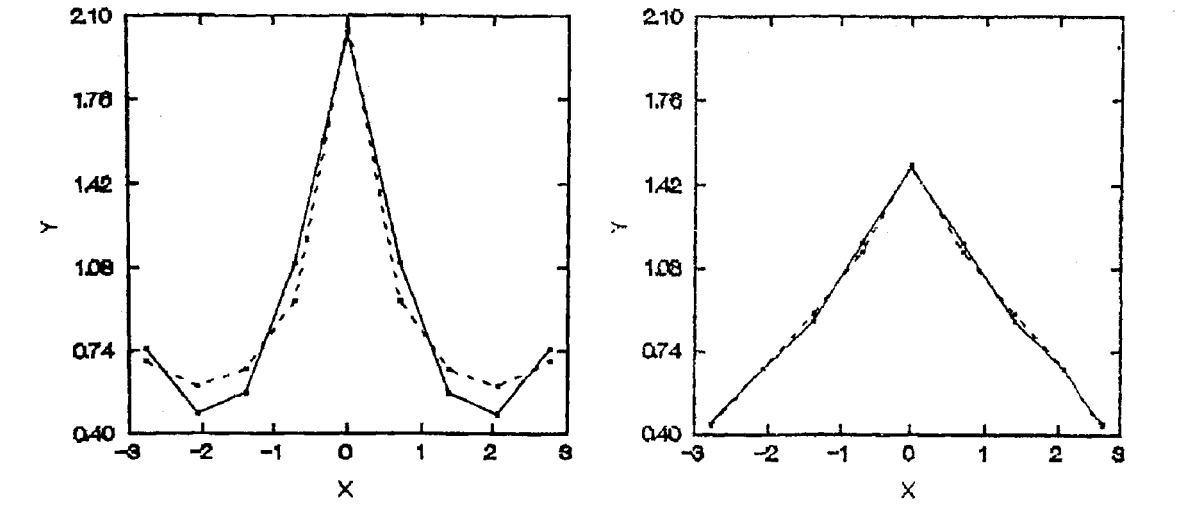
$V_{min}(C'_{ij}) = 0.2434$

$V_{min}(C'_{ij}) = 0.2434 = X^2(2, N = 532) = 0.2434, p > .75$

$R = .998$

$V_{min}(C_{ij}) = 0.6641$

**Fig. 6. Coauthorship Patterns From the Bibliography of the Quantized-Hall-Effect from 1980-1985, Data Base : INSPEC, T = 2276. [ Empirical Data (—) In Comparison with the Estimated Values (- - -), the Left Coauthorship Pattern is Based on the Matrix $C'_{ij}$ and the Right One on the Matrix $C_{ij}$ ]**




$R' = .972$

$p' = 0551$

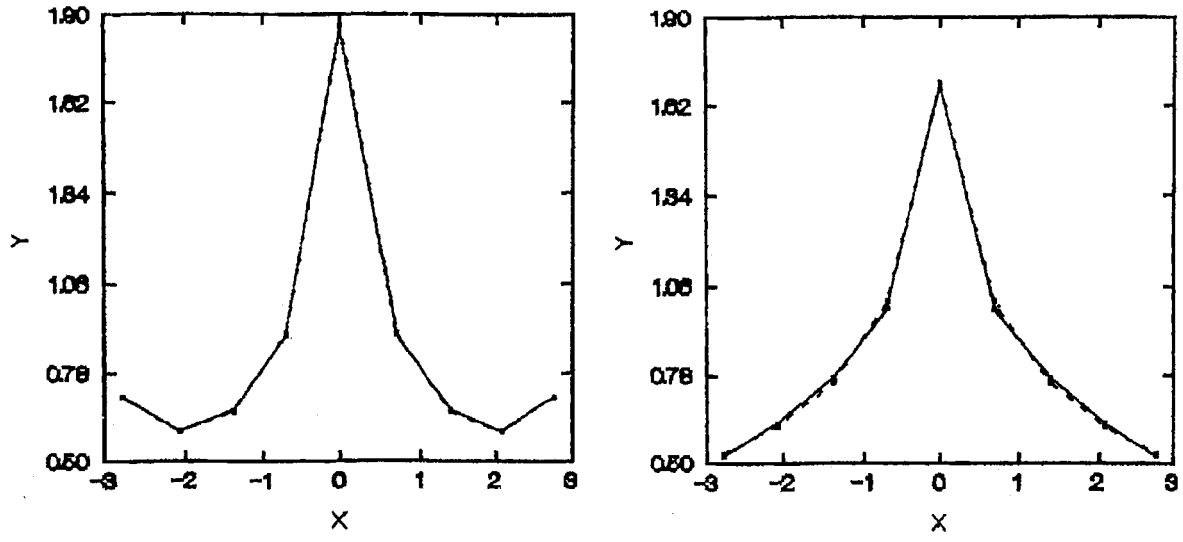
$V_{min}(C'_{ij}) = 2.6084$

$V_{min}(C'_{ij}) = X^2(2, N = 182) = 2.6084, p > .25$

$\dot{R} = .999$

$V_{min}(C_{ij}) = 0.5346$

**Fig. 7. Coauthorship Pattern from the Bibliography of the Monte-Carlo-Simulaton in Lattice-Field Theories from 1979-1984, Data Base : INIS, T = 1664. [ Empirical Data (—) In Comparison with the Estimated Values (- - -), the Left Pattern is Based from the Matrix $C'_{ij}$ and the Right One from the Matrix $C_{ij}$ ]**




$R' = 1.000$

$p' = .0003$

$V_{min}(C'_{ij}) = 0.0416 = X^2(2, N = 1063) = 0.0416, p > .975$

$R = .999$

$V_{min}(C_{ij}) = 0.5835$

**Fig. 8. Coauthorship Pattern of the *Summarized Bibliography*, T = 5014. [Empirical Data ( — ) In Comparison with the estimated values ( - - - ), the Left Coauthorship Pattern is Based on the Matrix $C'_{ij}$ and the Right One from the Matric $C_{ij}$ ]**

But to revert to the proverbs :

There is also a proverb with opposite meaning, such as *Opposites attract*. It should be possible to expect this pattern of relations to be found in coauthoship networks. In a paper that is presently in preparation it will be shown that also the second proverb is involved in above-indicated non-linear function. This pattern of relations will be shown to emerge also in scientific communities.

In conclusion, the question is important to ask whether the results of this paper could be taken as proof of the fact that the number of publications could serve as a criterion of evaluation of an individual scientist. I am afraid but this question must be answered with a clear *No*.

Just as it is impossible to make a safe prediction on an individual molecule, the number of publications of an individual scientist is no safe prophecy about the contacts to be established by him with other scientists. Although there is no exact prognosis on each individual molecule, there are lawful regularities existing on the entire system of molecules, for example, on the thermodynamics in a gas.

Analogous to it, structural developments should be reviewed in scientific communities : But the smaller the number of scientists under study, the lower is the predictability. This does not only apply to the individual scientist as the smallest entity, but also to the empirical values of small bibliographies which, like the *Zeitschrift fuer Sozialpsychologie*, are only loosely adapted to the non-linear function. The larger the bibliographies - isolated or summarized ones -, the more they approach Bernoulli's Law of Great Numbers concerning empirical and theoretical values, until their correlation will be 1.000.

Even if no findings can be drawn from this kind of science research in the interest of evaluating individual scientists, it will be increasingly more imperative in the present time marked by the merger of the world into a great systems to gain knowledge on the process of science as such and to identify regularities by analogy to the systems of the myriads of molecules.

## References


1. Blau, P. M. (1977). A macrosociological theory of social structure. *American Journal of Sociology*, **83**, 26-54.
2. Feger, H. (1987). Netzwerkanalyse in Kleingruppen : Datenarten, Strukturregeln und Strukturmodelle. In : F. U. Pappi (Hrsg.), Techniken der empirischen Sozialforschung, 1. Band : Methoden der Netzwerkanalyse ( S. 203-251). München : R. Oldenbourg Verlag GmbH.
3. Hofstätter, P. R. (1957) Gruppendynamik. Hamburg : Rowohlt.
4. Kretschmer, H. (1992). Significance of the logarithm of number of publications for the measurement of social stratification in coauthorship networks. In : I. K. Ravichandra Rao (Ed), Informetrics 91. Selected Papers from the Third International Conference on Informetrics, 9-12 August, 1991, Bangalore (289-331). Bangalore : Sarada Ranganathan Endowment for Library Science.
5. Marsden, P. V. (1981). Models and method for characterizing the structural parameters of groups. *Social Networks*, **3**, 1-27.
6. Price, D. J. de Solla. (1963). Little Science, Big Science. New York : Columbia Univrsity Press.
7. Wolf, Ch. (1996). Ein Simulationsmodell sozialer Netzwerke und einige unerwartete Einsichten zur Ähnlichkeit sozialer Beziehungen. Unveröff. Beitrag zur Tagung*Netzwerkanalyse*, eine gemeinsame Veranstaltung der Sektion *Methoden* und *Modellbildung und Simulation* der Deutschen Gesellschaft für Soziologie am 22. und 23. März 1996 in Köln.

# *Revealing the Hot Topics in Life Sciences Using Index to Scientific Reviews*

**Maria M Klavdieva**
Institute for the History of Science and Technology, Russian Academy of Sciences
Staropansky per. 1/5 103012 Moscow, Russia.

**The areas of breakthroughs in life sciences are determined using the publication of the relevant high-weight scientific reviews (citing 5 or more core-papers of this areas) as an indicator. Index of Scientific Reviews published by ISI is used as a source for selecting such reviews and areas (Research Front Specialties, or RFS). Disciplinary distribution of selected RFS is shown, which reflects the structure of the frontiers in modern science. Analysis of the geographical distribution of reviews showed the difference in both scientific policy and potential in different countries. Association of RFS with two or more major disciplines allows to cluster these disciplines on the principle of their interdisciplinary relations similar to the principle of co-citation.**

***Key-terms : Co-citation; Scientific Reviews; Index of Scientific Reviews; Research Front Specialty Index; Hot Topics; Disciplinary Distribution; High-weight Reviews; Topical RFS; Threshold of Supertopicality; Geographical Distribution; Interdisciplinary Science.***

## 0. Introduction

Among the goals of scientometrics are the elucidation of the current structure of research fields and assessment of the involvement of scientific communities in world research front. These approaches were reviewed in [1, 2]. Another important task for a scientometric study is revealing the "hot" problems and topics of research, these breakthrough fields which attracted most attention of the world scientific community at the moment. The aim of the present study was to determine the "hottest" topics of modern science using the scientometric approach independent of experts' opinion. This approach includes the usage of unique material of bibliographic editions and databases issued by the Institute for Scientific Information (Philadelphia, USA).

## 1. Studies of Research Fields Using *Index to Scientific Reviews*

The term "hot topics" infers those research fields that had produced the results, theoretical concepts and and/or new methods regarded by the international scientific community as important, promising and attracting attention of a great number of scientists.

The scientific output in life sciences is reflected in publications that can be divided into three groups corresponding to different levels of the development of a research field or a discipline [3]. The process of cognitive institutionalization of a new scientific specialty begins with the publication of original papers, letters to editor, abstracts, patents, etc., all of which have the same objective — to claim the priority of new knowledge. When this new knowledge is amassed sufficiently and being or seeming in agreement with some general pattern or idea, it becomes the subject of review articles. The next stage of maturation of a new specialty is marked by the appearance of monographs or collections of articles by the different authors dedicated to the same topic (for instance, symposium proceedings). This reflects the accumulation of a bulk of data (the knowledge not quite new, but

verified and validated), the separation of a special scientific community united by the work in the same special field and its claim for the acknowledgement of a new specialty by the outer, more wide scientific community. The final stage of the cognitive institutionalization of a discipline is marked by the publication of textbooks or special sections in the textbooks of larger disciplines. As a rule, these sources provide validated, verified information that is presented as an unambiguously determined fact or a regular pattern. This form of knowledge presentation is largely determined by the specific orientation of these publications towards students, and the knowledge consumed by students in the process of training is assimilated as an irrefutable truth.

We regard scientific reviews as a very important stage in the emergence of a new specialty. Directional sorting of information, interpretation and evaluation of a data are specific features of reviews. Quite a few important concepts were formulated for the first time in scientific reviews. The review covers the fields of extensive research activity with many researchers involved. Thus we can say that they may carry the germs of new scientific fields. Another specific feature of scientific reviews is that in most cases they are written and published "to order" on request of scientific community. It is common practice that reviews are written by the acknowledged scholars, the recognized authorities on the subject, on the editor's request. It means that the reviews satisfy the need of scientific communities in the cognitive integration of the new knowledge, its classification, generalization and expertise, and also "seal" the acknowledgement of its validity and importance. Thus, if several reviews devoted to the same topic are published simultaneously (during the same year), this would indicate that the international scientific comunity recognized this topic as highly important. This is even more so if these reviews appear in the journals associated with different disciplines. The latter case may indicate the recognition of the potential applicability of the new knowledge to the problems within different disciplines.

The technique suggested in the present paper is based on the use of the material provided by the *Index* of *Scientific Reviews* (ISR) published by the Institute or Scientific Information (ISI) since 1974. Beside the traditional author and corporate indexes, beginning with 1982 issue it contains a special section, Research Front Specialty Index (RFSI). *ISR* and *Compumath* — are the only widely available (at least in the Eastern Europe) sources of information on research front specialties (areas of intense research activity) or co-citation clusters.

The idea that the co-citation of two papers by the third paper reflects (or rather creates) the link between the first two was originally expressed in 1973 independently by Marshakova and Small [4, 5]. However, it is difficult to agree with the idea [6, pp.180-181] that the connection between the two works actually exists only when they are co-cited by a high number of citing documents, significantly exceeding mathematical expectation of such co-citation. We think that even a single citation of one document by another document cannot be regarded a random variable. In our opinion, citing is always the result of conscious choice determined by certain conceptual relationship between the two documents.

Co-citation creates invisible links between the papers allowing to aggregate them into clusters that provide the graphical representation of cognitive structure of a given research field. The core papers would comprise the nodes of these clusters and the links between them would be formed by the articles co-citing the pairs of core-articles, the more articles co-citing the core-papers, the stronger the bonds between these. Patterns of citation and co-citation change constantly, thus the corresponding clusters also change.

ISI has been identifying the clusters of core-papers representing the frontiers of research activity since the early 80s using their enormous *Science Citation Index* database, with the utmost goal to map science at different levels of aggregation. A cluster of current-year articles and the earlier publications that they consistently co-cited articles makes up an RFS — a specialty area of significant activity. Lists of RFS are compiled annually basing on the current co-citation patterns of the massifs of scientific publications. A special computer program developed by ISI extracts and names each specialty using words that appear in the titles of current papers citing into the specialty [7, 8, 9].

The cluster analysis is based on the co-cita-

tion frequency, which is the number of times two documents are cited together by current papers [7]. This association measure is normalized by dividing by the square root of the product of citation frequencies of the co-cited documents [10]. The clustering algorithm is called the single-link method and has the advantage of simplicity in application involving large files and disadvantage that clusters can be highly chained. To limit the chaining a cut-off in cluster size of 60 cited documents is used, as well as the variable co-citation level to increase the size of small clusters.

The strategy of cluster analysis for 1988 file is described in [11] :

(i) All the papers cited no less than 5 times a year were selected;

(ii) Of these, only highly cited papers were selected all linked together by at least one single citing and thus comprising one giant cluster — as a result 70000 papers were selected from 5 million papers;

(iii) The co-citation threshold was elevated gradually until manageable clusters were obtained (including 60 or fewer core-documents). As a result, about 9000 C1 (lowest level of iteration) clusters were obtained with about 50000 of originally selected 70000 core-documents. A list of sources citing the maximal number of core-documents was comprised for each cluster.

According of Garfield, ISR covers the reviews and original articles citing more than 50 sources, which are selected from more than 3000 of the most important scientific journals related to more than 100 areas of the life sciences. Garfield states that 3-4% of the total number of scientific publications are reviews [9]. It is also stated in the preface to ISR that all indexed documents are inspected by experts and meet the criteria of a review article. However, many of the publications indexed in the ISR can hardly be called reviews. Particularly, publications selected as reviews for ISR and related to such fields as theoretical physics, immunocytochemistry, theoretical biology, and others usually have very long lists of references being based on the analysis (or comparative analysis) of the material from numerous works performed by other authors. Studying the texts of such publications in the above mentioned scientific areas in order to check their nature. I found that quite often all "reviews" associated with the corresponding RFS are just typical experimental or theoretical articles. Citing numerous articles related to the subject, such articles naturally can be used a reviews in a sense of a source of information, but they lack other specific features of reviews and cannot be regarded as indicators in a sense considered herewith.

Beginning with the 1982 issue, ISR includes an RFSI section, which enables us to study the RFS using this Index. Some statistical data related to RFS and the sources included in ISR is presented in Table 1. It can be seen from this table that only 2000 RFS were included in ISR for 1982, more than 3000 RFS — for 1983, and beginning with 1984 — more than 6500 RFS yearly. The actual number of RFS is higher because RFSI in

**Table 1. Comparative Statistical Summary of 1982-1991 Semiannuals (From ISR - 1991)**

| Semi annual | 1982 | 1983 | 1984 | 1985 | 1986 | 1987 | 1988 | 1989 | 1990 | 1991 |
|---|---|---|---|---|---|---|---|---|---|---|
| TSI - I | 10740 | 13131 | 12654 | 15435 | 17172 | 17365 | 18588 | 19938 | 21285 | 23490 |
| - II | 15319 | 16071 | 15350 | 16438 | 16406 | 17034 | 18646 | 20364 | 21096 | 23230 |
| USI - I | 5931 | 7166 | 19265 | 12804 | 14411 | 14404 | 15834 | 17156 | 18500 | 20013 |
| - II | 8325 | 8827 | 12564 | 13009 | 13866 | 14520 | 15807 | 16906 | 18188 | 20195 |
| RFSN - I | 2228 | 3042 | 6200 | 6666 | 6668 | 6529 | 6570 | 6675 | 6473 | 6612 |
| - II | 2346 | 3102 | 6550 | 6671 | 6520 | 6493 | 6554 | 6634 | 6374 | 6460 |
| ANSI - I | 7.15 | 7.00 | 5.01 | 5.78 | 7.07 | 6.92 | 7.66 | 8.19 | 9.01 | 0.39 |
| - II | 9.96 | 8.15 | 5.75 | 5.85 | 6.90 | 6.99 | 7.46 | 7.93 | 8.71 | 10.45 |

TSI - Total source items USI - Unique source items citing RFS

RFSN - RFS names Ansi - Average number of source items per RFS name

cludes only these RFS, the core-documents of which were cited by at least one review or review-like article (RLA). Thus, about 9000 RFS were identified for 1988 and of these only about 6,500 (72%) were included in RFS.

The goal of our study was to construct the database of the "hot topics" of modern science for 1991, the first year of the last decade of the century. In the beginning, the relation of every RFS included in the first semiannual 1991 issue to one or more major disciplines was determined. In uncertain cases (especially in cases of interdisciplinary RFS, such as those which could be attributed to both physics and chemistry), the association of RFS with a discipline was defined according to the profile of the journals in which the corresponding reviews or RLA were published. The final evaluation was performed by the experts including the specialists in scientific information and the editors of the corresponding scientific journals. It was found that a great deal of RFS could be related to more than one major discipline. Most often these were medicine and biology, medicine and pharmacology, chemistry and physics.

It follows from the data presented in Table 2 that 32% of all RFS found in 1991 issue of ISR are related to medical sciences, 31,4% — to biological, 20,9% — to chemical, and 17,6% — to physical sciences, 6% — to pharmacology, 5% — to materials sciences, 4% — to geosciences, and 3,8% — to agriculture, diary and animal science.

An example of the disciplinary distribution of the 1979 SCI file clusters can be found in [6, p.208-209]. 61,0% of C1 clusters were related to biomedicine and biochemistry, 20,8% — to physics, 8,7% — to chemistry, 4,5% — to mathematics, 1,4% — to geology and 3,6% — to other fields (the total amount of clusters adds up to 100%). It is pointed out that such classification suggested by ISI is based on the nature of the predominating journals (the corresponding discipline) in which the core-documents of the clusters have been published. In [7] it is stated that clusters are distributed by disciplines defined more or less in the traditional sense. One can hardly agree with such one-to-one corespondence between the clusters and the disciplines. Moreover, putting such large complexes of sciences as "biochemistry and biomedicine" into one group while separating physics from chemistry seems artificial for one-to-one correspondence.

**Table 2. Relation of All RFS for the First 1991 Semiannual to Major Disciplines in Life Sciences.**

| Disciplines | No. of RFS |
|---|---|
| Medical sciences | 2116 |
| Biological science | 2078 |
| Chemical sciences | 1380 |
| Physical sciences | 1160 |
| Pharmacology | 392 |
| Materials sciences | 320 |
| Geosciences | 274 |
| Agriculture, dairy and animal science | 251 |
| Computer science, mathematics, statistics | 220 |
| Psychology | 204 |
| Biotechnology | 110 |
| Environmental sciences | 104 |
| Astronomy | 103 |
| Soil science | 75 |
| Food Science | 68 |
| Total No. of RFS identified for 1991 | 8243 |
| Total No. of RFS in the first 1991 semiannual | 6612 |

Every cluster in RFSI has a code number and a title. Below is the list of reviews or RLA citing at least one of the core-documents of this cluster, each characterized by the corresponding short bibliographical reference and the number of cluster core-documents cited by this review (the review "weight", Wsp). This feature was used as a criterion for selection of reviews. It reflects the involvement of works cited by this review in the cognitive structure of the corresponding RFS and thus may serve as an indicator of topicality of this review. Moreover, the analysis of the weight of reviews corresponding to certain RFS allows the assessment of this RFS [12]. It is important to point out that reviews indexed in ISR usually cite 1-3 core-documents. Statistical data on the clusterization of 1979 SCI file is given in [6]. It is pointed out that the mean C1 cluster consists of 5,38 documents and 48,3% of all clusters include only 2 documents. Thus we can say that if a review cites 5 core-documents of an RFS, not only this reflects that this review is highly relevant to this RFS, but this also means that this is a rather well developed RFS. We can say that the existence of high-weight

review reflects the high level of scientific activity in this field. Thus, we regard the appearance of scientific reviews with a rather high weight (citing 5 core-documents or more) an indicator of the advanced stage of development of a corresponding RFS. We should emphasize that if a review supposedly devoted to a certain RFS cites none of its key-documents, it does not mean that it is a "bad" review, it can only mean that the topic of this review claimed by its title (or in its headings) does not unequivocally correspond to its contents.

**Table 3. Disciplinary Distribution of Topical RFS.**

| Disciplines | No. of of FRS | No. of reviews per RFS |
|---|---|---|
| Medical sciences | 369 | 3.32 |
| Biological sciences | 277 | 4.35 |
| Chemical sciences | 173 | 2.38 |
| Physical sciences | 171 | 2.30 |
| Pharmacology and Pharmacotherapy | 142 | 3.60 |
| Materials sciences | 58 | 2.24 |
| Geosciences | 29 | 2.37 |
| Psychology | 27 | 2.29 |
| Agriculture, dairy and animal sciences | 25 | 2.64 |
| Environmental science | 18 | 1.66 |
| Astronomy | 16 | 1.81 |
| Computer science | 15 | 1.86 |
| Biotechnology | 10 | 2.0 |
| Food science | 6 | 2.33 |
| Soil science | 6 | 2.12 |
| Total | 766 | 3.08 |

In the beginning we selected from RFSI section of ISR-1991 (lst semiannual issue) all reviews and RLA with Wsp ≥ 5 RFS and all the corresponding RFS, and then checked whether these papers are reviews or just original articles with long reference lists. There exist many types of "Review" as well as the papers with "what is ...?", "Critical analysis of literature", "Overview of..." "Plenary lecture", "Development of ...", "...Coming of age..", "Advances in ...", "Survey of ...", "Update of ...", "Editorial", "Status update", "State of art", "Trends in ...", "Current trends in ...", "What's new in ...", "Past, present and future of ...", "New agents in ..." in their titles. A total of 2078 reviews was selected which cited core-documents of 766 RFS (called topical RFS) that were included in the database designed by S. Shapovalova. If all high-weight documents citing core-documents of an RFS cluster proved to be original articles, we did not include this RFS in our database. Selected topical RFS were classified by disciplines and subdisciplines. The disciplinary distribution of topical RFS is given in Table 3.

**Table 4. Association with the Different Disciplines of the Journals Used in ISR as the Sources for Reviews and Review-like Articles.**

| Discipline | ISR-1990 (II semi-annual | ISR-1991 II semi-annual) |
|---|---|---|
| Medicine | 783 | 780 |
| Biology | 740 | 802 |
| Chemistry | 260 | 256 |
| Physics | 231 | 230 |
| Geosciences | 125 | 136 |
| Pharmacology | 104 | 106 |
| Mathematics, Computer Sciences, statistics | 100 | 100 |
| Agriculture, dairy and animal science | 117 | 115 |
| Biotechnology | 23 | 24 |
| Materials sciences | 66 | 86 |
| Environmental science | 70 | 62 |
| Astronomy and astrophysics | 23 | 21 |
| Food science | 29 | 29 |
| Soil science | 12 | 11 |
| **Others** | 62 | 56 |

To determine whether there is a correlation between this distribution and the number of articles published within the corresponding disciplines, we studied the distribution of journals listed as sources in the ISR by major disciplines which roughly reflects the distribution of the bulk of articles. The results of these studies are presented in Table 4. We choose the second semiannuals of 1990 and 1991 because each second semiannual includes more sources than the first. Though most of the journals could be more or less unequivocally related to one of the major disciplines, some are clearly interdisciplinary (we are not speaking about the multidisciplinary journals like *Nature* or *Science*, etc.). We included such journals in all the

corresponding groups of journals. The distributions or 1990 and 1991 are similar.

To illustrate the relationship between the different disciplines we calculated the proportions of the journals related to the corresponding disciplines (from Table 4), as well as the proportions of RFS (from Table 2) and the topical RFS (from Table 3) expressed as percentage of the number of medical journals or RFS (or topical RFS), respectively. The results are given in Table 5. It is clear that the distribution of sciences by the journals differs significantly from their distribution by RFS and the topical RFS. The journal and RFS rates for biology are similar to these for medicine, but the topical RFS rate for biological sciences is 25% lower than that for medical sciences which indicates that in 1991 medical sciences generated more of the most important breakthroughs than biological sciences.

**Table 5. The Proportions of the Journals (A), RFS (B) and Topical RFS (C) Related to Different Disciplines Expressed as the Percentage of the Number of Medical Journals RFS and Topical RFS, respectively.**

| Discipline | A (from Table 4) | | B (from Table 2) | C (from Table 3) |
|---|---|---|---|---|
| | 1990 | 1991 | | |
| Medicine | 100% | 100% | 100% | 100% |
| Biology | 94.5 | 102.9 | 98.0 | 75.0 |
| Chemistry | 33.2 | 32.8 | 65.2 | 46.8 |
| Physics | 29.5 | 29.5 | 54.8 | 46.3 |
| Pharmacology | 13.3 | 13.6 | 18.5 | 38.5 |
| Materials sciences | 8.4 | 11.0 | 15.1 | 15.7 |
| Geosciences | 16.0 | 17.4 | 13.0 | 7.8 |
| Agrosciences | 14.9 | 14.8 | 11.9 | 6.7 |
| Environmental sciences | 12.8 | 12.8 | 10.4 | 4.0 |

The number of both physical and chemical journals is 3 times smaller than that of both medical and biological journals, but the proportions of physical and chemical RFS are almost twice as high, though the corresponding rates of topical RFS are lower. In geosciences, agriculture, dairy and animal science, mathematics, computer science and statistics, the RFS rate is a somewhat lower than the journal rate and the topical RFS rate is more than twice as low. The opposite tendency can be seen for pharmacology : the RFS rate and the topical RFS rates are 1,4 and 2,5 times higher than the journal rate. This is partly due to the fact that articles related to pharmacology are widely published in medical and also in chemical journals. On the other hand, many of the articles related to pharmacology are also related to medical sciences which generate the highest proportion of topical RFS.

The number of relevant topical RFS and the corresponding revies, as well as the mean number of high-weight reviews per RFS was calculated for each discipline. The last value was shown to be highest for the biological (4.3) and medical (3.2) sciences.

To determine the hottest (supertopical) RFS in different sciences we studied the distribution of topical RFS by the number of corresponding high-weight reviews (Table 6). From the graphs presented in Figures 1-5 it can be seen that the character of reviewing of the scientific literature varies in different disciplines. This kind of distribution is usually presented in logarithmic co-ordinates, which allows to determine the index of Zipfian distribution ($\alpha$) that can be calculated from the slope of asymptote approximating the distribution of high values of the variables in logarithmic co-ordinates [1]. Thus calculated, the values of $\alpha$ for the distributions of selected RFS by the number of corresponding high-weight reviews (total and related to medicine, biology, chemistry, and physics) amounted to $1.3 \pm 00.47$, $1.04 \pm 0.054$, $0.64 \pm 0.038$, $1.84 \pm 0.14$, and $2.24 \pm 0.17$, respectively.

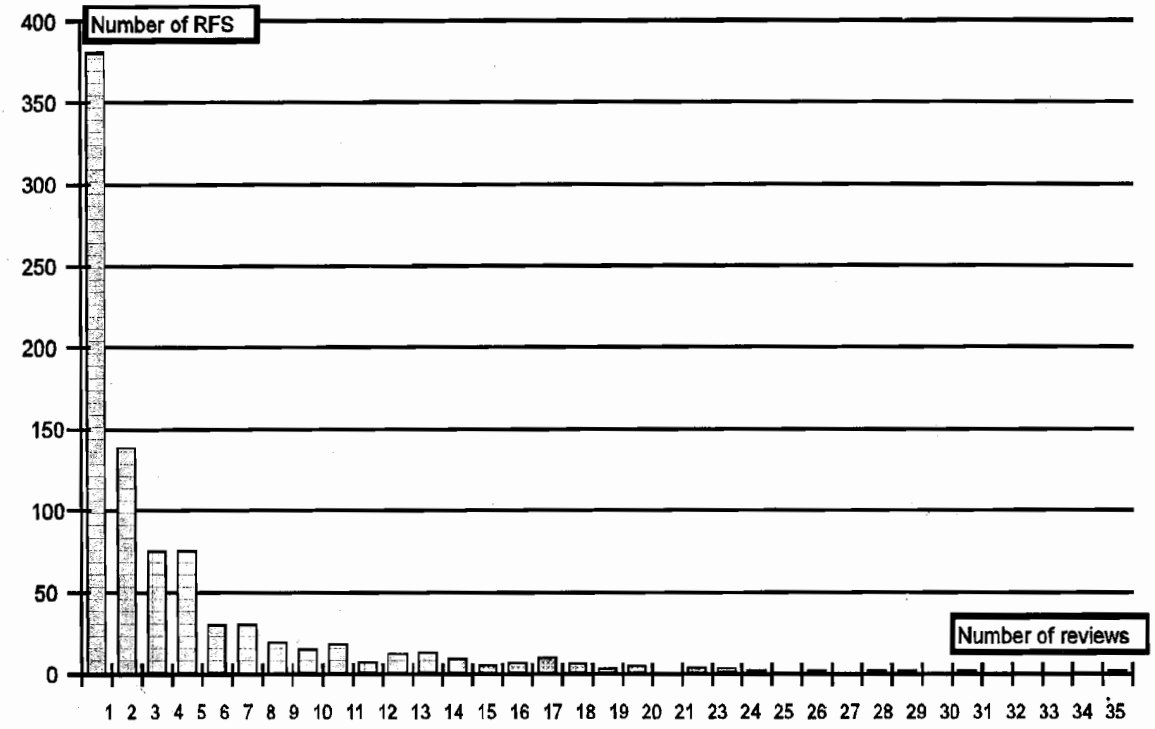
It follows from the distribution of RFS by the number of corresponding reviews presented in Table 6 that one high-weight review corresponds to 41% of medical RFS, 2 — to 18%, 3 — 10%, 4 — 10,6%, 5 — to 3,2%, 6 — to 3, and 7% — to

5%, and 7 — to 2,4% of medical RFS. Core-documents of remaining 10% (38) of RFS are cited by 8-35 reviews. RFS covered by 8 or more high-weight reviews we named supertopical, and the threshold of supertopicality (TS) for medical RFS of ISR-1991 we took equal to 8. The list of 48 supertopical medical RFS is given in Appendix 1.

**Table 6. Distribution of Topical RFS Relevant to Different Disciplines by the Number of Corresponding High-Weight Reviews.**

| No. of high-weight reviews | All RFS | Medicine | Biology | Chemistry | Physics |
|---|---|---|---|---|---|
| 1 | 363 | 154 | 109 | 98 | 85 |
| 2 | 139 | 67 | 50 | 22 | 38 |
| 3 | 72 | 37 | 25 | 18 | 14 |
| 4 | 66 | 40 | 27 | 13 | 17 |
| 5 | 23 | 11 | 8 | 6 | 4 |
| 6 | 24 | 12 | 8 | 5 | 5 |
| 7 | 15 | 9 | 7 | 3 | 3 |
| 8 | 12 | 6 | 5 | 3 | 1 |
| 9 | 6 | 5 | 3 | 1 | 1 |
| 10 | 9 | 6 | 2 | 2 | 2 |
| 11 | 4 | 4 | 3 | | |
| 12 | 5 | 3 | 4 | | |
| 13 | 3 | 2 | 3 | | |
| 14 | 2 | 2 | 1 | | |
| 15 | 3 | 1 | 2 | 1 | 1 |
| 16 | 4 | 3 | 4 | 1 | |
| 17 | 3 | 1 | 3 | | |
| 18 | 1 | 1 | 1 | | |
| 19 | 2 | 2 | 2 | | |
| 20 | | | | | |
| 21 | 2 | | 2 | | |
| 22 | 2 | | 2 | | |
| 23 | 1 | | 1 | | |
| 24 | | | | | |
| 25 | 1 | | 1 | | |
| 26 | | | | | |
| 27 | 1 | 1 | 1 | | |
| 28 | 1 | | 1 | | |
| 29 | | | | | |
| 30 | 1 | 1 | 1 | | |
| 31 | | | | | |
| 32 | | | | | |
| 33 | | | | | |
| 34 | | | | | |
| 35 | 1 | 1 | 1 | | |

A




B

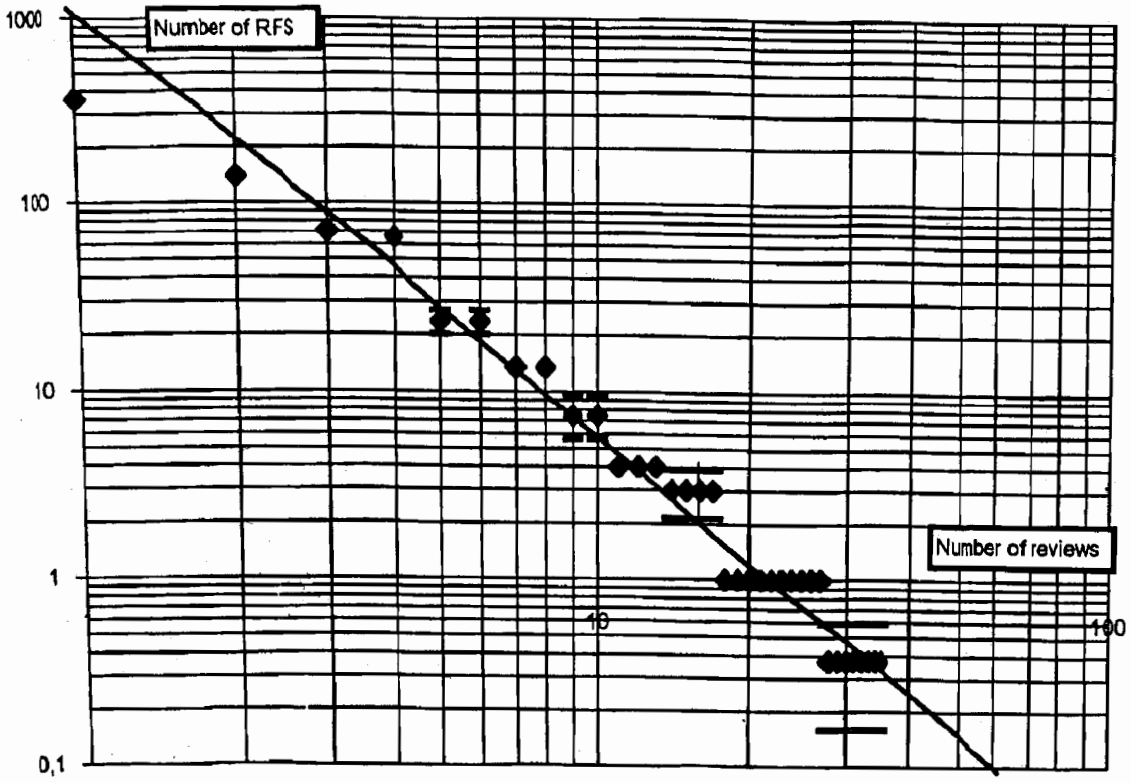
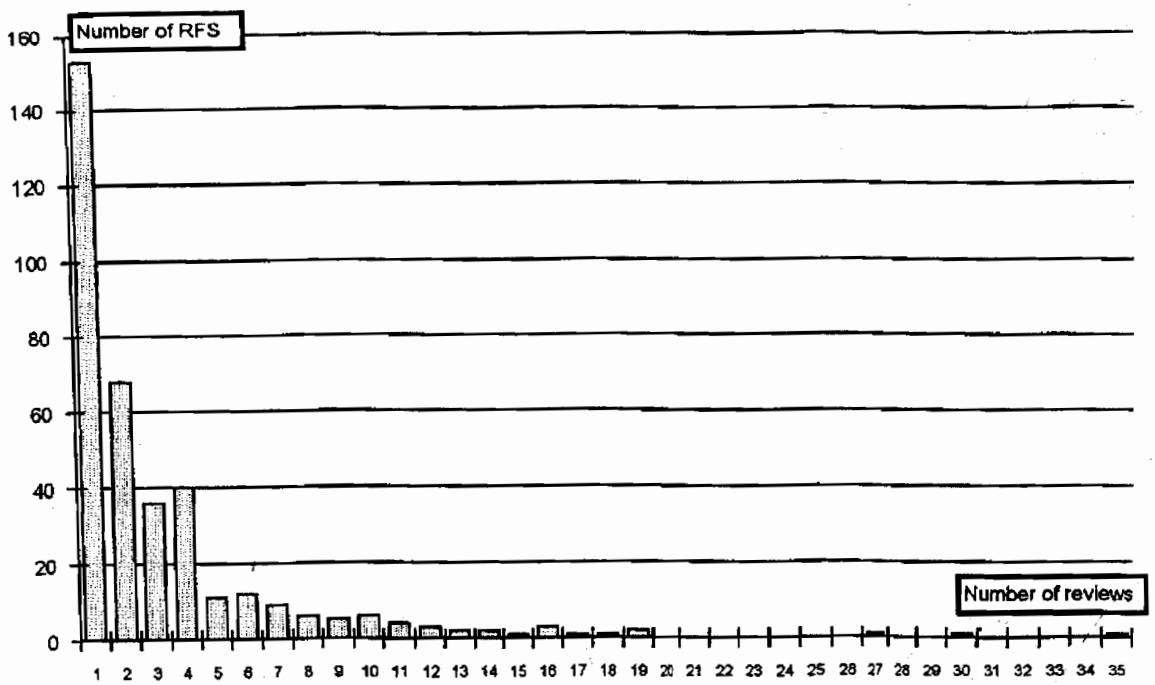



**Fig. 1. Distribution of All Selected RFS by the Number of Corresponding Reviews ($\alpha = 1.3 \pm 0.047$) A**




B

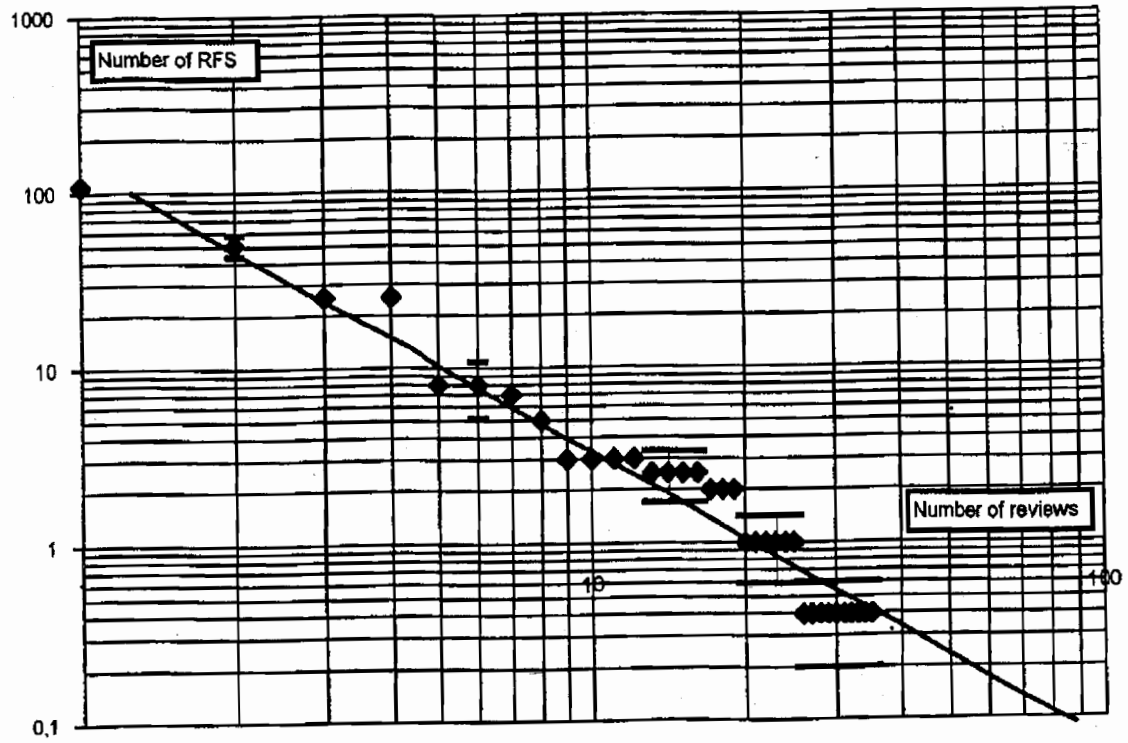
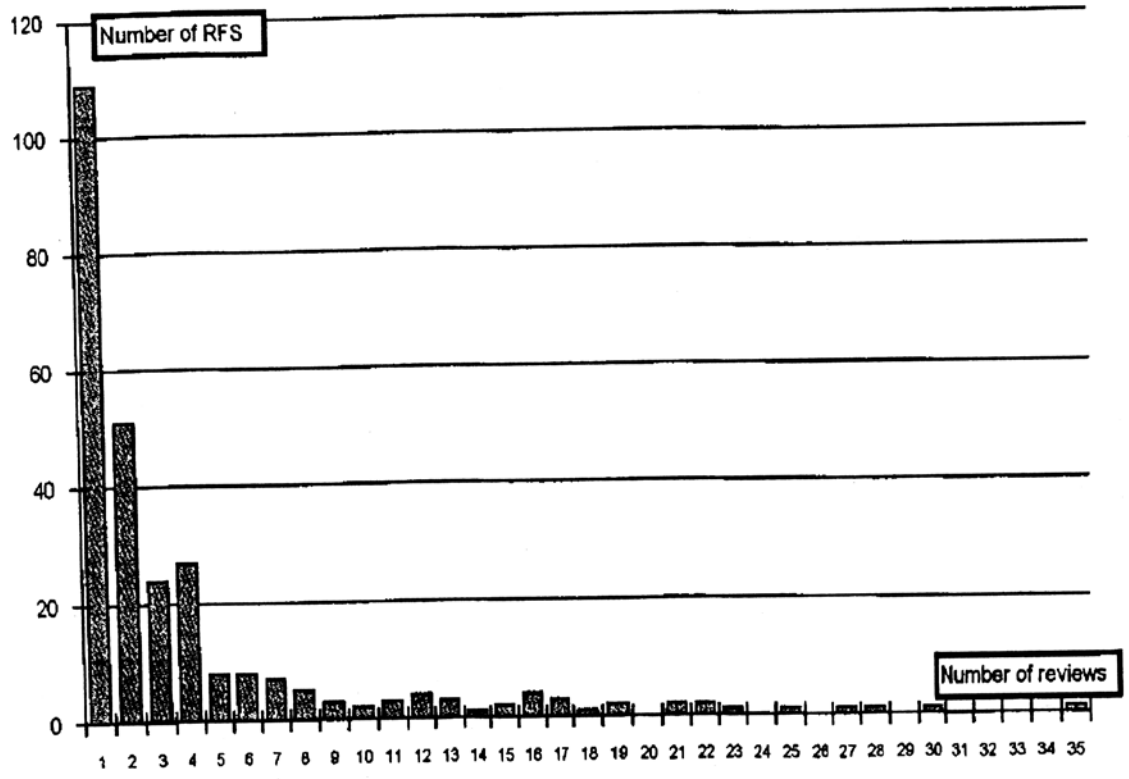



**Fig. 2. Distribution of Medical RFS by the Number of Corresponding Reviews ($\alpha = 1.04 \pm 0.054$)**

A




B

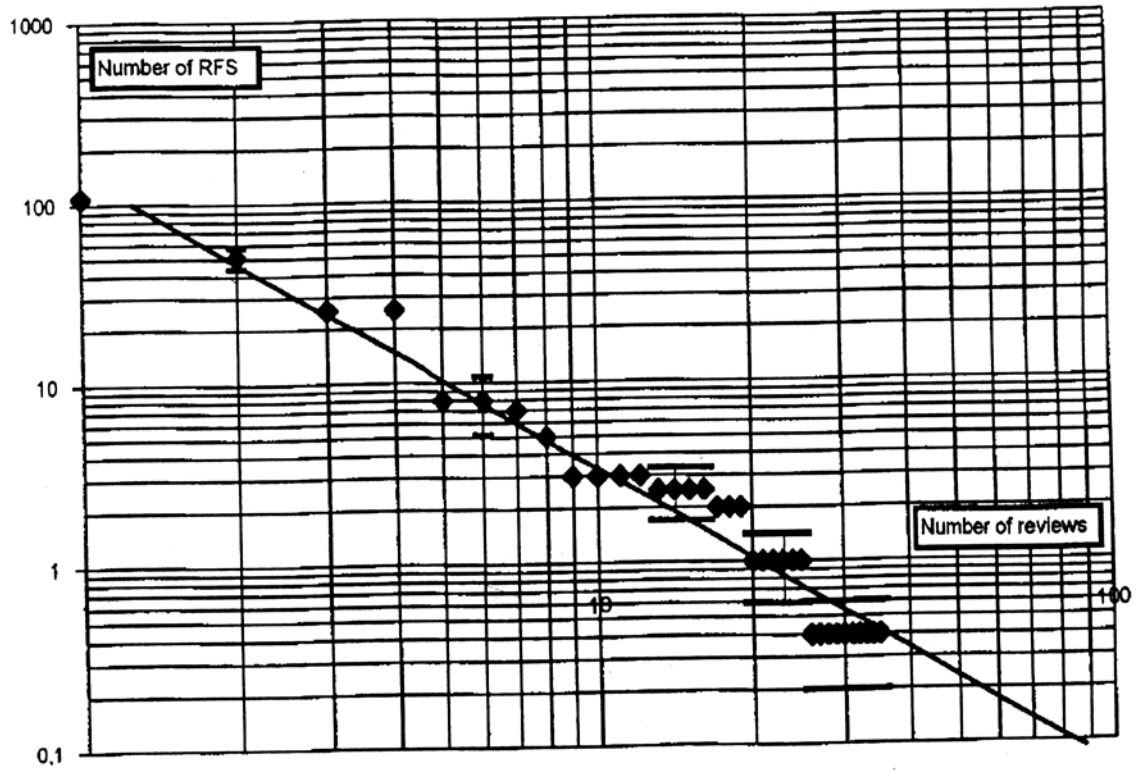
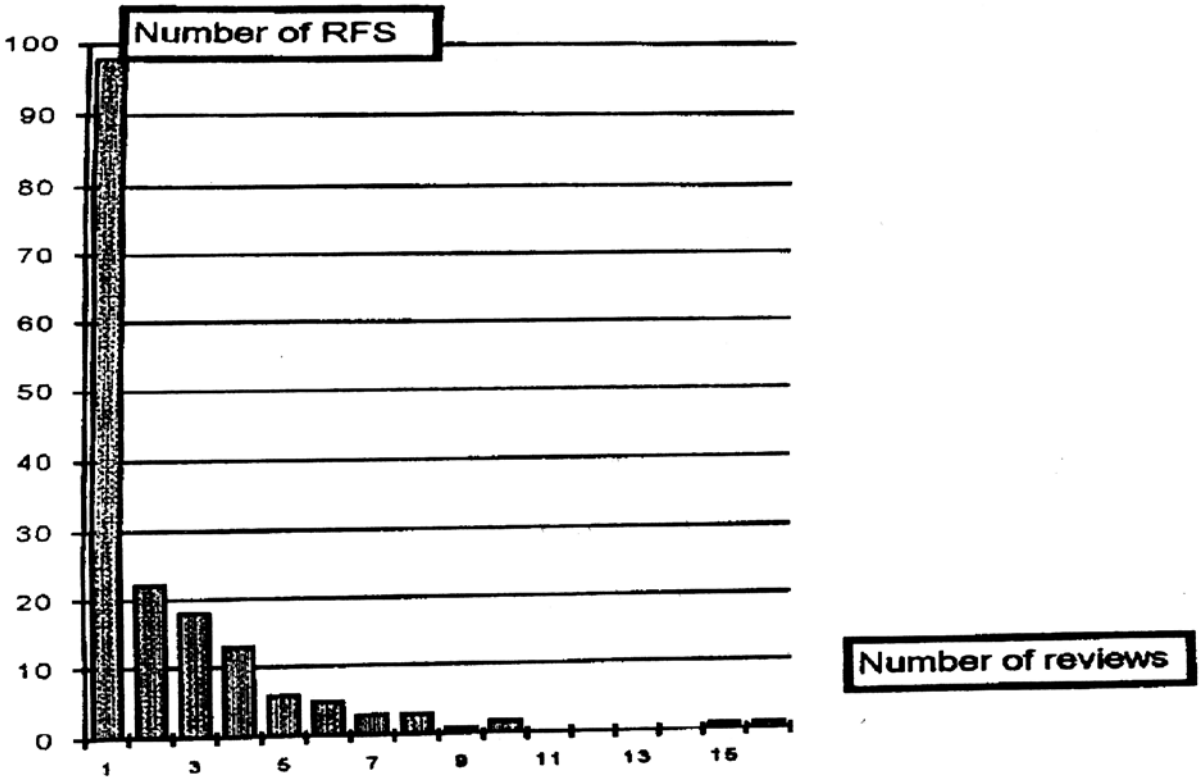



**Fig. 3. Distribution of Biological RFS by the Number of Corresponding Reviews ($\alpha = 0.64 \pm 0.04$)**

A




B

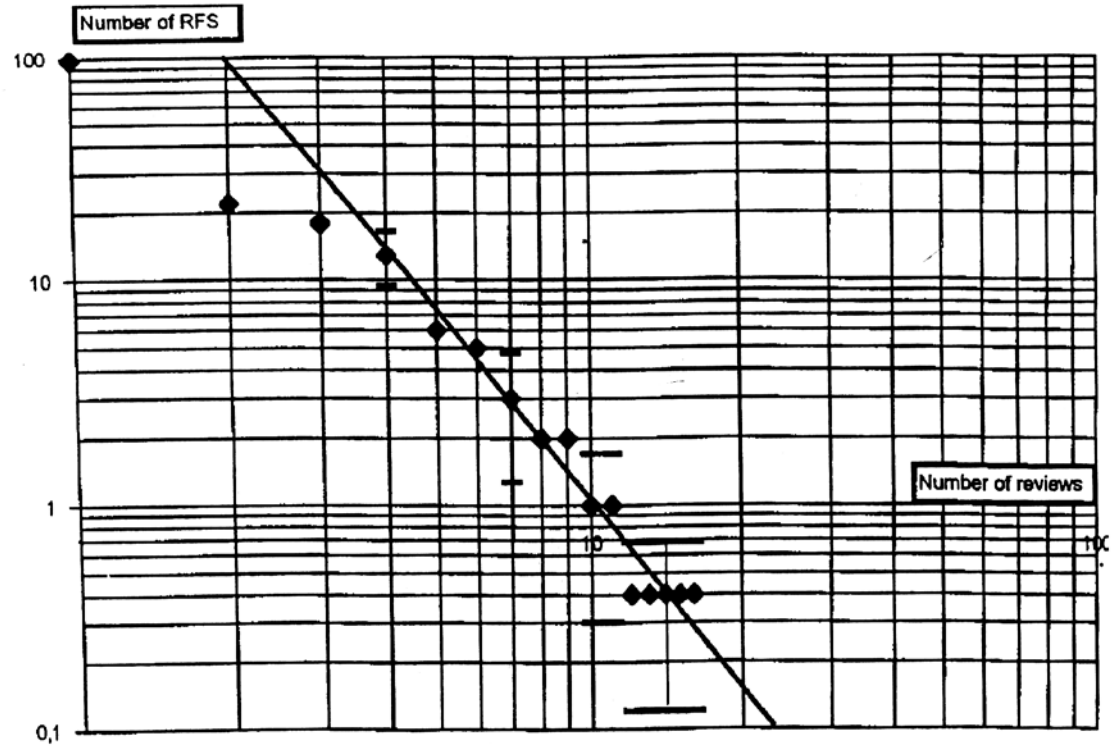
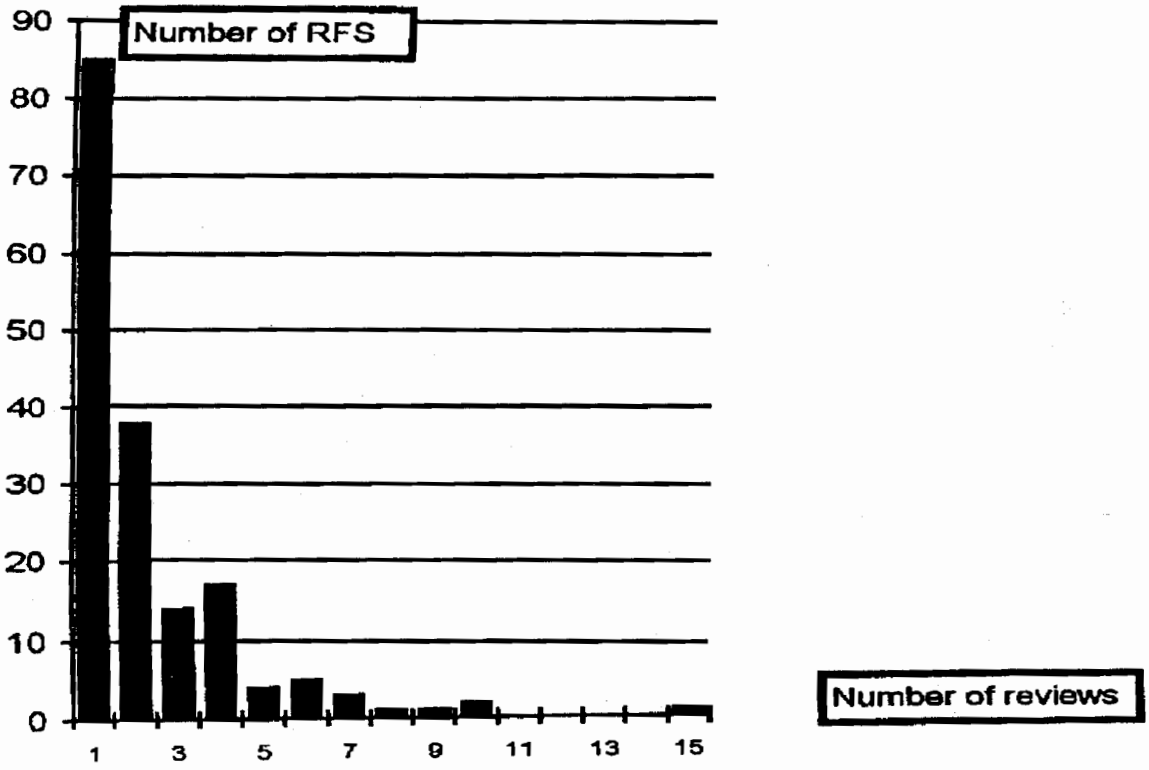



**Fig. 4. Distribution of Chemical RFS by the Number of Corresponding Reviews ($\alpha = 1.84 \pm 0.14$)**

A




B

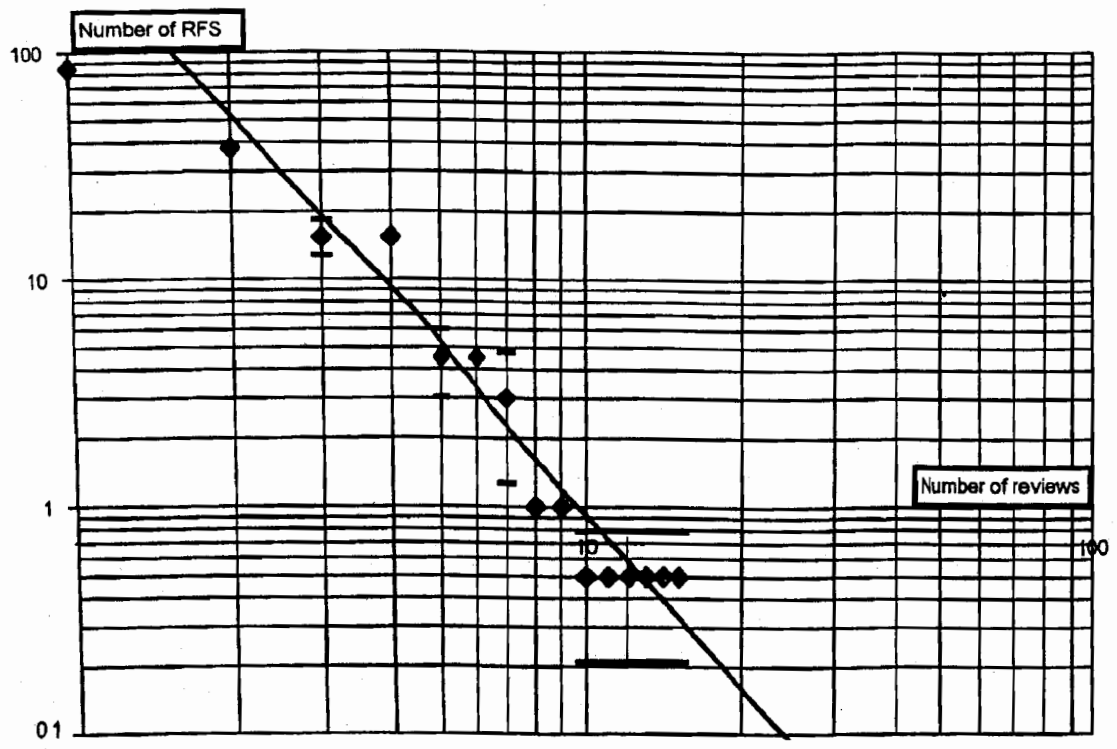



**Fig. 5. Distribution of Physical RFS by the Number of Corresponding Reviews ($\alpha = 2.24 \pm 0.17$)**

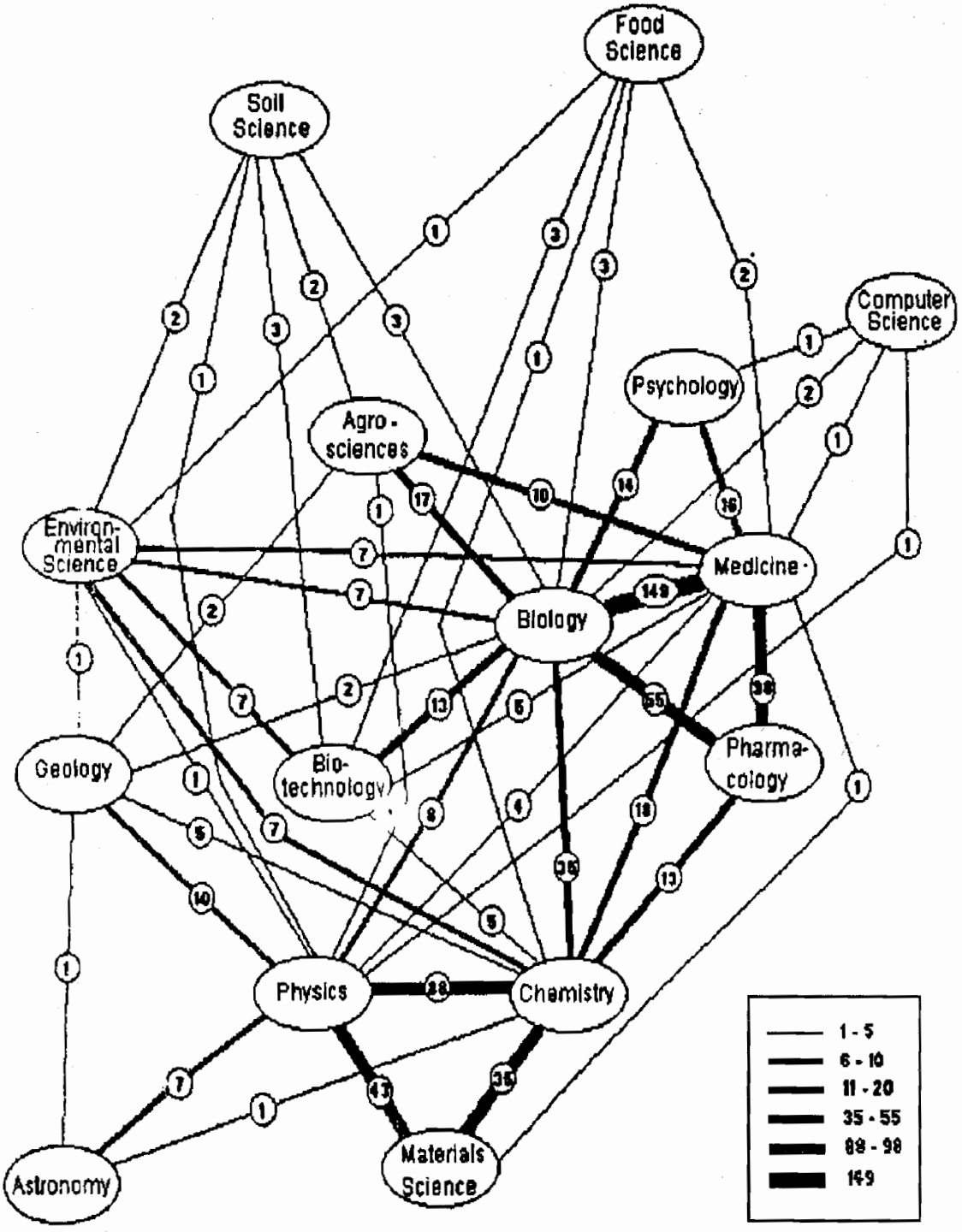



**Fig. 6. Cluster of Interdisciplinarity in Life Sciences**

Similarly, it follows from Table 6 that 40% of selected biological RFS are covered by one review, 18,3% — by 2, 8,5% — by 3, 9,5% — by 3, 2,8% — by both 5 and 6 reviews, and 2,5% — by 7 high-weight reviews. The remaining 15,8% are covered by 8-35 reviews. We considered the TS for biological RFS to be equal to that of medical sciences. The list of 51 supertopical biological RFS is given in Appendix 2.

In case of RFS associated with chemical sciences, 56% of these are covered by one high-weight review, 13%, 10,2% and 7,3% — by 2, 3, and 5 reviews respectively. The remaining 13% of chemical RFS (21 RFS) are covered by 5 — 16 high-weight reviews. In physical sciences 49,5% of selected RFS are covered by one review, 22,7%, 8% and 9,7% — by 2, 3 and 4 high-weight reviews respectively. The remaining 9,9% of RFS in the physical sciences (17 RFS) are covered by 5-15 reviews. Thus we took the TS for chemical and physical sciences for 1991 equal to 5. The lists of 21 supertopical chemical RFS and 17 supertopical physical RFS are given in Appendices 3 and 4.

## 2. Analysis of Topical RFS Related to Medical and Biological Sciences

As can be seen from Tables 2 - 4 medical sciences occupy the leading position in what is called life sciences which is dictated by indisputable and essential importance of these disciplines for survival of humanity and their priority in financing by the state and by private sources. The data on the distribution of selected RFS by relevant medical disciplines and highly important problem areas is given in Table 7. The order of disciplines in the Table is governed by the number of topical RFS related to each discipline. The number and the percentage of supertopical RFS is specified for each discipline and problem area. The highest number of supertopical RFS was associated with oncology (12 RFS), clinical immunology (12 RFS) and infectious diseases (7 RFS).

**Table 7. Distribution of Topical RFS by Medical Disciplines and Important Problem Areas.**

| Discipline or special important problem area | No. of RFS | Of these : supertopical RFS (≥ 7 rev/ RFS |
|---|---|---|
| Oncology | 78 | 12 (15.4%) |
| Surgery | 46 | 4 (15.4%) |
| Neuropathology | 23 | 5 (21.7%) |
| Infectious Diseases | 44 | 7 (15.9%) |
| Clinical Immunology | 43 | 12 (27.9%) |
| Cardiology | 36 | 4 (11.1%) |
| Psychiatry | 30 | 3 (10%) |
| Clinical Chemistry and Biochemistry | 30 | 3 (10%) |
| Hematology | 25 | 7 (28%) |
| Hepatogastroenterology | 24 | 5 (20%) |
| Pediatrics and Perinatology | 21 | 3 (14.3%) |
| Human Genetics and Heredity | 20 | 1 (5%) |
| Clinical Endocrinology | 20 | 4 (20%) |
| Medical Instruments | 19 | - |
| Transplantology | 19 | 3 (15.8%) |
| Nephrology | 18 | 3 (16.7%) |
| Radiology | 18 | 1 (5.6%) |
| Angiology | 17 | 4 (24%) |
| Aids | 16 | 4 (25%) |
| Clinical Virology | 16 | NB 3(18.8%) |
| Hypertension | 15 | NB 3 (20%) |
| Medical Physiology | 14 | - |
| Allergology | 13 | 1 (7.7%) |
| Medical Histology | 13 | 2 (15.4%) |
| Risk Factors, Preventive Medicine | 13 | - |
| Medical Bacteriology | 12 | 2 (16.7%) |
| Respiratory Diseases | 12 | 1 (8.3%) |
| Toxicology | 12 | - |
| Metabolism Disorders | 11 | 1 (9.1%) |
| Nutrition | 11 | 1 (9.1%) |
| Computed Tomography | 10 | 1 (10%) |
| Drug and Substance Abuse | 10 | 1 (10%) |
| Obstetrics | 9 | 1 (11.1%) |
| Rheumatology | 9 | 2 (22.2%) |
| NMR in Medicine | 8 | - |
| Anesthesia and Intensive Care | 8 | 1 (12.5%) |
| Dermatology | 7 | - |
| Orthopedics | 7 | 1 (14.3%) |
| Ultrasound in Medicine | 7 | - |

(Contd....)

| | | |
|---|---|---|
| Hygiene | 6 | 1 (16.6%) |
| Gynecology | 6 | - |
| Ophthalmology | 6 | - |
| Electrocardiography | 6 | - |
| Endoscopy | 5 | 1 (20%) |
| Andrology | 3 | - |
| Otolaringology | 3 | - |
| Lasers in Medicine | 3 | - |
| Trauma | 3 | 1 (33.3%) |
| Urology | 3 | - |
| Fetal Diagnostics | 2 | - |
| Gerontology | 1 | - |
| Emergency and Intensive Care Medicine | 1 | - |
| Contraception | 1 | - |
| Helminthology | 1 | - |
| Stomatology | 1 | - |
| History of Medicine | 1 | - |

The mean percentage of supertopical RFS in medical disciplines was about 10%. Some of these were characterized by the high concentration of supertopical RFS :: 28% of RFS associated with hematology (7 of the 25 RFS); 27.9% of RFS related to clinical immunology (12 of the 43 of RFS);; 25% of RFS related to AIDS (4 of the 16 RFS); 24% of RFS related to angiology (4 of the 17 RFS) ; 21.7% of RFS related to neuropathology (5 of the 23 RFS); 20.1% of RFS associated with hepatogastroenterology (5 of the 24 RFS); 20% of RFS related to clinical endocrinology (4 of the 20); 18.8% of RFS related to clinical virology (3 of the 16 RFS) and 22.2% of RFS in rheumatology (2 of the 9). We can say that these fields could be currently characterized by extensive research activity.

The data on biology RFS was processed the same way as the data on medical RFS. The information about the distribution of the biological RFs is presented in Table 8. The highest number of supertopical RFS was associated with neurobiology and animal biochemistry (12 RFS), molecular biology and immunology (10 RFS), and cytology/histology (9 RFS). The list of supertopical biological RFS is presented in Appendix 2. The mean percentage of supertopical RFS in biological sciences was calculated to be about 10%. The highest concentration of supertopical RFS was found in the studies of neuropeptides (41.7% of RFS), then immunology and molecular biology (both 31.25% of RFS), tumor cell biology (29.8% of RFS), bioorganic chemistry (27.8%), animal biochemistry (27.3% of RFS), animal physiology (by animal physiology we meant physiology of respiration, blood and blood circulation, digestion, reproduction, musculoskeletal system, etc. — all divisions of physiology minus neurophysiology and endocrinology), cytology/histology, and endocrinology and metabolism (22.6%, 20%, and 20%, respectively). These disciplines could be currently characterized by extensive research activity. One of the four RFS in embryology and developmental physiology (20%) and one of the five RFS (25%) in plant biochemistry were also supertopical.

**Table 8. Distribution of Topical RFS by Biological Disciplines and Important Problem Areas.**

| Discipline or important problem area | No. of RFS | Of these : supertopical RFS (≥ 7 rev/ RFS |
|---|---|---|
| Neurobiology | 70 | 12 (17.1%) |
| Cytology/Histology | 45 | 9 (20%) |
| Animal Biochemistry | 44 | 12 (27.3%) |
| Immunology | 32 | 10 (31.25%) |
| Molecular Biology | 32 | 10 (31.25%) |
| Animal Physiology (minus neurobiology & endocrinology) | 31 | 7 (22.6%) |
| Tumor Cell Biology | 24 | 7 (29.2%) |
| Bioorganic Chemistry | 18 | 5 (27.8%) |
| Enzymology | 16 | 2 (12.5%) |
| Animal Genetics | 16 | 1 (6.25%) |
| Microbiology | 16 | 1 (6.25%) |
| Endocrinology & Metabolism | 15 | 3 (20%) |
| Ecology | 12 | - |
| Neuropeptides | 12 | 5 (41.7%) |
| Plant Physiology | 11 | 1 (9.1%) |
| Virology | 11 | 1 (9.1%) |
| Zoology | 8 | - |
| Hydrobiology | 7 | - |
| Evolution | 6 | - |
| Parasitology | 6 | - |
| Entomology | 5 | 1 (20%) |
| Embryology & Developmental Biology | 4 | - |
| Plant Biochemistry | 4 | 1 (25%) |
| Ethology | 4 | - |
| Botany | 3 | - |
| Genosystematics | 3 | - |
| Plant Genetics | 2 | - |
| Biophysics | 1 | - |
| Ornitology | 1 | - |
| Paleobotany | 1 | - |
| Theoretical biology | 1 | - |

Twenty eight supertopical RFS were common for medical and biological sciences, which comes to 58.3% of supertopical RFS in medicine and 54.9% in biology. All but one supertopical RFS covered by more than 10 high-weight reviews were associated with both medicine and biology. Of these the highest number was related to immunology (7 RFS) tumor cell biology (6 RFS), animal biochemistry (5 RFS), and molecular biology (4 RFS).

### 3. Analysis of the Topial RFS in Chemical and Physical Sciences

The information concerning the distribution of topical RFS by chemical and physical disciplines is given in Tables 9 and 10 respectively.

**Table 9. Distribution of Topical RFS by Chemical Disciplines and Important Problem Areas.**

| Discipline or important problem area | No. of RFS | Of these : supertopical RFS (≥ 5 rev/ RFS) |
|---|---|---|
| Physical Chemistry | 50 | 9 (18%) |
| Analytical Chemistry | 50 | 8 (16%) |
| Organic Chemistry | 28 | 1 (3.6%) |
| Coordination Chemistry | 20 | 2 (10%) |
| Surface Chemistry | 20 | 3 (15%) |
| Polymer Chemistry | 19 | 1 (5.3%) |
| Catalysis | 16 | 1 (6.25%) |
| Organometallic Chemistry | 15 | - |
| Stereochemistry | 10 | 1 (10%) |
| Electrochemistry | 9 | 1 (11.1%) |
| Inorganic Chemistry | 8 | - |
| Photochemistry | 8 | - |
| Thermochemistry | 7 | 1 (14.2%) |
| Colloid Chemistry | 7 | - |
| Theoretical Chemistry | 7 | - |
| Chemical Synthesis | 6 | - |
| Radio Chemistry | 1 | - |

The highest number of chemical RFS was related to physical and analytical chemistry (50 RFS each). The highest number of supertopical RFS (9 and 8 RFS respectively) was also associated with these disciplines. The mean percentage of supertopical RFS per chemical discipline was calculated to be about 6.6%. The highest percentage of supertopical RFS was found for physical (18%), analytical (16%), and surface chemistry (15%).

**Table 10. Distribution of Topical RFS by Physical Disciplines and Important Problem Areas.**

| Discipline or important problem area | No. of RFS | Of these : supertopical RFS (≥ 5 rev/ RFS) |
|---|---|---|
| Spectroscopy | 34 | 5 (14.7%) |
| Chemical Physics | 29 | 6 (20.7%) |
| Thin Films | 22 | - |
| Superconductors | 20 | 2 (10%) |
| Optics and Lases | 18 | 1 (5.6%) |
| Quantum Electrodynamics | 15 | 2 (13.3%) |
| Physics of Semiconductors | 15 | - |
| Physics of Materials | 12 | 3 (25%) |
| Gravitation Theory, Cosmology | 11 | 1 (9.1%) |
| Surface Physics | 10 | 2 (20%) |
| Magnetics | 8 | - |
| Nuclear Physics | 7 | - |
| Polymer Physics | 6 | - |
| Fusion/Plasma | 6 | - |
| Atmosphere Physics | 6 | - |
| Liquid Crystals | 5 | - |
| Hydrodynamics | 5 | - |
| Crystallography | 5 | 1 (20%) |
| Disordered Systems | 5 | - |
| Amorphous State Physics | 5 | - |
| Atomic Physics | 4 | - |
| Electron Structure of Solids | 4 | - |
| Acoustics | 3 | - |
| Low Dimension Systems | 3 | 3 (100%) |
| Instruments and Methods in Physics Research | 3 | - |
| Chaos, Fractals | 3 | 1 (30%) |
| Mechanical Properties of Solids | 2 | - |
| Lattice Vibrations | 1 | - |
| Processes of Transfer | 1 | - |
| Superfluidity | 1 | - |
| Physics of Dielectrics | 1 | - |

The highest number of supertopical RFS in physical sciences and research areas was associated with chemical physics and spectroscopy (6 and 5 RFS, or 20.7% and 14.7% respectively). The average percentage of supertopical RFS per physical discipline or important probnlem area was about 9%. In constrast to chemical sciences, several directions of high concentration of supertopical RFS could be defined in physical sciences. These in-

cluded low dimension systems (all 3 RFS, or 100%, were supertopical), chaos and fractals (1 of the 3 RFS were supertopical, 1 - almost supertopical with 4 high-weight reviews relevant to it), physics of materials (3 of the 12 RFS, or 25%, were supertopical, spectroscopy (20.7% of RFS), and surface physics (2 of the 10 RFS, or 20%, were supertopical).

## 4. Topical RFS in Other Disciplines

In materials science (Table 13), the highest number of topical RFS was associated and with the chemistry and physics of materials (18 RFS), metals and alloys (14 RFS), and with the studies of ceramics (13 RFS).

**Table 11. Distribution of Topical RFS in Materials Sciences**

| Discipline or Important problem area | No. of RFS |
|---|---|
| Chemistry & Physics of Materials | 18 |
| Metals and Alloys | 14 |
| Ceramics | 13 |
| Plastic Materials | 7 |
| Composite Materials | 4 |
| Fiber Optics | 2 |
| Optics of Metals | 2 |

**Table 12. Distribution of Topical RFS in Geosciences**

| Discipline or Important problem area | No. of RFS |
|---|---|
| Geophysics | 9 |
| Geochemistry | 9 |
| Physics of Atmosphere | 8 |
| Tectonics | 6 |
| Meteorology | 5 |
| Marine Geology | 4 |
| Petrology | 4 |
| Seismology | 3 |
| Volcanology | 1 |
| Stratigraphy | 1 |

In geosciences (Table 12), the highest number of RFS was related to geophysics and geochemistry (9 RFS each), and atmosphere physics (8 RFS), and in agriculture, dairy and animal science (Table 13) — to veterinary science and plant-breeding (8 RFS in each and to silviculture (5 RFS).

**Table 13. Distribution of Topical RFS in Agriculture, Dairy and Animal Science**

| Discipline or Important problem area | No. of RFS |
|---|---|
| Veterinary Science | 8 |
| Agronomy and Plant Breeding | 8 |
| Silviculture | 5 |
| Plant Diseases and Crop Protection | 3 |
| Livestock-Production | 2 |
| Aquaculture | 2 |

## 5. Participation of Scientists From Different Countries in Adequate Reviewing of Topical RFS

Participation of authors from different countries in adequate reviewing of topical RFS could be one of the most important indicators of performance of national scientific communities within the limits of the cognitive base of the world scientific community and thus — an indirect indicator of the efficiency of national scientific potential. Distribution of high-weight scientific reviews (total and in medical, biological, chemical, physical, materials and geological sciences) by the countries (from the address of the first author) are presented in Tables 14-20 respectively. The character of the distribution of high-weight reviews by the authors from the same country can tell a good deal about the scientific policy in this country.

In the list of the countries listed by the total number of high-weight reviews, works of the Russian scientists occupied the 9th line after the American, English, French, Canadian, Japanese, Dutch, and Australian authors (Table 14). However, in similar distributions of reviews in medical sciences (Table 15), Russian authors occupied the 13th line, and biological sciences - the 12th line (Table 16), whereas in chemical sciences Russian Scientists occupied the 4th line (Table 17, and in physical sciences - the 6th line (Table 18).

Thus we can see that the works of Russian authors show a strong bias towards the chemical and physical sciences. Thus, 31.1% of the high-weight reviews by Russian authors are related to medical, 28.9% — to biological, 46.7% — to chemical, and 26.7% — to physical sciences. However, the distribution of reviews by the authors from other developed countries is essentially different: thus, 58.8% of reviews by the American authors

are associated with medical sciences, 52.2% — with biological, 14.7% — with chemical, and 12.9% — with physical sciences; similar is the distribution of reviews by :

(a) English authors : 58.5%, 58.5%, 14.5% , 11.6%

(b) Australian authors : 52.2%, 54.35%, 10.9%, 8.7%

(c) Dutch authors : 62.5%, 54.2%, 16.7%, 16.7%

(d) Canadian authors : 57.3%, 34.15%, 15.9%, 11%

(e) Swedish authors : 59.1%, 61.4%, 22.7%, 13.6%

(f) Italian Authors : 53.5%, 51.2%, 16.3%, 20.1%

(g) French authors : 51.7%, 56.2%, 19.1%, and 19.1% respectively.

This character of the distribution of scientific priorities indicates that the scientific policy in these countries is humanly oriented with health care being of the primary significance. The typical examples of this are the distributions of reviews by the authors from Denmark (65.6%, 37.5%, 9.4%, and 3.1% respectively), Finland and Israel.

A somewhat different kind of distribution, more oriented towards the chemical and physical sciences can be observed for the reviews by :

*(a)* Swiss authors : 48.7%, 56.4%, 25.6%, 25.6%

*(b)* German authors : 42.5%, 52.8%, 22.8%, 28.35%

*(c)* Japanese : 32.14%, 57.14%, 28.6%, 23.2% respectively.

A strong bias towards chemical and physical sciences was observed for India (21%, 21%, 42%, and 52.6%, respectively) and Poland (9.1%, 9.1%, 54%, and 63.6%, respectively), which was even stronger than in case of the USSR.

**Table 14. Distribution of High-Weight Reviews by the Countries (from the Address of the First Author)**

| Country | Number of reviews | | | | | | |
|---|---|---|---|---|---|---|---|
| | Total | Med | Biol | Chem | Phys | Mater | Geo |
| USA | 1027 | 604 | 536 | 151 | 132 | 44 | 42 |
| Britain | 207 | 121 | 121 | 30 | 24 | 8 | 5 |
| Germany | 127 | 54 | 67 | 29 | 36 | 10 | 4 |
| France | 89 | 46 | 50 | 17 | 17 | 9 | 3 |
| Canada | 82 | 47 | 28 | 13 | 9 | 6 | 5 |
| Japan | 56 | 18 | 32 | 16 | 13 | 8 | - |
| Holland | 48 | 30 | 26 | 8 | 8 | 3 | - |
| Australia | 46 | 24 | 25 | 5 | 4 | - | 2 |
| USSR | 45 | 14 | 13 | 21 | 12 | 6 | 1 |
| Sweden | 44 | 26 | 27 | 10 | 6 | 2 | - |
| Italy | 43 | 23 | 22 | 7 | 9 | - | 1 |
| Switzerland | 39 | 19 | 22 | 10 | 10 | 5 | - |
| Denmark | 32 | 21 | 12 | 3 | 1 | - | - |
| India | 19 | 4 | 4 | 8 | 10 | 7 | 1 |
| New-Zealand | 18 | 13 | 9 | - | 2 | - | - |
| Belgium | 17 | 9 | 10 | 4 | 3 | 1 | 1 |
| Finland | 16 | 13 | 11 | 1 | - | - | - |
| Israel | 15 | 10 | 8 | 2 | 2 | 2 | - |
| Poland | 11 | 1 | 1 | 6 | 7 | 1 | - |
| Czechoslovakia | 10 | - | 6 | 8 | - | 1 | - |
| Austria | 9 | 5 | 3 | 2 | 1 | - | - |
| Spain | 7 | 3 | 3 | 1 | 3 | - | - |
| China | 6 | 1 | 1 | 6 | 1 | - | 1 |
| South African Republic | 5 | 2 | 3 | 1 | - | - | 1 |

**Table 15. Distribution of High-Weight Reviews in Medical Sciences by the Countries (from the First Author's Address)**

| Country | No. of reviews |
|---|---|
| USA | 604 |
| Britain | 121 |
| Germany | 54 |
| Canada | 47 |
| France | 46 |
| Holland | 30 |
| Sweden | 26 |
| Australia | 24 |
| Italy | 23 |
| Denmark | 21 |
| Switzerland | 19 |
| Japan | 18 |
| USSR | 14 |
| Finland | 13 |
| New Zealand | 13 |
| Israel | 10 |
| Belgium | 9 |
| Austria | 5 |

**Table 16. Distribution of High-Weight Reviews in Biological Sciences by the Countries (from the First Author's Address)**

| Country | No. of reviews |
|---|---|
| USA | 536 |
| Britain | 121 |
| Germany | 67 |
| France | 50 |
| Japan | 32 |
| Canada | 28 |
| Sweden | 27 |
| Holland | 26 |
| Australia | 25 |
| Italy | 22 |
| Switzerland | 22 |
| USSR | 13 |
| Denmark | 12 |
| Finlnd | 11 |
| Belgium | 10 |
| New Zealand | 9 |
| Israel | 8 |
| Czechoslovakia | 6 |

**Table 17. Distribution of High-Weight Reviews in the Chemical Sciences by the Countries (from the First Author's Address)**

| Country | No. of reviews |
|---|---|
| USA | 151 |
| Britain | 30 |
| Germany | 29 |
| USSR | 21 |
| France | 17 |
| Japan | 16 |
| Canada | 13 |
| Sweden | 10 |
| Switzerland | 10 |
| India | 8 |
| Czechoslovakia | 8 |
| Holland | 8 |
| Italy | 7 |
| China | 6 |
| Poland | 6 |
| Australia | 5 |
| Belgium | 4 |

**Table 18. Distribution of High-Weight Reviews in Physical Sciences by the Countries (from the First Author's Address)**

| Country | No. of reviews |
|---|---|
| USA | 132 |
| Germany | 36 |
| Britain | 24 |
| France | 17 |
| Japan | 13 |
| USSR | 12 |
| India | 10 |
| Switzerland | 10 |
| Canada | 9 |
| Italy | 9 |
| Holland | 8 |
| Poland | 7 |
| Sweden | 6 |
| Australia | 4 |

**Table 19. Distribution of High-Weight Reviews in Materials Sciences by the Countries (from the First Author's Address)**

| Country | No. of reviews |
|---|---|
| USA | 44 |
| France | 10 |
| Germany | 10 |
| Britain | 8 |
| Japan | 8 |
| India | 7 |
| Canada | 6 |
| USSR | 6 |
| Switzerland | 5 |
| Holland | 3 |
| Israel | 2 |
| USSR | 1 |

**Table 20. Distribution of High-Weight Reviews in Geosciences by the Countries (from the First Author's Address)**

| Country | No. of reviews |
|---|---|
| USA | 42 |
| Canada | 5 |
| England | 5 |
| Germany | 4 |
| France | 3 |
| Australia | 2 |
| Belgium | 1 |
| India | 1 |
| Italy | 1 |
| China | 1 |
| South African Republic | 1 |

## 6. Interdisciplinarity of Topical RFS

We have mentioned already that the share of interdisciplinary RFS is rather large. Actually these comprise 44% of the total number of RFS. The highest number of interdisciplinary RFS is related to both biology and medicine. These 149 RFS comprise 39.5% of all medical and 52.5% of all biological RFS. Eighty-eight RFS are associated with both chemical and physical sciences thus comprising 50% and 51% respectively. Materials science is closely related with physics and chemistry. Of 51 RFS related to materials science, 43 are also related to physics and 35 — to chemistry. Pharmacology has the highest number of intersections with other disciplines : among the 130 pharmacological RFS 55 are also related to biology, 98 — to medicine and 13 — to chemistry. Another highly interdisciplinary discipline is environmental science. Its 17 RFS are related to biology, biotechnology, medicine and chemistry (7 RFS related to each discipline).

Clustering of the interdisciplinary RFS related to two or more disciplines allowed to obtain a single cluster (or a map) of interdisciplinary of topical science. We must emphasize that we are talking here only about topical science since the map was constructed on the base of selected topical RFS. The map thus obtained is presented in Fig. 6. The nodes of the cluster are comprised by major disciplines and the connections between them are defined by the number of RFS (given in circles) related to every pair of sciences.

## 7. Conclusive Remarks

The hot topics in life sciences, that is the areas of breakthroughs, can be determined using the publication of the relevant topical reviews as an indicator. Scientific review as a direct or indirect response to a social order by scientific community is regarded as a proof of validity and urgency of the new knowledge. We consider topical the reviews that cite 5 or more core documents of any RFS. If RFS is covered by the high number of topical reviews, the new knowledge obtained within this RFS is considered quite important and promising by the internatioal scientific community. "Index for Scientific Reviews" published by ISI may be used and subdisciplines reflects the structure of modern topical science. Analysis of geographical distribution of reviews provides information about both scientific policy and potential in different countries.

Interdisciplinary character of many topical RFS, i.e., association of RFS with two or more major disciplines, allows to cluster these disciplines on the principle of their interdisciplinary relations.

## Appendices

*Appendix 1.*

***Hypertopical RFS in medicine***

RFS 0312 35 rev. MED - oncol BIOL - tumor cell biol; immunol PHARMACOL TUMOR NECROSIS FACTOR; ANTIVIRAL ACTIVITY; EFFECTS OF CYTOKINES

RFS 0573 30 rev. MED - oncol BIOL - tumor cell biol; molec. biol
SUPPRESSOR GENES; HUMAN-BREAST CANCER; MEDULLARY-THYROID CARCINOMA; MOLECULAR MECHANISMS; BIOLOGY OF DISEASE

RFS 4383 27 rev. MED - hypertension; angiol BIOL - animal physiol PHARMACOL ENDOTHELIUM-DERIVED NITRIC OXIDE; L-ARGININE IN THE CENTRAL NERVOUS SYSTEM; VASCULAR TONE

RFS 0627 19 rev. MED - clin. immunol BIOL - immunol
ANTIGEN RECOGNITION SITE OF MAJOR HISTOCOMPATIBILITY COMPLEX CLASS - II MOLECULES; T-CELL RECEPTOR; SYNTHETIC PEPTIDES

RFS 0577 19 rev. MED - oncol BIOL - molec. biol; tumor cell biol
PLATELET-DERIVED GROWTH-FACTOR; PDGF RECEPTOR; ONCOGENE EXPRESSION; C-ERBB-2 AMPLICATION IN HUMAN-BREAST CARCINOMA; NON-SMALL CELL LUNG-CANCER

RFS 1051 18 rev. MED - oncol; hematol; clin. immunol; transplantol, surgery BIOL - tumor cell biol
RECOMBINANT HUMAN GRANULOCYTE-MACROPHAGE COLONY-STIMULATING FACTOR; HEMATOPOIETIC RECONSTRUCTION FOLLOWING AUTOLOGOUS BONE-MARROW TRANSPLANTATION

RFS 0580 17 rev. MED - aids BIOL - molec. biol; animal biochem
HUMAN IMMUNODEFICIENCY VIRUS; HIV INFECTION; RECOMBINANT SOLUBLE CD4 RECEPTOR; PROTEIN EXPRESSION VIA A CIS-ACTING SEQUENCE

RFS 5495 16 rev. MED - clin. immunol BIOL - immunol
T-CELL RECEPTOR V-BETA EXPRESSION; POSITIVE SELECTION; TRANSGENIC MICE

RFS 2468 16 rev. MED - oncol, infect. dis. BIOL - immunol AGRO - vet
TUMOR NECROSIS FACTOR; MURINE INTERLEUKIN-1 ALPHA PROTECTS MICE; SEVERE SEPSIS

RFS 0090 16 rev. MED - neurol; psychiat BIOL - cyt/hist; animal biochem; neurobiol; molec. biol
ALZHEIMERS-DISEASE AMYLOID PRECURSOR PROTEIN; DIFFERENTIAL BRAIN EXPRESSION; REACTIVE ASTROCYTES FOLLOWING NEURONAL DAMAGE

RFS 2002 15 rev. MED - clin. immunol; oncol BIOL - immunol
RECOMBINANT INTERLEUKIN-2; LYMPHOKINE-ACTIVATED KILLER CELLS; ADOPTIVE IMMUNOTHERAPY MEDICINE

RFS 0237 14 rev. MED - nephrol; hematol BIOL - animal biochem PHARMACOL RECOMBINANT HUMAN ERYTHROPOIETIN; ANEMIA OF CHRONIC RENAL-FAILURE; IN VIVO METABOLISM

RFS 0104 14 rev. MED - clin. immunol; allergol BIOL - immunol
PLATELET-ACTIVATING FACTOR; INCREASING ASTHMA MORTALITY; AIRWAY HYPERRESPONSIVENESS IN ALLERGIC SHEEP; LATE RESPONSE

RFS 1708 13 rev. MED - clin. immunol BIOL - tumor cell biol; immunol; cyt/hist
T-CELL RECEPTOR DELTA-GENES IN HUMAN T-CELL LEUKEMIAS; ANALYSIS OF JUNCTIONAL DIVERSITY; IMPLICATIONS FOR THYMIC DIFFERENTIATION

RFS 1398 13 rev. MED - hypertension; angiol; nephrol BIOL - animal physiol; neuropeptide PHARMACOL ENDOTHELIN STIMULATES RELEASE; RAT CARDIAC MEMBRANES; VASCULAR SMOOTH-MUSCLE; SYSTEMIC VASOCONSTRICTOR PEPTIDE

RFS 0825 12 rev. MED - nephrol; clin. endocrinol; clin. biochem
DIABETIC NEPHROPATHY; ANGIOTENSIN CONVERTING ENZYME-INHIBITORS; URINARY ALBUMIN

RFS 0778 12 rev. MED - oncol BIOL - animal biochem; cyt/hist; tumor cell biol PHARMACOL MULTIDRUG RESISTANCE; P-GLYCOPROTEIN EXPRESSION; CHINESE-HAMSTER OVARY CELLS

RFS 0170 12 rev. MED - infect. dis.; pediatrics BIOL - entomol AGRO - vet
DIAGNOSIS OF LYME-DISEASE; BORRELIA-BURGDORFERI IN CHILDREN; DIFFERENT CLINI CAL MANIFESTATIONS

RFS 1721 11 rev. MED - hematol; nutrition BIOL - animal biochem
FISH OIL; OMEGA-3 FATTY ACIDS; DIETARY SUPPLEMENTATION

RFS 1095 11 rev. MED - Neurol; hereditary dis.; clin. biochem. BIOL - animal genetics
DUCHENNE MUSCULAR-DYSTROPHY; MOLECULAR DELETION PATTERNS; DMD GENE

RFS 0744 11 rev. MED - neurol BIOL - neuropeptides; neurobiol
CALCITONIN GENE-RELATED PEPTIDE; LATERAL SPINAL NUCLEUS OF THE RAT CERVICAL SPINAL-CORD

RFS 0119 11 rev. MED - aids; clin. virol
HIV TYPE-1; ZIDOVUDINE PHARMACOKINETICS; ANTIRETROVIRAL THERAPY; ANTI-HIV NUCLEOSIDE ANALOGS; MANAGEMENT OF AIDS PATIENTS.

RFS 3007 10 rev. MED - traumatol; clin. endocrinol; tomography; radiol; orthopedics
BONE MASS; HIP FRACTURE IN WOMEN; POSTMENOPAUSAL OSTEOPOROSIS; SPINAL QUAN TITATIVE COMPUTED-TOMOGRAPHY; PROXIMAL FEMUR

RFS 2155 10 rev. MEDICINE - rheumatol; clin. immunol
WEGENERS GRANULOMATOSIS; SYSTEMIC VASCULITIS; NEUTROPHIL CYTOPLASMIC ANTI BODIES

RFS 1247 10 rev. MED - clin. bacteriol; gastroenterol; clin. histol
CAMPILOBACTER-PYLORI INFECTION; HISTOLOGIC GASTRITIS; ULCER DISEASE

RFS 1115 10 rev. MED - oncol; clin. biochem BIOL - enzymol; cyt/hIst; animal biochem
DNA TOPOISOMERASE-II; DRUG-RESISTANT CHINESE-HAMSTER OVARY CELL-LINE; BIOCHEMICAL BASIS

RFS 0756 10 rev. MED - cardiol; angiol; surgery; hematol
ACUTE MYOCARDIAL-INFARCTION; THROMBOLYTIC THERAPY; CORONARY ANGIOPLASTY FOR UNSTABLE ANGINA

RFS 0623 10 rev. MED - endoscopy; gastroenterol
ENDOSCOPIC SPHINCTEROTOMY FOR BILE-DUCT STONES; BILIARY LITHOTRIPSY; GALL STONE DISEASE

RFS 2862 9 rev. BIOL - MICROBIOL; ANIMAL BIOCHEM MED - infect. dis.; clin. bacteriol; oncol; clin. immunol; gastroenterol CHEM - coord.chem
SHIGA-LIKE TOXIN-II; ESCHERICHIA-COLI SEROTYPE O-157-H7; RADIOLABELLED MONOCLONAL- ANTIBODIES; HEMOLYTIC UREMIC SYNDROME; RIBOSOME-INACTIVATING PROTEINS

RFS 2487 9 rev. MED - oncol; clin. virol; clin. immunol; hematol BIOL - virol
ADULT-CELL LEUKEMIA-LYMPHOMA AMONG HUMAN T-LYMPHOTROPIC VIRUS TYPE-1 CARRIERS; TROPICAL SPASTIC PARAPARESIS; HTLV-1 INFECTION

RFS 1513 9 rev. BIOL - neurobiol; cyt/hist MED - neurol; transplantol; surgery
FETAL DOPAMINE NEURAL GRAFTS; RAT MODEL OF PARKINSONS-DISEASE; INTRACEREBRAL TRANSPLANTATION; POLYMERIC BRAIN IMPLANT; BOVINE ADRENAL CHROMAFFIN CELLS

RFS 0600 9 rev. MED - aids; infect. dis.
PNEUMOCYSTIS-CARINII PNEUMONIA; AIDS; AIDS PATIENTS

RFS 0151 9 rev. MED - infect. dis.; clin. virol
CHRONIC FATIGUE SYNDROME; PLASMA-EXCHANGE IN DERMATOMYOSITIS POLYMYOSITIS; ACTIVE EPSTEIN-BARR VIRUS-INFECTION

RFS 2363 8 rev. MED - hematol; ; angiol BIOL - animal biochem; cyt/hist
PLASMINOGEN-ACTIVATOR INHIBITOR; ROLE OF VASCULAR ENDOTHELIAL-CELLS; ABNORMAL FIBRINOLYSIS IN HEALTHY MALE CIGARETTE SMOKERS

RFS 1984 8 rev. MED - gastroenterol
SMALL BOWEL MOTILITY; CLINICAL EPIDEMIOLOGY OF CHRONIC CONSTIPATION; IDIOPATHIC COLONIC DYSMOTILITY

RFS 1080 8 rev. MED - oncol; transplantol; surgery
BONE-MARROW TRANSPLANTATION; ACUTE MYELOGENOUS LEUKEMIA; CHRONIC GRAFT-VERSUS-HOST DISEASE

RFS 0890 8 rev. MED - oncol; clin. immunol; clin. histol
T-CELL RECEPTOR GENE REARRANGEMENT; ACUTE LYMPHOBLASTIC-LEUKEMIA; LYMPHOPROLIFERATIVE DISORDERS; LARGE GRANULAR LYMPHOCYTES; LYMPHOID MALIGNANCIES

RFS 0389 8 rev. MED - infect. dis.; clin. immunol PHARMACOL
INVASIVE PULMONARY ASPERGILLOSIS; ;EMPIRIC AMPHOTERICIN-B THERAPY; CHRONIC SYSTEMIC CANDIDIASIS; IMMUNOCOMPROMISED HOST; INFECTION IN NEUTROPENIC PATIENTS

RFS 0065 8 rev. MED - drug and substance abuse; obstetrics; pediatrics
IN-UTERO COCAINE EXPOSURE; PERINATAL OUTCOME; DRUG-USE IN PREGNANCY

RFS 1296 7 rev. MED - clin. immunol; rheumatol; hematol
ANTIPHOSPHOLIPID ANTIBODIES; LUPUS ANTICOAGULANT IN SYSTEMIC LUPUS-ERYTHEMATOSUS; THROMBOTIC DISEASE

RFS 1053 7 rev. BIOL - neurobiol; neuropeptides MED - psychiat; clin. endocrinol
CORTICOTROPIN-RELEASING FACTOR; MAJOR DEPRESSION; DUODENAL BICARBONATE SECRETION DURING STRESS

RFS 0887 7 rev. BIOL - animal physiol; neuropeptides MED - cardiol; hypertension
ATRIAL NATRIURETIC PEPTIDE; CHRONIC LEFT-SIDED HEART-FAILURE; NACL-RESISTANT SPONTANEOUSLY HYPERTENSIVE RATS

RFS 0822 7 rev. MED - aids; neurol; pediatrics
HUMAN IMMUNODEFICIENCY VIRUS; HIV INFECTION; CENTRAL NERVOUS-SYSTEM PATHOLOGY IN PEDIATRIC AIDS; ACQUIRED IMMUNE DEFFICIENCY SYNDROME

RFS 0758 7 rev. MED - cardiol
MYOCARDIAL REPERFUSION INJURY; INFARCT SIZE; CONTRACTILE FUNCTION

RFS 0444 7 rev. BIOL - neurobiol; animal physiol MED - cardiol PHARMACOL
CONGESTIVE HEART-FAILURE; REGULATION OF CARDIAC BETA-ADRENERGIC RECEPTORS; CARDIOVASCULAR VARIABILITY SIGNALS

RFS 0416 7 rev. MED - gastroenterol PHARMACOL
NON-STERIODAL ANTI-INFLAMMATORY DRUGS; PEPTIC-ULCER DISEASE; MISOPROSTOL HEALS GASTRODUODENAL INJURY

RFS 0264 7 rev. BIOL - neurobiol; neuropeptides MED - psychiat; clin. endocrinol
NEURO-ENDOCRINE SYTEMS; RECOMBINANT INTERLEUKIN-1; PSYCHOSOCIAL STRESS; BIDIRECTIONAL INTERACTION; CORTICOTROPIN-RELEASING HORMONE

RFS 0002 7 rev. MED - infect. dis. PHARMACOL
CHRONIC HEPATITIS-B; INTERFERON THERAPY; RANDOMIZED CONTROLLED TRIAL OF ADENINE-ARABINOSIDE 5'-MONOPHOSPHATE

*Appendix 2*

***Hypertopical RFS in biology***

*- RFS related also to medical sciences are marked by*

RFS 0312 35 rev.*

RFS 0573 30 rev.*

RFS 1985 28 rev.* BIOL - enzymol; animal biochem
PROTEIN KINASE-C; INOSITOL PHOSPHATES REGULATES CALCIUM SIGNALING; BIOCHEMICAL MECHANISMS OF PLATELET ACTIVATION

RFS 4383 27 rev.*

RFS 1013 25 rev. BIOL - molec. biol
SEQUENCE-SPECIFIC DNA INTERACTION IN THE FOS JUN PROTEIN COMPLEX; EARLY GENE INDUCTION; TRANSCRIPTION FACTOR AP-1; LEUCINE ZIPPER DOMAIN

RFS 1028 23 rev. BIOL - neurobiol PHARMACOL
EXCITATORY AMINO-ACID NEUROTOXICITY AT THE N-METHYL-D-ASPARTATE RECEPTOR; [H-3] MK-801 IN RAT CEREBRAL-CORTEX; GLYCINE MODULATORY SITE

RFS 2063 22 rev. BIOL - neurobiol; animal physiol
CALCIUM CHANNELS; SMOOTH-MUSCLE CELLS; ISOLATED RAT HYPOTHALAMIC NEURONS

RFS 1298 22 rev. BIOL - neurobiol; cyt/hist PHARMACOL
MUSCARINIC RECEPTOR SUBTYPES; MURINE FIBROBLAST B82 CELLS; FUNCTIONAL INTERRELATIONSHIPS

RFS 3405 21 rev. BIOL - cyt/hist; immunol; tumor cell biol
FIBRONECTIN RECEPTOR; ALKALINE-PHOSPHATASE ACTIVITY IN HUMAN OSTEOSARCOMA CELLS; BETA-1 INTEGRINS; EXPRESSION OF VLA-5

RFS 1275 21 rev. BIOL - bioorganic chem; molec. biol
CELLULAR HEAT-SHOCK PROTEIN; MITOCHONDRIAL MEMBER OF THE HSP70 FAMILY; MAMMALIAN BIP/GRP78 GENE

RFS 0627 19 rev.*

RFS 0577 19 rev.*

RFS 1051 18 rev.*

RFS 0580 17 rev.*

RFS 1322 17 rev. BIOL - endocrinol; animal physiol;animal biochem
TRANSFORMING GROWTH FACTOR-BETA; INHIBIN FAMILY OF PROTEIN HORMONES; INTRAGONADAL REGULATION

RFS 0418 17 rev. BIOL - endocrinol; molec. biol
GLUCOCORTICOID RECEPTOR; MOLECULAR MECHANISMS OF THYROID-HORMONE ACTION; DNA-BINDING DOMAIN; NEGATIVE REGULATION

RFS 0090 16 rev.*

RFS 2468 16 rev.*

RFS 5495 16 rev.*

RFS 0235 16 rev. BIOL - molec. biol; plant biochem; plant physiol Chem - physical chem
BACTERIAL PHOTOSYNTHETIC REACTION CENTER; PRIMARY CHARGE SEPARATION IN PHOTO-SYSTEM-II; ELETRON-TRANSFER DYNAMICS; RUTHENIUM (II) COMPLEXES; WATER OXIDATION

RFS 2202 15 rev.*

RFS 1245 15 rev. BIO - neurobiol PHARMACOL
HYDROXYTRYPTAMINE RECEPTORS; HUMAN-BRAIN MEMBRANES; SEROTONIN BINDING-SITES

RFS 0104 14 rev.*

RFS 0237 14 rev.*

RFS 1398 13 rev.*

RFS 1708 13 rev.*

RFS 4000 13 rev. BIOL - animal biochem; endocrinol
INSULIN-LIKE GROWTH FACTOR-I; TYROSINE PHOSPHORYLATION; KINASE DOMAIN; RECEPTOR INTERNALIZATION

RFS 0170 12 rev.*

RFS 0778 12 rev.*

RFS 1771 12 rev. BIOL - animal physiol
RYANODINE RECEPTOR-CA-2+RELEASE CHANNEL COMPLEX OF SKELETAL-MUSCLE SARCOPLASMIC-RETICULUM; CALCIUM REGULATION; MAMMALIAN CARDIAC VENTRICULAR CELLS

RFS 3817 12 rev. BIOL - molec. biol; ;bioorganic chem
POLYMERASE CHAIN-REACTION AMPLIFIED GENOMIC DNA; DIRECT SEQUENCING

RFS 0744 11 rev.*

RFS 1095 11 rev.*

RFS 1721 11 rev.*

RFS 1392 10 rev. BIOL - bioorganic chem CHEM - analytical chem; physical chem PHYS-spectroscopy
NUCLEAR MAGNETIC-RESONANCE SPECTROSCOPY; C-13 ASSIGNMENTS; HYBRID DISTANCE GEOMETRY DYNAMICAL LSIMULATED ANNEALING

RFS 1115 10 rev.*

RFS 1513 9 rev.*

RFS 2487 9 rev.*

RFS 2363 8 rev.*

RFS 2862 8 rev.*

RFS 0055 8 rev. BIOL - animal biochem CHEM - analytical chem
IMPLANTABLE GLUCOSE SENSOR; AMPEROMETRIC ENZYME ELECTRODE; FIBER-OPTIC PROBES; PH OPTRODE

RFS 0812 8 rev. BIOL - Bioorganic chem; molec. biol CHEM - analytical chem
PULSED-FIELD GEL-ELECTROPHORESIS OF DNA; ACCURATE SIZE DETERMINATION; SEPARATION PROCESS

RFS 3943 8 rev. BIOL - immunol
MURINE INTERLEUKIN-4 RECEPTOR; REGULATION OF IGE SYNTHESIS; HELPER T-CELLS; RECOMBINANT IFN-GAMMA; CD 23 EXPRESSION; DIFFERENTIAL ACTIVATION

RFS 4141 8 rev. BIOL - immunol; cyt/hist
DNA FRAGMENTATION; APOPTOSIS IN THE PROGRAMMED CELL-DEATH; IMMATURE THYMOCYTES

RFS 0264 7 rev.*

RFS 0444 7 rev.*

RFS 0887 7 rev.*

RFS 1053 7 rev.*

RFS 0174 7 rev. BIOL - neurobiol INFO PSYCHOL
NEURAL NETWORKS; ASSOCIATIVE MEMORY; NEURONAL MODELS OF COGNITIVE FUNCTIONS

RFS 2261 7 rev. BIOL -bioorganic chem CHEM - physical chem PHYS - crystallography
PROTEIN STABILITY; 2.0-A RESOLUTION; CRYSTALLOGRAPHIC REFINEMENT; SIMULATED ANNEALING

RFS 3847 7 rev. BIOL - neurobiol PHARMACOL
DELTA OPIOID RECEPTORS; MODULATION OF MU-MEDIATED ANTINOCICEPTION; RAT SPINAL CORD; PHENCYCLIDINE BINDING-SITES; SELECTIVE POTENTIATION

*Appendix 3*

***Supertopical RFS in chemistry***

*, ** - RFS related also to medical and biological sciences are marked by * and **, respectively.

RFS 0235 16 rev.**

RFS 0280 15 rev. CHEM - physical chem; electrochem PHYS - chemical phys
SOLVENT DYNAMICAL EFFECTS IN ELECTRON-TRANSFER; BLUE COPPER PROTEINS; INTRAMOLECULAR CHARGE SEPARATION

RFS 0435 10 rev. CHEM - Thermochem PHYS - superconductivity MATERIALS - phys. chem. mater.; ceramics
HIGH-TC SUPERCONDUCTORS; MIXED-VALENCE COPPER OXIDES; 110 K PHASE IN PB-SUBSTITUTED BI-SR-CA-CU-O

RFS 1392 10 rev.**

RFS 0812 8 rev.**

RFS 2862 8 rev.*

RFS 1069 8 rev. CHEM - surface chem; physical chem PHYS - chemical phys; physic properties of materials; mechanical properties of solids MATERIALS - metals/alloys
HYDROGEN EMBRITTLEMENT OF NI3AL ALLOYS; GRAIN-BOUNDARIES IN BCC METALS; NI3(AL,X) SINGLE-CRYSTALS; ORIENTATION DEPENDENCE; (100) PLANES; STRENGTH ANOMALY

RFS 0055 8 rev.**

RFS 2261 7 rev.**

RFS 0640 7 rev. CHEM - analytical chem; stereochem PHARMACOL
STEREOSELECTIVE HIGH-PERFORMANCE LIQUID-CHROMATOGRAPHIC ASSAY; CHIRAL STATIONARY PHASE; RESOLUTION OF ENANTIOMERS; POLYMORPHYC DEBRISOQUINE HYDROXYLATION

RFS 1778 7 rev. CHEM - physical chem; surface chem PHYS - physicl properties of materials; surface phys MATERIALS phys. chem. mater.
SCANNING TUNNELING MICROSCOPY; GRAPHITE SURFAC LATTICE ATOMIC IMAGE; SI(111)7X7 DIMER ADATOM STACKING-FAULT STRUCTURE

RFS 0403 6 rev. CHEM - physical chem PHYS - chemical phys; low dimension systems; spectroscopy
LENNARD-JONES CLUSTERS; STABILITY OF SMALL METAL PARTICLES; PHOTOIONIZATION SPECTRA; ELECTRONIC STATES; POTENTIAL-ENERGY SURFACES

RFS 0589 6 rev. CHEM - physical chem BIOL - bioorganic chem
DNA CONFORMATION; CHROMATIN FOLDING; MOBILE HISTONE TAILS IN NUCLEOSOMES

RFS 1046 6 rev. CHEM - polymer chem PHYS - spectroscopy; chemical phys GEO - geochem
P-31 SOLID-STATE NUCLEAR MAGNETIC-RESONANCE SPECTROMETRY; C-13 NMR; COAL PYROLISIS; NITROGEN CHEMICAL SHIELDING ANISOTROPIES

RFS 1761 6 rev. CHEM - analytical chem PHYS - spectroscopy BIOL - neuropeptides; animal biochem
FAST ATOM BOMBARDMENT MASS-SPECTROMETRY; BOVINE POSTERIOR INTERMEDIATE PITUITARY PEPTIDES; SECONDARY ION EMISSION

RFS 3617 6 rev. CHEM - analytical chem ENVIR. SCIENCE SOIL SCIENCE
DISSOLVED ORGANIC-MATTER; MODELING SOLUTE TRANSPORT; SORPTION OF CHLORINATED HYDROCARBONS; SOIL COLUMNS; NATURAL AQUATIC HUMIC SUBSTANCES

RFS 1913 5 rev. CHEM - coord. chem; organic chem; catalysis
ALKENE EPOXIDATION; PORPHYRIN COMPLEXES; CYTOCHROME-P-450 MODELS; BIOMIMETIC ALKANE OXIDATION; FUNCTIONALISATION OF SATURATED-HYDROCARBONS; HEME IRON

RFS 2580 5 rev. CHEM - surface chem MATERIALS - metals/alloys PHYS - surface phys; spectroscopy
EXTENDED ENERGY-LOSS FINE-STRUCTURE SPECTROSCOPY; EPITAXIAL SILICIDE INTERFACES; SHORT-RANGE LOCAL ORDER; NICKEL SURFACE; (111) SI

RFS 2951 5 rev. CHEM - analytical chem BIOL - genosystematics; animal genetics; zool MED - medical genetics
DNA FINGERPRINTING VARIABLE NUMBER OF TANDEM REPEAT (VNTR) SEQUENCES; 14 GENOMIC PROBES FOR HIGHLY POLYMORPHIC LOCI

RFS 3716 5 rev. CHEM - physical chem PHYS - optics; chemical phys MATERIALS - composit
NONLINEAR OPTICAL-PROPERTIES; 2ND-HARMONIC GENERATION; ORGANIC MATERIALS

RFS 3937 5 rev. CHEM - analytical chem PHARMACOL
CAPILLARY ZONE ELECTROPHORESIS; ELECTROKINETIC CHROMATOGRAPHY FOR DRUG ANALYSIS; RAPID SEPARATION OF DNA RESTRICTION FRAGMENTS

*Appendix 4*

***Hypertopical RFS in physics***

*, **, *** - RFS related also to medical, biological and chemical sciences are marked by *, ** and *** respectively.

RFS 0280 15 rev.***

RFS 0435 10 rev.***

RFS 1392 10 rev.**

RFS 0041 9 rev. PHYS - quantum dynamics
HIGGS BOSON; QUARK MASS MATRICES; CP VIOLATION; PHYSICS OF B-MESONS; WEAK DECAYS; CHARGE ASYMMETRIES

RFS 1069 8 rev.***

RFS 1778 7 rev.***

RFS 0300 7 rev. PHYS - superconductivity; low dimension systems
HIGH-TC SUPERCONDUCTIVITY; FRACTIONAL STATISTICS; TWO-DIMENSIONAL QUANTUM HEISENBERG-ANTIFERROMAGNET; CHERN-SIMONS TERM; RESONATING-VALENCE-BOND STATE

RFS 2261 7 rev.**

RFS 0403 6 rev.***

RFS 1046 6 rev.***

RFS 1761 6 rev.***

RFS 0009 6 rev. PHYS - chaos; low dimension systems MATERIALS - metals/alloys
DIFFUSION-LIMITED AGGREGATION; TWO DIMENSIONAL FRACTAL GROWTH; CHAOTIC ATTRACTORS

RFS 0557 6 rev. PHYS - chemical phys; physical properties of materials
COLD FUSION; PAIR INTERACTION ENERGY OF HYDROGEN ISOTOPES; PALLADIUM CATHODES IN ELECTROLYTIC CELLS

RFS 2580 5 rev.***

RFS 3716 5 rev.***

RFS 0010 5 rev. PHYS - cosmol ASTRON/ASTROPHYS
COLD DARK MATTER; TOPOLOGY OF LARGE-SCALE STRUCTURE; COSMOLOGICAL VELOCITY CORRELATIONS; EXTENDED INFLATION; COSMIC MICROWAVE ANISOTROPY; ABELL CLUSTERS

RFS 0198 5 rev. PHYS - quantum dynamics
SKYRME MODEL; CHIRAL ANOMALIES; MESON SOLITION SCATTERING; NUCLEUS-NUCLEUS COLLISIONS

## References


1. Haitun, S.D. (1983). Naukometriya : Sostoyanie I Perspektivy (Scientometrics : State of Art and Prospects). *Nauka* : Moscow.
2. King, J. (1987). A review of bibliometric and other science indicators and their role in research evaluation. *Journal of Information Science* **13 (5)** 261-276.
3. Mirsky, E.M. (1980). Mezhdisciplinarnye Issledovaniya I Disciplinarnaya Organizaciya Nauki (Interdisciplinary Research and Disciplinary Organization of Science). *Nauka* : Moscow.
4. Marshakova, I.V. (1973). Systema svyazei Mezhdu dvumya dokumentami, Postroyennaya na osnove ssylok (Po dannym Science Citation Index). (Document Linking System Based on References Taken from" Science Citation Index"). Nauchno-Tehnicheskaya Informatsiya, Ser. 2. *Informatsionnye processy I Sistemy* **(6)** 3-8.
5. Small, H. (1973). Co-citation in the scientific literature : A new measure of the relationship between two documents. *Journal of the American Society for Information Science* **24**, 256-269.
6. Marshakova, I.V. (1988). Sistema Tzitirovaniya Nauchnoi Literatury kak Sredstvo Siezhenia za Razvitiem Nauki (System of Citations in Scientific Literature as the Means of Science Monitoring). *Nauka* : Moscow.
7. Small, H., E. Garfield. (1985). The Geography of Science : Disciplinary and National Mappings. *Journal of Information Science* **11**, 147-159.
8. Garfield, E. (1988). New Tools for Studying the History of Science. *Current Contents* (No 4, January 25), 3-11.
9. Index to Scientific Reviews, 1991. *ISI* : Philadelphia. 5-7.
10. Salton, G., D. Bergmark. (1979). A Citation Study of Computer Science Literature. *IEEE Transactions of Professional Communication*, 22, 146-158.
11. ISI Atlas of Science : *Pharmacology*, 1988. ISI : Philadelphia, ISI.
12. Rozhkov, S.A., Yavelov, B.E., S.A. Melamed. (1986). Czechoslovakian *Journal of Physics B.*, **36**, 118-120.

# *Scientometric Portrait of Nobel Laureate S. Chandrasekhar*

**B.S. Kademani, V. L. Kalyane and A. B. Kademani**
Turning Point Centre, 2:2, C 5-29, Sector 5,CBD Belapur,
Konkan Bhavan, New Bombay 400 614.

**S. Chandrasekhar, the well known Astrophysicist is widely recognised as a very successful Scientist. His publications were analysed by year, domain, collaboration pattern, channels of communications used, keywords etc. The results indicate that the temporal variation of his productivity and of the types of papers published by him is of such a nature that he is eminently qualified to be a *role model* for the younger generation to emulate. By the end of 1990, he had to his credit 91 papers in *Stellar Structure and Stellar atmospheres*. 80 papers in *Radiative transfer and negative ion of hydrogen*, 71 papers in *Stochastic, statistical hydromagnetic problems in physics and astronomy*, 11 papers in *Plasma Physics*, 43 papers in *Hydromagnetic and Hydrodynamic Stability*, 42 papers in *Tensor-virial theorem*, 83 papers in *Relativistic astrophysics*, 61 papers in *Mathematical theory of Black holes and colloiding waves*, and 19 papers of *general* interest.**

**The highest Collaboration Coefficient was 0.5 during 1983-87. Productivity coefficient was 0.46. The mean Synchronous self citation rate in his publications was 24.44. Publication density was 7.37 and Publication concentration was 4.34.**

*Keywords/Descriptors:* **Biobibliometrics; Scientometrics; Bibliometrics; Collaboration; Individual Scientist; Scientometric portrait; Sociology of Science, History of Science.**

## 1. Introduction

Subrahmanyan Chandrasekhar was born in Lahore (then a part of British India) on 19 October 1910. He had his early education by private tution till he was twelve. He had his high school education in the Hindu High School, Triplicane during the years 1922-25. He had his University education at the Presidency College during 1925-30 and received his Bachelor's degree, B.Sc.(Hon.), in physics in June 1930. He was awarded a Government of India Scholarship for graduate studies in Cambridge, England in July 1930 to work in theoretical physics, more specifically in the theory of stellar structure, the field which was dominated then by Arthur Eddington.

He became a research student under the supervision of professor R. H. Fowler (who was responsible for his admission to Trinity College). On the advice of Professor P. A. M. Dirac, he spent his three undergraduate years at the Institute for Theoretisk Fysik in Copenhagen.

He was awarded Ph. D. degree by Cambridge University in 1933. He was elected as a fellow at Trinity College for the period 1933-37. He was a Research Associate at Yerkes Observatory, Chicago during 1936-38. He became Assistant Professor, Chicago University during 1938-41, Associate Professor (1942-43), Professor (1943-47), Distinguished Service Professor of Theoretical Astrophysics (1947-52), Morton D. Hull Distinguished Service Professor of Theoretical Astrophysics (1952-1986). He was Professor Emeritus (1986-95). He died of heart failure in Chicago on 21st August 1995.

He was an editor of the journal *Astrophysical Journal* during 1952 - 1971. When he took over, the journal was nothing more than a private journal of Chicago University. By the time he resigned it

had become an official journal of the American Astronomical Society.

There is no doubt that he was influenced by his illustrious uncle Sir C. V. Raman the Nobel Laureate for 1930 well known for his invention on *Raman Effect*.

Many honours and awards were bestowed on him in recognition of his contribution in the field of Astrophysics. Important ones being :

1. Fellow of the Royal Society of London - 1944.
2. Adams prize (Cambridge University) - 1947.
3. Bruce Medal of the Astronomical Society of the Pacific - 1952.
4. Gold Medal of the Royal Astronomical Society of London - 1952.
5. Elected to the National Academy of Sciences - 1955.
6. Rumford Medal of the American Academy of Arts and Sciences - 1957.
7. Srinivasa Ramanujan Medal of the Indian National Science Academy - 1962.
8. Royal Medal of the Royal Society - 1962.
9. National Medal of Science (United States) - 1968.
10. Padma Vibhusan Title (India) - 1968.
11. Henry Draper Medal of the National Academy of Sciences - 1971. Smoluchowski Medal (Polish Physical Society).
12. Dannie Heinemann Prize of American Physical Society - 1974.
13. Nobel Prize - 1983.
14. Dr. Tomalla Prize (ETH, Zurich).
15. Copley Medal of Royal Society -1984.
16. R. D. Birla Award - 1984.
17. Vainu Bappu Medal of the Indian National Science Academy - 1985.

He was also a member of following Academies :

- National Academy of Sciences
- American Academy of Arts and Sciences
- Royal Astronomical Society
- American Astronomical Society
- Royal Society

As a student Chandrasekhar had received as a prize, Eddington's famous book *The Internal Constitution of the Stars* which left a lasting impression on young Chandrasekhar's mind. This perhaps was responsible for his taking up research in the field of Astronomy and Astrophysics.

Chandrasekhar's contribution is particularly multi-faceted and covers many aspects of the evolution of stars. An important part of his work is a study concerning the problems of stability in different phases of their evolution. He has studied relativistic effects, which became important because of the extreme conditions which arise during the later stages of the star's development. One of Chandrasekhar's most well known contributions is his study of the *Structure of White Dwarfs*. In recent years he had worked on *The Mathematical Theory of Black Holes*.

His books : *An Introduction to the Study of Stellar Structure* (1939); *Principles of Stellar Dynamics* (1942); *Radiative Transfer* (1950); *Plasma Physics* (1960); *Hydrodynamic and Hydromagnetic Stability* (1961); *Ellipsoidal Figures of Equilibrium* (1969); *and The Mathematical Theory of Black Holes* (1983) have become classics in the fields of Astronomy and Space research.

He had wide interest in music and literature, and he wrote a book entitled *Truth and beauty : aesthetics and motivations in science*. His final book was a commentary on, *Newton's principia for the common reader*, published early 1995.

Chandrasekhar had to face several humiliating experiences in the hands of noted astronomers which did not dampen his zeal, spirit and scientific temper which was in him by birth.

It is noteworthy to mention that Chandrasekhar's students Tsung - Dao Lee and Chen Ning Yang were awarded Nobel prize in physics for 1957 at their age 31 and 35 respectively for their investigation of the so-called parity laws which led to the discoveries regarding elementary particles. S. Chandrasekhar was awarded Nobel prize jointly with A. Fowler for his contributions on the evolution of stars in 1983 when he was 73 years of age mainly for his well known discovery 'Chandrasekhar limit' named after him which states that *Some stars are too massive to become white dwarf star which is formed with mass greater than a limiting value (1.4 M.)*.

The Nobel prize is regarded not only by laymen but also by scientists as the most honorific

recognition of scientific achievement. The prestige of the Nobel prize is so great that it enhances the standing of nations and institutions as well as the reputation of its *laureates* [1-3]. His works have been well documented [4-9].

Citation analysis of some important contributions of S. Chandrasekhar has already been carried out [10]. This study deals with six citation classics which have been identified based on the citations received to the papers of S. Chandrasekhar. These six papers received 53% of total 10,359 citations during the period under study and concluded that there is a high correlation in quantity, quality of works, citedness and receiving honours and awards.

## 2. Objectives

Objectives of present work are to highlight quantitative aspects of the research communications :

(a) authorship pattern,
(b) domainwise contribution,
(c) author productivity,
(d) use of channels of communication
(e) bibliographic characteristics of publications, and
(f) documentation of keywords from title

The main concept of working on individual scientist is to provide *Role Model Scientist* for younger generation of science graduates and post graduates who have become frustrated due to various reasons. To show them light or hope or new direction towards success. Success of others may teach many things to follow their path. The attempt however small, may prevent them to make suicide of their creativity, and channelise aggressive energies of youth towards constructive ideas [11].

A successful scientist is one who keeps on publishing his ideas or works. To be successful, capacity to communicate effectively and efficiently is most fundamental. Scientific communications have their own regime and regimentation crossing all political and geographical boundaries.

## 3. Methodology

Scientific publication, seems to provide the best available basis for measuring research output. One of the first writers to suggest scientific publication as a measure of research productivity was Nobel Laureate William Shokley [12] who was interested in measuring research.productivity among individual within a group by analysing their publications. A few studies have been recently published on individual scientists [10, 13-41].

Bibliographic details of the publications of S. Chandrasekhar were documented on cards from the list appended at the end of volume six of Selected papers of S. Chandrasekhar [42] and sorting was done as per requirements of the study.

Normal count procedure [43] was followed. Full credit was given to each author regardless of whether he happens to be the first or the last author. It is widely recognised that scientists all over the world look at their own papers exclusively in that way. Similarly titles of the articles were analysed and one score was alloted for each keyword, subject, journal, etc.

The degree of collaboration [44] in a discipline was defined as the ratio of the number of collaborative research papers to the total number of research papers published in the discipline during a certain period of time (Figure 3).

Vinkler [45] defined (Table - 3) Publication Density as the ratio of the total number of papers published to the total number of journals in which the papers were published, and Publication Concentration as the ratio in percentage of the journals containing half of the papers published to the total number of journals in which those papers were published during the period under study.

Sen and Gan [46] defined Productivity Coefficient as the ratio of 50 percentile age to the total productivity age.

Lawani [47] defined (Table - 8) Synchronous Self Citation rate :

$$\text{Synchronous rate} = \frac{\text{Self references in an article}}{\text{Total no. of references in an article}} \times 100$$

Frequency of keywords from the titles of the articles were recorded. Data obtained from above study were presented in tables and figures.

## 4. Results and Discussion

During 1928 - 1990 S. Chandrasekhar had published 380 research communications in the

following domains :

**A** = Steller structure and stellar atmospheres
**B** = Radiative transfer and negative ion of hydrogen
**C** = Stochastic, Statistical hydromagnetic problems in physics and astronomy
**D** = Plasma physics
**E** = Hydromagnetic and hydrodynamic stability
**F** = Tensor - Virial theorem
**G** = Relativistic astrophysics
**H** = Mathematical theory of Black holes and colloiding waves
**I** = General

**Table 1** shows author productivity and distribution of authors in various domains. The research group of S. Chandrasekhar has the credits of number of authorships in various domains : A(91), B(80), C(71), D(11), E(43), F(42), G(83), H(61), and I(19). Total number of authors in the research group were 48. Researchers and their authorships in collaboration with S. Chandrasekhar in Chronological order of their association (in first publication with S. Chandrasekhar) are depicted in Figure 1. Most active researchers and their contributions with S. Chandrasekhar were N. R. Lebövitz (22) and D. D. Elbert (15). Other active collaborators with S. Chandrasekhar and their contributions were B. C. Xanthopoulos (10), G. Münch (8), and F. H. Breen (6). Other collaborators having three papers each were 12, two papers each were 20, and single paper each were 28.

B. C. Xanthopoulos had collaborated with S. Chandrasekhar in the domain H only. D. D. Elbert had collaborated with him in the A, B, E, F and G. whereas N. R. Lebövitz had collaborated in the domains E, F, G and H.

Domainwise Collaboration of S. Chandrasekhar with his 47 Collaborators and their status of authorship in various domains is provided in Table - 2. S. Chandrasekhar had single authored papers in various domains as A(63), B(34), C(39), D(2), E(30), F(14), G(43), H(25) and I(17). He had collaborations in various domains as A(28), B(46), C(32), D(9), E(13), F(28), G(40), H(36) and I(2).

**Table 1. Author Productivity and Distribution of Authors in Various Domains**

| No. of papers | Domainwise Authorships | | | | | | | | | No. of Authors | Total No. of Authorships | Prominent Collaborators |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | **A** | **B** | **C** | **D** | **E** | **F** | **G** | **H** | **I** | | | |
| 1 | 8 | 4 | 5 | - | 2 | 2 | 5 | 1 | 1 | 28 | 28 | |
| 2 | 5 | 7 | 4 | | | | 4 | | | 10 | 20 | |
| 3 | | | | 6 | | | | 6 | | 4 | 12 | |
| 6 | | 6 | | | | | | | | 1 | 6 | Breen, F. H. |
| 8 | | 1 | 7 | | | | | | | 1 | 8 | Münch, G. |
| 10 | | | | | | | | 10 | | 1 | 10 | Xanthopoulos, B. C. |
| 15 | 1 | 7 | | | 4 | 1 | 2 | | | 1 | 15 | Elbert, D. D. |
| 22 | | | | | 1 | 11 | 9 | 1 | | 1 | 22 | Lebovitz, N. R. |
| 380 | 77 | 55 | 55 | 5 | 36 | 28 | 63 | 43 | 18 | 1 | 380 | Chandrasekhar, S. |
| Total | 91 | 80 | 71 | 11 | 43 | 42 | 83 | 61 | 19 | 48 | 501 | |

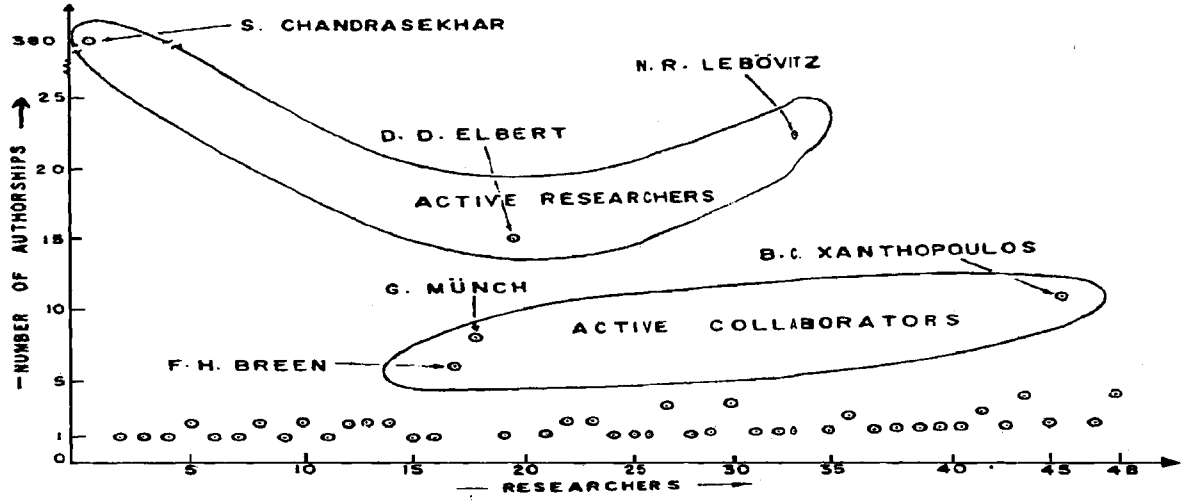



**Fig. 1. Researchers Association in Chronological Order**

Percentagewise contribution of authorships to various domains include A(18.16), G(16.57), B(15.97), C(14.17), H(12.18), E(8.58), F(8.38), I(3.79) and D(2.20).

He had published two papers in collaboration with the Nobel Laureate Enrico Fermi in the domain C during 1953.

His domainwise cumulative number of publications, his age, and scientific career advancements are depicted in Figure - 2.

A feature of Chandrasekhar's career was that he would write a very long series of papers in a particular research field and once he felt that he has exhausted everything in that particular field then he would summarise the whole work in the form of an authoritative monograph and then move on to another field.

It is clearly visible from the Figure - 2 that Chandrasekhar shifted his research domains very frequently. That is how he continued to remain very active in the field.

How does one not become an expert? Astrophysicist S. Chandrasekhar gave a remarkable television interview a few years ago.

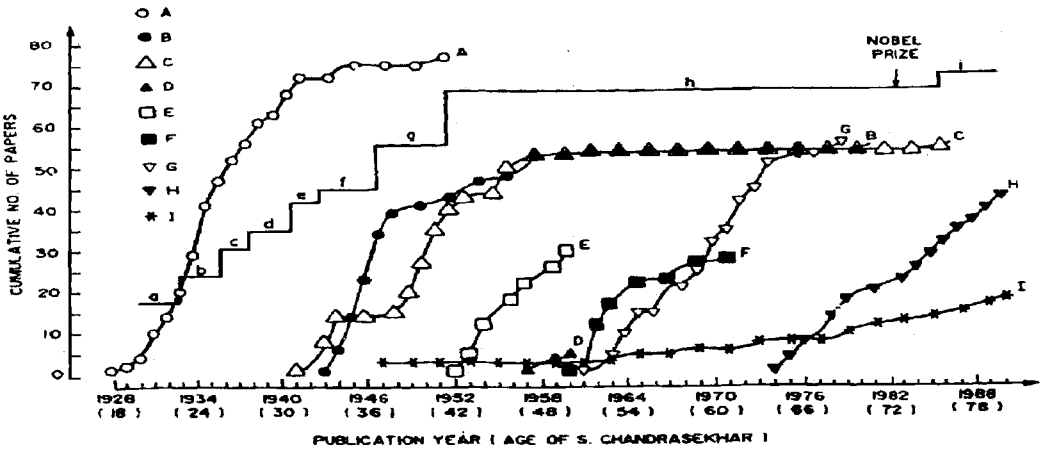



**Fig. 2. Domainwise Publication Productivity of S. Chandrasekhar**

Scientific Career Advanements : a = Govt. of India Scholar, Cambridge Univ.; b = Fellow, Trinity College, Cambridge Univ.; c = Res.Assoc., Yerkes Observatory, Chicago; d = Asst. Prof., Chicago Univ.; e = Assoc. Prof., Chicago Univ.; f = Prof., Chicago Univ., g = Disting. Sergice Prof. of Theoretical Astrophysics; h = Morton D. Hull Disting. Service Prof. of Theoretical Physics; i = Prof. Emeritus.

He had lead a Scientific Career notable for a rate of productivity that has not slowed down at all into his 70s. When asked how he has avoided the drop in creativity and productivity that plagues many scientists, he replied that approximately every seven years he takes up a new topic. He found that he would run out of new ideas after working in an area for too long. This pattern lead him to tackle such topics as the dynamics of stellar systems, white dwarfs, relativity and radiative transfer. Although all these subjects are in astrophysics, they are different enough to present unique problem [48].

With advances in research, vision of scientist expands, one island of superspecialisation or micro-theme expands and bridges connection with another island of micro-theme. A creative researcher travels through the bridges to other island and instead of returning to his original island such scientists continue to colonise and work on the latest theme of fresh interest due to intrinsic motivations which accelerate vigorous activities further and exploit new idea resources. Natives (Super Specialists) of that island (micro-theme) may have become complacent because of inbreeding of their thoughts. Creativity predominant in scientists is of two types: Convergent thinking creativity and Divergent thinking creativity [49].

The most productive researchers have changed research field more often than the less productive researchers [50].

However, no two individuals can be identical in their creativity i.e. each individual scientist has his/her own *Stereotype* [51] and *Mentor* [52-53]. Hence, attempts to generalise may fail.

With time and advances in research a creative scientist builds-up his/her own research team. As pioneer has already established himself he becomes pivote around whom entire team revolves in spirals (not in circle, because in circle there is no advancement as end meets the beginning) the direction and progressive movement of the spiral shifts its progress slowly to next higher stratum every time. Leader or conductor of the orchestra has the responsibility to bring forth best in every individual. Thus with advancing age many individuals and groups join such an individual for their own individual success as well as to satisfy affiliation needs.

Quinquennial publication productivity of S. Chandrasekhar is shown in Figure - 3. Highest Collaboration coefficient was 0.5 during 1983-87. His productivity coefficient was 0.46 which is clear indication of his high publication productivity behaviour during early period of 29 years of research publication career.

His first paper was published in 1928 in *Indian*

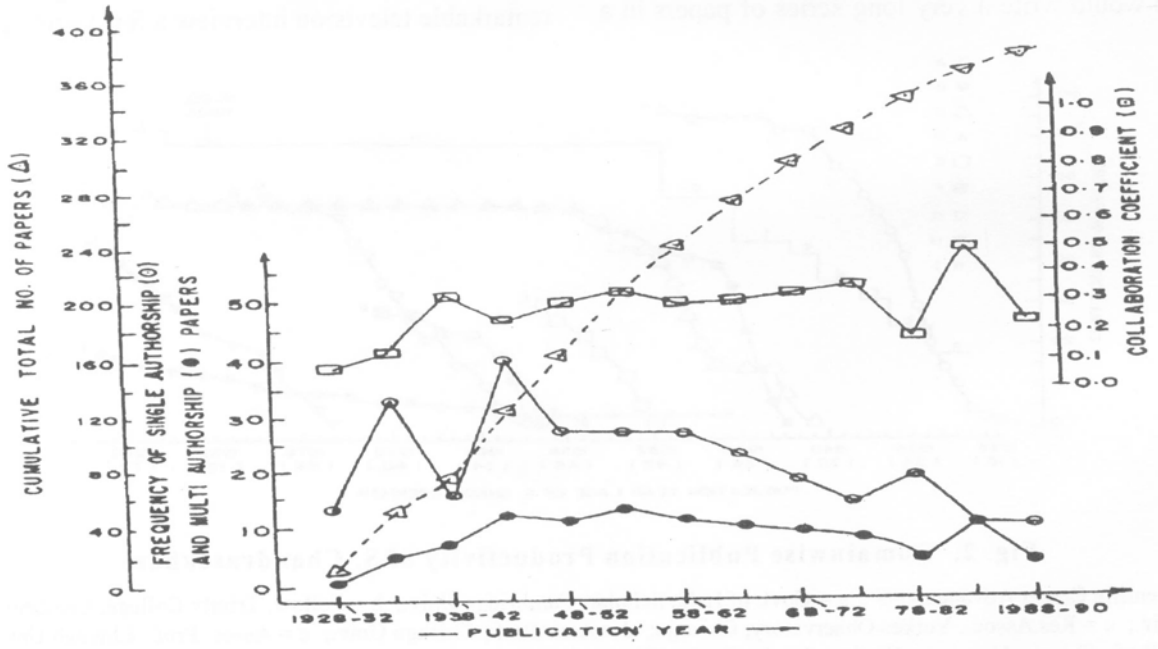



**Fig. 3. Quinquennial Publication Productivity of S. Chandrasekhar**

*Journal of Physics* at 18 years of his age in the domain A.

Distribution of his 339 publications were in 46 journals, 16 chapters in books, 16 conference proceedings and nine books.

Journalwise scattering of publications of S. Chandrasekhar in various journals is provided in Table - 3. He has published 139 papers in *The Astrophysical Journal*, 59 papers in *Proceedings of the Royal Society A*, 31 papers in *Monthly Notices of the Royal Astronomical Society*, 14 papers in *Proceedings of the National Academy of Sciences*. He has published 10 papers in the journal *Observatory*.

**Table 3. Journalwise Scattering of Publications of S. Chandrasekhar**

| Sl. No. | Journal | No.of Papers | Percentage | Cumulative percentage | Period of Journal usage | | | SCI JCR 1992 | | Country of publication |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | FPY | LPY | TOTAL | IF | II | |
| 1. | Astrophys. J. | 139 | 41.0 | 41.0 | 1931 | 1975 | 45 | 2.931 | 0.152 | US |
| 2. | Proc. Roy. Soc. A. | 59 | 17.4 | 58.4 | 1929 | 1990 | 62 | 1.673 | 0.289 | UK |
| 3. | Month. Notic. Roy. Astron. Soc. | 31 | 9.1 | 67.5 | 1931 | 1984 | 54 | 2.579 | 0.460 | UK |
| 4. | Proc. Natl. Acad. Sci. | 14 | 4.1 | 71.6 | 1956 | 1963 | 8 | 10.480 | 1.436 | US. |
| 5. | Observatory | 10 | 3.0 | 74.6 | 1933 | 1972 | 40 | 0.814 | 0.227 | UK |
| 6. | Philos. Mag. | 9 | 2.7 | 77.3 | 1930 | 1957 | 28 | - | - | UK |
| 7. | Nature | 7 | 2.1 | 79.4 | 1935 | 1990 | 56 | 22.139 | 5.224 | UK |
| 8. | Phys. Rev. | 7 | 2.1 | 81.4 | 1949 | 1971 | 23 | - | - | US |
| 9. | Proc. Camb. Phil. Soc. | 6 | 1.8 | 83.3 | 1935 | 1955 | 21 | - | - | UK |
| 10. | Zeit. Astrophys. | 6 | 1.8 | 85.1 | 1931 | 1937 | 7 | - | - | Gemany |
| 11. | Rev. Mod. Phys. | 5 | 1.5 | 86.6 | 1943 | 1949 | 7 | 14.071 | 1.759 | US |
| 12. | Science | 4 | 1.2 | 87.8 | 1944 | 1981 | 38 | 20.967 | 3.600 | US |
| 13. | Am. J. Phys. | 3 | 0.9 | 88.7 | 1969 | 1972 | 4 | 0.563 | 0.134 | US |
| 14. | Contemp. Phys. | 3 | 0.9 | 89.6 | 1973 | 1980 | 8 | 1.541 | 0.111 | US |
| 15. | Ann. Phys. | 2 | 0.5 | 90.1 | 1957 | 1958 | 4 | 0.608 | 0.509 | UK |
| 16. | Mathematika | 2 | 0.5 | 90.6 | 1954 | 1957 | 4 | 0.694 | 0.000 | UK |
| 17. | Philos. Trans. Roy. Soc. London | 2 | 0.5 | 91.1 | 1950 | 1952 | 3 | 1.182 | 0.237 | UK |
| 18. | Proc. Am. Philos. Soc. | 2 | 0.5 | 91.6 | 1939 | 1964 | 26 | - | - | US. |
| 19. | Am. Math. Monthly | 1 | 0.3 | 91.9 | 1954 | 1954 | 1 | 0.193 | 0.101 | US |
| 20. | Ann. New York Acad. Sci. | 1 | 0.3 | 92.2 | 1943 | 1943 | 1 | 0.830 | 0.141 | US |
| 21. | Astrofisika | 1 | 0.3 | 92.5 | 1988 | 1988 | 1 | - | - | Russia |
| 22. | Astron. J. Sov. Union | 1 | 0.3 | 92.8 | 1934 | 1934 | 1 | - | - | Russia |
| 23. | Astrophys. Norvegic | 1 | 0.3 | 93.1 | 1964 | 1964 | 1 | - | - | Norway |
| 24. | Bull. Am. Acad. Arts & Sci. | 1 | 0.3 | 93.4 | 1989 | 1989 | 1 | - | - | US |
| 25. | Bull. Am. Math. Soc. | 1 | 0.3 | 93.7 | 1947 | 1947 | 1 | 0.857 | 0.137 | US |

continued...

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 26. Can. J. Phys. | 1 | 0.3 | 94.0 | 1951 | 1951 | 1 | 0.461 | 0.099 | Canada |
| 27. Commun.Pure Appl. Math. | 1 | 0.3 | 94.3 | 1967 | 1967 | 1 | 1.080 | 0.167 | US |
| 28. Curr. Sci. | 1 | 0.3 | 94.6 | 1985 | 1985 | 1 | 0.253 | 0.075 | India |
| 29. Ind. J. Phys. | 1 | 0.3 | 94.9 | 1928 | 1928 | 1 | - | - | India |
| 30. J. Astrophys. Astron. | 1 | 0.3 | 95.2 | 1984 | 1984 | 1 | 0.464 | 0.105 | India. |
| 31. J. Ind. Math. Soc. | 1 | 0.3 | 95.5 | 1960 | 1960 | 1 | - | - | India |
| 32. J. Math. Anal. Appl. | 1 | 0.3 | 95.8 | 1960 | 1960 | 1 | 0.291 | 0.081 | US |
| 33. J. Math. Mech. | 1 | 0.3 | 96.1 | 1961 | 1961 | 1 | - | - | US |
| 34. J. Ration. Mech. Anal. | 1 | 0.3 | 96.4 | 1954 | 1954 | 1 | - | - | US |
| 35. Mem.Soc.Roy.Soc. deLiege | 1 | 0.3 | 96.7 | 1935 | 1935 | 1 | - | - | France |
| 36. Nord. Astron. Tidskr. | 1 | 0.3 | 97.0 | 1935 | 1935 | 1 | - | - | Norway |
| 37. Notes. Record. Roy. Soc. | 1 | 0.3 | 97.3 | 1976 | 1976 | 1 | - | - | UK |
| 38. Phys. Rev. Lett. | 1 | 0.3 | 97.6 | 1965 | 1965 | 1 | - | - | US |
| 39. Physics Today | 1 | 0.3 | 97.9 | 1971 | 1971 | 1 | - | - | US |
| 40. Proc. Am. Acad. Art. Sci. | 1 | 0.3 | 98.2 | 1957 | 1957 | 1 | - | - | US |
| 41. Proc. Lond. Math. Soc. | 1 | 0.3 | 98.5 | 1959 | 1959 | 1 | 0.649 | 0.188 | UK |
| 42. Pub. Astron. Soc. Pacific. | 1 | 0.3 | 98.8 | 1952 | 1952 | 1 | 1.047 | 0.006 | France |
| 43. Quart. J. Mech. Appl. Math. | 1 | 0.3 | 99.1 | 1955 | 1955 | 1 | 0.567 | 0.115 | UK |
| 44. Quart. J. Roy. Astron. Soc. | 1 | 0.3 | 99.4 | 1980 | 1980 | 1 | 0.514 | 0.042 | UK |
| 45. Scientific Month. | 1 | 0.3 | 99.7 | 1947 | 1947 | 1 | - | - | US |
| 46. Trans. Am. Philos. Soc. | 1 | 0.3 | 100.0 | 1954 | 1954 | 1 | - | - | US |
| Total | 339 | | | | | | | | |

FPY = First Paper Publishing Year; LPY = Last Paper Publishing Year; IF = Impact Factor; II = Immediacy Index; IF and II values taken from SCI Journal Citation Reports 1992.

In the highest Impact Factor (22.139) journal *Nature* he has published seven papers. In other highest Impact Factor (20.967) journal *Science* he has published four papers; *Reviews in Modern Physics* having Impact Factor (14.017) where he has published five papers.

The journals from various countries publishing S. Chandrasekhar's research papers were : 21 from USA (45.65%), 13 from UK (28.26%), four from India (8.70%), whereas from France, Norway and Russia two each, and Canada and Germany one each.

Average Bradford multiplier was 3.46. Publication density was 7.37 and Publication concentration was 4.34.

The frequency and cumulative number of papers published journalwise is depicted in Figure - 4.

Keywords in the titles of the articles were counted. The data are provided in Tables 4 and 5. From the data it is revealed that the titles were very compact and expressive [54].

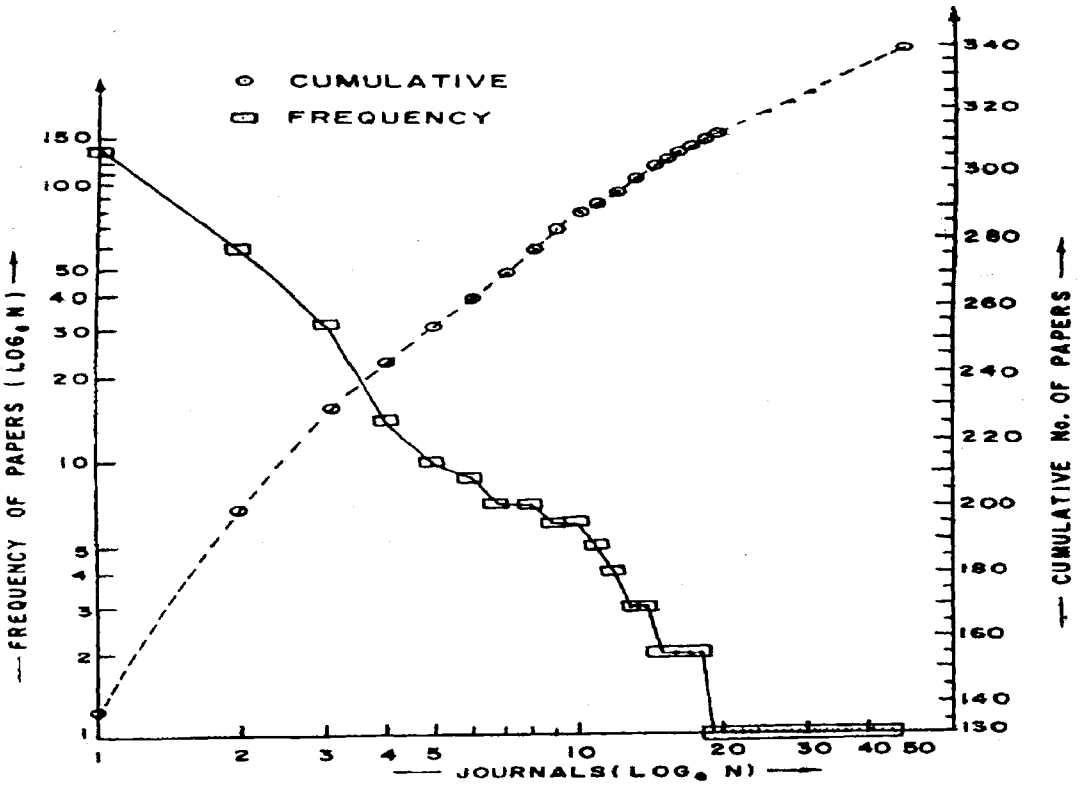



**Fig. 4. Bibliograph on Papers of S. Chandrasekhar**

**Table 4. Length of Article Titles in Terms of Number of Keywords in the Titles of Publications of S. Chandrasekhar**

| No. of Keywords | No. of publications | Percentage |
|---|---|---|
| ONE | 52 | 13.69 |
| TWO | 166 | 43.68 |
| THREE | 99 | 26.05 |
| FOUR | 40 | 10.53 |
| FIVE | 10 | 2.63 |
| SIX | 12 | 3.16 |
| EIGHT | 1 | 0.26 |

**Table 5. Domainwise Keywords in the Titles of Research Papers of S. Chandrasekhar**

| Domain | Total No. of Words | Total No. of Keywords | Mean Per Title | | Proportion of Keywords to No. of Words |
|---|---|---|---|---|---|
| | | | No. of Words | No. of Keywords | |
| A | 623 | 166 | 8.09 | 2.16 | 1 : 3.75 |
| B | 470 | 115 | 8.55 | 2.09 | 1 : 4.09 |
| C | 513 | 140 | 9.33 | 2.55 | 1 : 3.66 |
| D | 39 | 12 | 7.80 | 2.40 | 1 : 3.25 |
| E | 412 | 110 | 11.44 | 3.06 | 1 : 3.75 |
| F | 283 | 81 | 10.11 | 2.89 | 1 : 3.49 |
| G | 781 | 200 | 12.40 | 3.17 | 1 : 3.91 |
| H | 459 | 112 | 10.67 | 2.60 | 1 : 4.10 |
| I | 102 | 28 | 5.67 | 1.56 | 1 : 3.64 |
| Total | 3682 | 964 | 84.06 | 22.48 | 9 : 33.64 |
| Mean | 409.11 | 107.11 | 9.84 | 2.50 | 1 : 3.74 |

The Keywords frequencies in the titles of the papers is provided in Tables 6 and 7. High frequency Keywords were *Stability* (39), *General Relativity* (35), *Radiative equilibrium* (30), *Stellar atmosphere* (30), *Equilibrium* (26), *Magnetic fields* (17), *Stars* (17), *Gaseous masses* (9) *and Kerr black hole* (9).

**Table 6. Keyword Frequencies in the Titles of Papers by S. Chandrasekhar.**

| | | | |
|---|---|---|---|
| Stability | 39 | Colloiding waves | 6 |
| General relativity | 35 | Dynamical friction | 6 |
| Radiative equilibrium | 30 | Gravitational waves | 6 |
| Stellar atmosphere | 30 | Hydrodynamics | 6 |
| Equilibrium | 26 | Interior of stars | 6 |
| Magnetic fields | 17 | Isotropic turbulence | 6 |
| Stars | 17 | Oscillations | 6 |
| Gaseous masses | 9 | Post-Newtonian effects | 6 |
| Kerr black hole | 9 | Thermal instability | 6 |
| Instability | 8 | Uniformly rotating bodies | 6 |
| Perturbation theory | 8 | Absorption coefficient | 5 |
| Rotating cylinders | 8 | Axisymmetric perturbations | 5 |
| Fluctuations | 7 | Brightness | 5 |
| Hydrodynamic stability | 7 | Deformed figures | 5 |
| Negative hydrogen ion | 7 | Equations | 5 |
| Statistical theory | 7 | Gravitational perturbations | 5 |
| Stelar dynamics | 7 | Jacobi ellipsoids | 5 |
| Viscous flow | 7 | | |
| Astronomy | 6 | | |
| Axisymmetric systems | 6 | | |

continued...

| | |
|---|---|
| Maclaurin spheroids | 5 |
| Milky way | 5 |
| Non-radial oscillations | 5 |
| Post-Newtonian approximation | 5 |
| Reissner - Nordstrom black hole | 5 |
| Stellar systems | 5 |
| Turbulence | 5 |
| Virial theorem | 5 |
| Distorted polytropes | 4 |
| Fluid motions | 4 |
| Gravitational field | 4 |
| Gravitational radiation | 4 |
| Gravitational stability | 4 |
| Hydromagnetics | 4 |
| Layer of fluid | 4 |
| Random distribution | 4 |
| Rotating gaseous masses | 4 |
| Stationary | 4 |
| Stellar configurations | 4 |
| Absorption | 3 |
| Absorption lines | 3 |
| Astrophysics | 3 |
| Beauty | 3 |
| Black holes | 3 |
| Coriolis force | 3 |
| Decay | 3 |
| Eddington, A. S. | 3 |
| Ellipsoidal figures | 3 |
| Force - free magnetic field | 3 |
| Four boundary conditions | 3 |
| Incompressible fluid | 3 |
| Ionization | 3 |
| Kerr geometry | 3 |
| Magneto hydrodynamics | 3 |
| Miline, Edward Arthur | 3 |
| New statistics | 3 |
| Opacity | 3 |
| Planetary nebulae | 3 |
| Schwarzschild black hole | 3 |
| Science | 3 |
| Stellar coefficient | 3 |
| Stellar mass | 3 |
| Stellar structure | 3 |
| Super potentials | 3 |
| Time relaxation | 3 |
| Uniformly rotating configuration | 3 |
| Universe | 3 |
| Variable density | 3 |
| Adiabatic invariants | 2 |
| Axisymmetric turbulence | 2 |
| β Canis Majoris stars | 2 |
| Cauchy horizon | 2 |
| Clusters | 2 |
| Coaxial cylinders | 2 |
| Compton Scattering | 2 |
| Congruent Darwin ellipsoids | 2 |
| Conservation laws | 2 |
| Dedekind ellipsoids | 2 |
| Degeneration cores | 2 |
| Density | 2 |
| Differentially rotating configurations | 2 |
| Diffuse reflection | 2 |
| Distribution | 2 |
| Dynamical instability | 2 |
| Dynamical stability | 2 |
| Dynamics | 2 |
| Einstein | 2 |
| Evolution | 2 |
| Expansion of functions | 2 |
| Extended stellar atmospheres | 2 |
| Fluid sphere | 2 |
| Forces | 2 |
| Functions Gn, $m^{(n)}$ | 2 |
| General variational principle | 2 |
| Homogeneous mass | 2 |
| Infinite homogeneous medium | 2 |
| Inhibition of convection | 2 |
| Internal motions | 2 |
| Invariant theory | 2 |
| Inviscid flow | 2 |

continued.....

| | | | |
|---|---|---|---|
| Ionized gas | 2 | Radiative transfer | 2 |
| Kerr metric | 2 | Rate of escape | 2 |
| Low density | 2 | Reflexion | 2 |
| Magnetic rotation | 2 | Reversing layers of stars | 2 |
| Motions | 2 | Riemann ellipsoids | 2 |
| Negative oxygen ion | 2 | Rotating configurations | 2 |
| Neutrino waves | 2 | Rotating fluid sphere | 2 |
| Newtonian gravitation | 2 | Roche ellipsoids | 2 |
| Oort, J. H. | 2 | Schwarzschild limit | 2 |
| Oxygen | 2 | Slow rotation | 2 |
| Polarization | 2 | Solar chromosphere | 2 |
| Post-Newtonian equations | 2 | Sunlit sky | 2 |
| Pressure | 2 | Thermodynamics | 2 |
| Pulsation | 2 | Transmission | 2 |
| Pursuit of science | 2 | Viscous dissipation | 2 |
| Radial acceleration | 2 | White dwarfs | 2 |

**Table 7. Keywords Used Only Once in the Titles of Papers by S. Chandrasekhar**

| | | | |
|---|---|---|---|
| Absorbing atoms | Blanketing effect | Constitution of stars | Distorted polytropes |
| Absorption continuum | Blended absorption lines | Continuous spectrum | Distorted stellar configurations |
| Adjoining media | Boundary value problem | Convection | Double periods |
| Adjoint differential systems | Brownian motion | Coriolis acceleration | Double - star problem |
| Aeasthetics | Carter's theorem | Corona | Einstein's field equations |
| Amplifications | Central Radiation pressure | Correlation | Einstein Maxwell equations |
| Angular distribution | Central temperature | Cosmic magnetic fields | Einstein - Maxwell space times |
| Arbitrary spin | Centrally condensed stars | Cosmological constants | Einstein - Maxwell theory |
| Astrophysical conditions | Centrifugal force | Cowlirg's theorem | Einstein - Vacuum space times |
| Astrophysical interest | Characteristic value problems | Curved channel | Electromagnetic perturbations |
| Astrophysicist | Charged particles | Cylindrical impulsive waves | Electron |
| Atoms | Chromosphere | Cylindrical waves | Electron pairs |
| Axisymmetric gravitational fields | Collapsed configuration | Darwin ellipsoids | Elements |
| Asymmetric homogeneous dynamos | Collision | Degenerate cores | Ellipticity |
| Axisymmetric magnetic fields | Compton effect | Density distribution | Energies |
| Axisymmetric motions | Condensation of stars | Dirac, P. A. M. | Eridani B. |
| Beats | Configurations | Dirac, equation | |
| Bell - Szekers space time | Connective instability | Dirac's views | |
| Binary system | Constants | Dispersion | |
| | | Dissociation formula | |

continued...

Evolution of stars
Extended photospheres
Finite distance
Fluid conductor
Flux integral
Fourier - Bessel-type expansions
Fowler, Ralph Howard
Frequency
Galactic evidences
Gaseous star
Geodesics
Godel's Universe
Ground states of Helium
Ground states of Lithium ions
Ground states of oxygen ions
Hartree field
Heavy viscous fluid
Heisenberg's elementary theory
High order differenial equations
High speed atoms
Higher order virial equations
Highly collapsed configurations
Historical account
Homogeneous compressible model
Homogeneous ellipsoids
Homogeneous turbulent medium
Horizones
Human culture
Hydrogen atom
Hydromagnetic oscillations
Hyperbolic equations
Illumination
Integral equation
Integral theorem
Interface
Invariant theory
Ionization formula
Isothermal cores
Isothermal function
Isotopes
Isotropic scattering
Jacobi sequences
Jeans, janqes hopwood
Jeans sequences
Jeans spheroids
Lane - Emsden function θ 325
Limiting case
Limiting mass
Linear perturbations
Lindbald's theory
Liquids
Maclaurin sequences
Magnetic stars
Main sequence stars
Massless particles
Maxwell's equations
Metric perturbations
Motions of charged particles
Motivations
Moving atmosphere
Multiple frequencies
Nebular luminosity
Nebullium emission
Newtonian theory
Non-axisymmetric mode of oscillation
Non-dissipative couette flow
Non-stationary perturbed systems
Novae
Null dust
Nutku-Halil solution
Odd - parity mode
One-dimensional potential barriers
Onset of convection
Operation
Orthogonal functions
Otto struve
Outer layers
Pencil radiation
Perception of beauty
Perfect fluid
Perturbation analysis
Photographs
Physical content
Physical state of matter
Physical theory
Physics
Pin river
Pinch
Plane gravitational forces
Plane - parallel atmosphere
Plasma
Plasma physics
Post - Galilean transformation
2½ Post - Newtonian equation
Post - Newtonian methods
Post-Newtonian theory of Einstein
Potential barriers
Potentials
Probability distribution
Probability method
Prominences
Quasi normal modes
Radial ejection
Radial oscillation
Radial speed
Radiation reaction
Radial temperature gradient
Radiation
Rajagopal, C. T.
Rayleigh scattering
Recombination
Reflexion coefficients
Relative abundances
Relativistic degeneracy
Relativistic equilibrium
Relativistc instability
Relativistic statistics
Relativistic systems
Relativistic theory
Richtmyer
Roche model
Roots of
$J - (1 + ½) (\lambda n)J1+ ½ (\lambda)$
$J_1 + ½ (\lambda n) J - (\lambda + ½) N = 0$
Roots of
$Yn(\lambda n) Jn(\lambda) - Jn(\lambda n) Yn(\lambda) = 0$
Rotating liquid drops
Rotating stars
Rotational distortion
Rotational problem
Rotational velocities
Rotational masses
Rotating polytropes
Royal Astronomical Society
Rumford Medel Lecture 1957
Russel, H. N.
Scattering of radiation
Schwarzchild geometrics
Scientific attitude
Scientist

continued...

Second harmonic oscillations
Second post - Newtonian equations
Secular stability
Semi-infinite atmospheres
Sequence
Simultneous action
Singularities
Smart, W. M.
Softening of radiation
Solar research
Solar system origin
Source of energy
Spatial correlations
Speed of fluctuations
Spherical shells
Spiral arms
Spiral flow
Star - Streaming
State of matter
Stationary perturbed systems
Statistical basis
Statistical turbulence
Stellar absorption lines
Stellar encounters
Stellar envelops
Stellar evolution
Stellar models
Stellar photospheres
Stellar Scintillation
Stellar statistics
Stochastic problems
Stochastic variation
Strings
Sun
$1^1$ S State of helium
Temperatures
Tensor virial equations
Tensors of high rank
Terrestrial conditions
Teukolsky's equation
Teukolsky - Starbinsky constant
Thermal convection
Theoretical astrophysics
Theory of relativity
Third harmonics
Tidal distortion
Tidal problem
Time - scale
Time - like singularities
Total eclipse of the Sun
Transfer of radiation
Transformation
Transmission coefficients
Trumpler's stars
Truth and beauty
Two black holes
Two centre problem
Two commuting killing fields
Uniform rotation
Uniformly rotating fluid masses
Vacuum metrics
Variable viscosity
Variational methods
Velocity ellipsoid
Victor ambarstsumian
Virial equations
Virial relations
Viscid flow
Viscosity
Viscous liquid globe
Weizsacker theory
Weyl's solution
White dwarf configuration
White dwarf stars
Wolf - Rayet stars
X - functions
Y - functions

These keywords indicate his wide spectrum of interest, materials, methods, instruments used and subjects addressed to in the course of his 63 years of research paper publishing life span.

Domainwise bibliographic characteristics of publications of S. Chandrasekhar are provided in Tables 8 and 9.

It is evident from the publications of S. Chandrasekhar that they are full of Mathematical equations. It is very difficult for an ordinary reader to understand them very easily. One is awed by the depth of his physical acumen the range of his mathematical vision and the sweep of his astronomical knowledge. He was a confluence of Mathematician, Physicist and Astronomer in himself.

Highest number of equations per paper were 127.4 in the domain D, 108.3 in the domain B, and 107.4 in the domain H.

**Table 8. Domainwise Bibliographical Characteristics per Publication of S. Chandrasekhar**

| Domain | No. of equations | No. of figures | No. of tables | Self citations | Citations to others | Synchronous self citation rate |
|---|---|---|---|---|---|---|
| A ( N = 47 ) | 46.5 | 1.6 | 2.1 | 0.8 | 6.3 | 11.14 |
| B ( N = 37 ) | 108.3 | 2.0 | 1.5 | 1.9 | 7.6 | 20.23 |
| C ( N = 9 ) | 84.8 | 1.0 | 1.0 | 2.0 | 8.8 | 18.46 |
| D ( N = 5 ) | 127.4 | 0.8 | 0.4 | 0.8 | 5.8 | 12.12 |
| E ( N = 26 ) | 57.7 | 1.7 | 1.9 | 3.3 | 4.9 | 40.57 |
| F ( N = 23 ) | 88.8 | 1.3 | 2.7 | 4.9 | 4.8 | 50.00 |
| G ( N = 63 ) | 61.7 | 0.6 | 0.5 | 3.9 | 5.1 | 43.02 |
| H ( N = 39 ) | 107.4 | 3.0 | 0.3 | 4.0 | 7.1 | 36.32 |

**Table 9. No. of Pages per Publication of S. Chandrasekhar**

| Domain | No. of pages |
|---|---|
| A ( N = 74 ) | 17.8 |
| B ( N = 55 ) | 14.9 |
| C ( N = 55 ) | 19.6 |
| D ( N = 5 ) | 20.4 |
| E ( N = 35 ) | 10.5 |
| F ( N = 27 ) | 14.3 |
| G ( N = 62 ) | 12.6 |
| H ( N = 42 ) | 19.4 |

Numbers of figures per paper were three in the domain H and two in the domain B.

Number of tables per paper were 2.7 in the domain F, 2.1 in the domain A, 1.9 in the domain E, and 1.5 in the domain B.

Self citations per paper were 4.9 in the domain F, 4.0 in the domain H, 3.9 in the domain G and 2.0 in the domain C.

Citations to other authors per paper were 8.8 in the domain C, 7.6 in the domain B, 7.1 in the domain H, and 6.3 in the domain A.

Synchronous self citation rate for the domains were A (11.14), B (20.23), C (18.46), D (12.12), E (40.57), F (50.00), G (43.02), and H (36.32). Mean Synchronous self citation rate was 24.44 whereas mean synchronous self citation rates were for C. V. Raman (15.05) [29] and for K. S. Krishnan (13.82) [33]. This has sociological implications indicating that S. Chandrasekhar was a highly productive and key figure in his research speciality [47].

Number of pages per publication of S. Chandrasekhar are provided in Table 9.

India inspite of its limitations has produced so many illustrious scientists like H. J. Bhabha, J. C. Bose, C. V. Raman, S. Ramanujan, M. N. Saha and can produce so many scientists of high calibre provided it provides congenial scientific climate for scientists to work.

Chandrasekhar admits : he sometimes wonders how his career would have unfolded had he remained in India. Like Raman, his uncle, he might have presided over his own institute, but he then would have become enmeshed in the orcane politics of India's scientific establishment [9].

## 5. Conclusion

S. Chandrasekhar had contributed 380 papers during the period under study to various domains : *Stellar structure and Stellar atmospheres* (77); *Radiative transfer and negative ion of hydrogen* (55); *Stochastic, Statistical hydrodynamic problems in physics and astronomy* (55); *Plasma physics* (5); *Hydromagnetic and hydrodynamic stability* (36); *Tensor - Virial theorem* (28); *Relativistic Astrophysics* (63); *Mathematical theory of Black holes and Colloiding Waves* (43); and *General* (18).

He had 267 single authorship papers, 105 two authorship papers, and eight three authorship papers to his credit.

His 47 collaborators have contributed 421 authorships and domainwise collaborative authorships were A (28), B (46), C (32), D (9), E (13), F (28), G (40), H (36), and I (2).

He has published 139 papers in Astrophysical Journal, 59 papers in *Proceedings of Royal Society-A*, 31 papers in *Monthly Notices of Royal Astronomical Society*, 14 papers in *Proceedings of the National Academy of Sciences*, and 10 papers in *Observatory*.

High frequency keywords in the title of his papers were : Stability (39); General relativity (35); Radiative equilibrium (30); Stellar atmosphere (30); Equilibrium (30); Magnetic fields (17); Stars (17).

Mean bibliographic characteristics ranged : Equations (47-127); Figures (1-3); Tables (1-3); Self Citations (1-5); Citations to others (5-9); Synchronous Self Citation rate (11-50); Pages (11-20).

Considering all above bibliometric indicators, he represented excellence in his performance and had set up very high standards for his followers to surpass it. His work can be considered as performance of a Role Model Scientist to be emulated by present and future generations.

## References


1. Zuckerman, H. (1967). The sociology of Nobel prizes. *Scientific American*, **217 (5),** 25-33.
2. Zuckerman, H. (1967), Nobel laureates in science : Patterns of productivity, Collaboration and authorship. *American Sociological Review.* **32**, 391-403.
3. Zuckerman, H. (1977). *Scientific elite : Nobel laureates in the United States.* Free Press : New York.
4. Singh, J. (1966). *Some eminent Indian Scientists.* Publication Division, Ministry of Information and Broadcasting : Delhi, 53-62.
5. Salpeter, E. E. (1983). The 1983 Nobel Prize in Physics. *Science.* **222 (4626),** 883-885.
6. Wali, K. C. (1991). *Chandra : a biography of S. Chandrasekhar.* University of Chicago Press : Chicago.
7. Vemlataraman, G. (1992). *Chandrasekhar and his limit.* Universities Press, Hyderabad.
8. Chadrasekhar, S. (1993). Subramanyan Chandrasekhar (in Nobel lectures in physics 1981-1990, edited by T. Frangsmyr and G. Ekspong. *World Scientific* : Singapore ) 133-164.
9. Horgan, J. (1994). Profile : Subrahmanyan Chandrasekhar. *Scientific American.* March, 16-17.
10. Gupta, D. K. (1983). Chandrasekhar : winner of the 1983 Nobel Prize for Physics : a citation analysis study of his works. *Ann. Lib. Sci. Doc.* **30 (3-4)**, 177-184.
11. Kalyane, V. L. (1995). Role model scientist (in Wither Indian Science : third National Convention on "What is wrong with Indian Science", *Sovenir of Indian Science Writers' Association*, **Feb. 18-19**, New Delhi) 31-35.
12. Shockley, W. (1975). On the statistics of individual variations of productivity in research laboratories. *Proceedings of the IRE.* **March**, 279-290.
13. Gupta, D. K. (1978). Plate tectonics : a case study of transmission of ideas. *Ann. Lib. Sci. Doc.* **25 (1-4),** 86-92.
14. Ruff, I. (1979). Citation analysis of a scientific career : a case study. *Social Studies of Science.* **9**, 81-90.
15. Cawkell, T. and E. Garfield. (1980). Assessing Einstein's impact on today's science by citation analysis (in Einstein : the first hundred years, edited by M. Goldsmith, A. Mackay and J. Woudhuysen. Pergamon Press : Oxford). 31-40.
16. Sinha, S. C. and I. M. S. Bhatnagar. (1980). The information profile of a plant pathologist : a bibliometric study. *Ann. Lib. Sci. Doc.* **21(1-4),** 106-113.
17. Fox, M. F. (1983). Publication productivity among scientists : a critical review. *Social Studies of Science. 13*, 285-305.
18. Gupta, D. K. (1983). Citation analysis : a case study of a most cited author and his most cited article on sea floor spreading. *IASLIC Bull.* **28 (1),** 1-12.
19. Gupta, D. K. and S. Gupta. (1983). A citography on Lepichon's article on Sea-floor Spreading and Continental Drift : application of Bradford's law. *IASLIC Bull.* **28 (2),** 49-58.
20. Simonton, D. K. (1985). Quality, quantity and age : the careers of ten distinguished psychologists. *Int. J. Aging and Human Development.* **21**, 241.
21. Dieks, D. and W. J. Slooten. (1986). Historic papers in physics - the case study of Hugo Martin Tetrode, 1895-1931. *Czech. J. Phys. B.* **36**, 39-42.
22. Kragh, H. (1990). *Dirac bibliometrics* (in Dirac : a Scientific biography, Cambridge University Press : Cambridge) 293-301.
23. Todorov, R. and M. Winterhager. (1991). An overview of Mike Moravcsik's publication activity in physics. *Scientometrics.* **20 (1),** 163-172.
24. Garg, K. C. and M. M. S. Karki. (1992). Bibliometrics of research communication of INSA fellows. *J. Sci. Ind. Res.* **51**, 929-935.
25. Lancaster, F. W., M. J. Seter and L. Metzler. (1992). Ranganathan's influence examined bibliometrically. *Libri.* **42 (3),** 268-281.
26. Kalyane, V. L. and S. V. Kalyane. (1993). Scientometric portrait of Vinodini Reddy. *Journal of Information Sciences,* **4 (1),** 25-47.
27. Sinha, S. C. and M. F. Ullah. (1993). Citation profile of Dr. V. S. Ramachandran : a bibliometric analysis of his highly cited articles and books in the area of cement and concrete chemistry. *Ann. Lib. Sci. Doc.* **40 (1)**, 21-31.
28. Kademani, B. S., V. L. Kalyane and M. R. Balakrishnan. (1994). Scientometric portrait of P. K. Iyengar. *Lib. Sci. with a Slant to Documentation and Information Studies.* **31(4),** 155-176.
29. Kademani, B. S., V. L. Kalyane and A. B

Kademani. (1994). Scientometric portrait of Nobel Laureate Dr. C. V. Raman. *Indian Journal of Information, Library and Society.* 7 **(3-4)**, 215-249.

30. Kalyane, V. L. and S. V. Kalyane. (1994). Scientometric portrait of M. S. Swaminathan. *Library Science with a Slant to Documentation and Information Studies.* **31 (1),** 31-46.
31. Kalyane, V. L. and R. S. Devarai. (1994). *Informetrics on C. S. Venkata Ram* (in New Horizons in Library and Information Science : Dr. Velaga Venkatappaiah Festschrift, edited by C. P. Vasishth, L. S. Ramaiah, N. V. Jagga Rao and T. V. Prafulla Chandra. T. R. Publications, Madras) 475-478.
32. Kalyane, V. L. and B. S. Kademani. (1994). Scientometric portrait of U. R. Murty. *Timeless Fellowship.* **16**, 1-23.
33. Kademani, B. S., V. L. Kalyane and A. B. Kademani. (1996). Scientometric portrait of Sir K. S. Krishnan. *Indian Journal of Information, Library and Society.* **9 (1-2)** (In press).
34. Kademani, B. S. and V. L. Kalyane. (1995). *Outstandingly cited and most significant publications of R. Chidambaram* (To be published).
35. Kademani, B. S. and V. L. Kalyane. (1995). *Scientometric portrait of R. Chidambaram : a citation analysis* (In press).
36. Kalyane, V. L. and B. S. Kademani. (1995). *Scientometric portrait of R. Chidambaram : a publication productivity analysis.* (In press).
37. Kademani, B. S. and V. L. Kalyane. (1996). Bibliometric indicators for publication productivity analysis of an individual scientist. National seminar on progress in Bibliometric Indicators, 28-29 Feb. 1996, Dept. of Library & Information Science, Annamalai University. (In press).
38. Kalyane, V. L. and S. S. Munnolli. (1995). Scientometric portrait of T. S. West. *Scientometrics*, **33 (2),** 233-256.
39. Kalyane, V. L. and R. K. Samanta. (1995). Informetrics on K. Ramiah (in New Vistas in Library and Information Science : Papers in honour of Prof. G. V. S. L. Narasimha Raju, edited by A. A. N. Raju, L. S. Ramaiah, N. Laxman Rao, and T. V. Profulla Chandra, Vikas : New Delhi) 565-578.
40. Kalyane, V. L., M. B. Hanji, S. V. Kalyane. (1995). *Scientific School of a Botanist (in International Dr. P. N. Kaula Felicitation Festschrift)* (In press).
41. Munnolli, S. S. and V. L. Kalyane. (1995). Scientometric portrait of R. G. Rastogi. *ILA Bulletin.* **31 (3).** (In press).
42. Chandrasekhar, S. (1989-91). *Selected papers of S. Chandrasekhar.* **Vol. 1-6.** University of Chicago Press; Chicago.
43. Pravdic, N. and C. Ouic - Vukovic. (1986). Dual approach to multiple authorship in the study of Collaboration/Scientific output relationship. *Scientometrics.* **10 (2-3)**, 259-280.
44. Subramanyam, K. (1983). Bibliometric Studies of research collaboration : a review. *Journal of Information Science.* **6 (1)**, 33-38.
45. Vinkler, P. (1990). *Bibliometric analysis of publication activity of a scientific research institute* (in Informetrics 89/90, edited by L. Egghe and Rousseau. Elsevier Science publishers. B. V.) 309-334.
46. Sen, S. K. and S. K. Gan. (1990). Biobibliometrics: concept and application in the study of productivity of scientists. *Int. Forum Inf. and Doc.* **15 (3)**, 13-21.
47. Lawani, S. M. (1982). On the heterogeneity and classification of author self-citations. *J. Am. Soc. Inf. Sci.* **33**, 281-284.
48. Loehle, C. (1990). A guide to increased creativity in research - inspiration or perspiration. *Bioscience.* **40 (2)**, 123-129.
49. Kalyane, V. L. and S. V. Kalyane. (1991). Scientometric dimension of innovative communication productivity system. *Annals of Library Science and Documentation.* **38 (1)**, 8-29.
50. Van Heeringen, A. and P. A. Dijkwell. (1986). Mobility and productivity of academic research scientists. *Czech. J. Phys.* **B 6**, 58-61.
51. Merton, R. K. (1973). *The sociology of science.* University of Chicago, Chicago, P.56.
52. Sindermann, C. J. (1985). *The joy of Science : excellence and its rewards, plenum*, New York.
53. Long, J. S. and R. McGinnis. (1985). The effects of the mentor on the academic career. *Scientometrics.* **7 (3-6)**, 255-280.
54. Mahapatra, G. and R. Kaul. (1994). Bibliometric analysis of citation classics in life sciences. *Library Science with a Slant to Documentation and Information Studies.* **31(3),** 129-134.

# *Bibliometric Study of Gender Differences in Psychological Research During 1976-1985 : A Pilot Study*

**Kamlesh Goel**
Documentation Officer, NASSDOC, ICSSR
35 Ferozshah Road New Delhi - 110 001,India

**A bibliometric study of gender differences in psychological research during 1976-1985 in reported, based on the entries noticed in Bibliography of doctoral dissertations from the year 1976-77 to 1985-86. Number of male research scholars (64.87%) dominates to female scholars (35.13%). 35% females and 65% males have done research in the educational psychology. 32.7% females and 37.58% males have published abstracts of their papers in Indian Psychological Abstracts from 1976 to 1985.**

*Keywords* : ***Bibliometric study; Gender Differences; Psychology; Research.***

## 1. Introduction

Bibliometrics is quantitative study of various aspects of literature of a subject. Bibliometric studies are undertaken now-a-days enormously at national and international level. The present study is conducted in the area of gender differences in psychological research.

She sleeps last and rises first. She feeds her family first and eats last. She is the last to receive education and the first to *drop out* when she is required to baby-sit or mind the house. She is the last to get a proper job and the first to be laid off during *structural adjustment*. She is the last to be consulted for decisions about her environment and the first to be affected.

## 2. Gender Difference

Gender is here understood as the socially defined and constructed roles of men and women. Unlike sex, which is bibliogically determined, gender roles change from one place and culture to another and across time. For example, in nineteenth century in Europe it was considered that only men were suited for office work, whereas by the mid-twentieth century, secretarial work in offices was considered a female occupation. The twentieth century has been rapid changes in many places.

The Human Development report admits that *in no society do women fare as well as men*. The involvement of women in decision-making is absolutely essential. The HDR' 95 also suggests that 30% reservations be made for women in all decision making positions of a national level by all countries as part of affirmative action tobring women into the political system.

With the 73rd Amendment India has already taken the first step in this direction

## 3. Male Dominance

A unique feature of the improvement of women's movement in India dates back to the last decade of the eighteenth century. In 1795 the British government passed two regulations to abolish the practice of female infanticidy. It has been the support received from male leaders. Raja Ram Mohan Roy, Ishwar Chandra Vidyasagar, Swami Dayanand, Gandhiji and Jawaharlal Nehru fought for the rights of women. Gandhiji's vision of freedom and equality for women inspired the beginnings of a truely feminist movement in the early decades of the twentieth century.

Writing in Young India in 1918 Gandhiji said

*Women is the companion of man, gifted with equal mental capabilities, she has the right to participate in the minutest details of the activities of man, and she has the same right of freedom and liberty as he... By sheer force of vicious custom, even the most ignorant and worthless men have been enjoying a superiority over women which they do not deserve and ought not to have.*

**4. Gender Difference And Literacy**

The Indian government feels that the focus of development is to be on the girl child and young women so that the benefits of education and health programmes come as natural increase to them.

However, a lot has to be done to improve the status of Indian women and girls as the situation is not satisfactory. Girls enrolment in primary education had shot up from a little under 25 per cent to 92 per cent in 40 years. However, the gap in the literacy rate for women and men was still high 39 per cent and 62 per cent respectively[3].

**5. Social And Psychological Aspects Of Indian Women**

While the Indian policy recognizes equality of rights between men and women, society implicity accepts a sharp distinction in their roles and spheres of activity. True parity will be possible only when the implications of the institutional equality are accepted in people's minds.

Women suffer continually in the *gender difference*. In other words, what is all right for men is all wrong for women.

The attitude of men towards women's work is different. Men do not take it seriously. If it is house work it is not work at all. *What do you do with yourself all days*? They ask the mother of three or two children. If it is an additional job outside the house, it is also not taken seriously. If there is any sort of family crisis, the woman must give up her job. If the children are sick, she is to stay home. Although exceptions are always there but it is the general phenomena.

Decision-making for community and the exercise of political power is still regarded as an elusive male preserve : This is clear from the entirely male composition of the traditional Panchayats, either of villages or of caste group (including some muslim castes). Men may engage in manual works outside the home but such work inside the house is considered derogatory and is expected to be done by women. Cooking, tailoring can be taken up by men as a profession but inside the house, these are left to women.

If all the gender differences, the domination, the dependence and lack of understanding and honest discussion could be done away with, a women would become equal to a man. If both could share in all the bad as well as the good things in life, both men and women could have so much to gain.

This traditional concept is changing with girls taking up white-coolar jobs. Parental inhabitions are breaking down where girls have to earn sometimes to provide their dowry and marriage expenses and sometimes to support their parents and younger members of the family. Home-making is raised to fine art. The precise activities may depend upon the locality, educational level, and extent of modernization, but the real differentiation remains.

Women's participation in the economic activity is important from their personal advancement and their status in the society. It has been suggested that the women must enter into the labour force of the country on an equal footing with men and get integrated into the system. Engels was one of the first theorists, to stress women's integration in the economy. Marx, also had opined that women should play an equal part in the country's development andthis was taken as an important precondition for the advancement not only of the women but the country as well.

**6. Factors Responsible For Gender Differences**

Three types of factors have been held responsible by different writers for influencing the early development of females and males; Biological, environmental and cognitive. No agreement has been reached as to which has the greater influence, but all three factors appear to have at least some effect[4].

[3] Social Welfare 47(7) October 1995

[4] Indian Journal of Industrial Relations 31(1) July 1995

**7. Significance of Psychology**

The subject psychology is becoming more and more popular day by day. The number of students opting for the study of this subject is yearly increasing in almost all the institutions of our country.

Psychology is the legitimate child of its mother philosophy. However, with passage of time, its nature has undergone a change from sheer speculation to the scientific procedure.

Nature of psychology is quite scientific. Like science, it believes in cause and effect relationship, utilise observation, experimentation and other scientific method for the study. It possesses an universally accepted body of facts and believes in the modification and alterations in its principles through future researches and findings. In fact it is a developing behavioural science that is trying hard to become as much objective, exact and accurate as possible to be on a far with developed sciences. Therefore, it is termed as a developing positive science (not as a science) of behaviour.

The scope of psychology is too wide. It studies, describes and explains the behaviour of all the living organisms. Living organisms and their life activities are countless. Therefore, no limit can be imposed upon the scope of the subject psychology. It has many branches and fields of study.

**8. Scope**

Indian Universities are covered by Bibliography of Doctoral Dissertations. Therefore my coverage of research is India only.

**9. Source of Data**

Dissertations are extremely important to research scholars therefore, their bibliographic control is of utmost important. Many countries bring out bibliographies of dissertations limited to their countries. Association of Indian Universities also bring out its publication as Bibliography of Doctoral Dissertations. All Ph. D. dissertations awarded by Indian Universities during the year are arranged according to various subjects with full bibliographical details as : Author, title, year of registration and award, name of university and with full address of guide.

**10. Methodology**

All the 10 volumes of *Bibliography of Doctoral Dissertations 1976-1985* have been scanned to collect the necessary data. Data has been recorded on 5" x 3" cards with full bibliographical details as : Author, title of the dissertations, name of the university, year of registration and awarding of degree and name of the guide with full address.

Data has been collected from psychology; sociology and education disciplines mentioned in the *Bibliography*. Social psychology is listed under social sciences and Educational psychology is listed under education.

Sample has been selected randomly as dissertations during the period of 1976-1985 have been selected.

The data has been organized to prepare various tables like : year-wise, according to subdisciplines of psychology and Male - Vs.- Female also:

The data has been matched with Indian Psychological Abstracts 1976-1985 to know the level of research scholar.

**11. Justification Of The Study**

The study will help in identification of the following aspects :

(i) Growth of Indian Psychology literature as reflected by the growth of theses submitted to different universities

(i) Growth of Indian Psychology literature as reflected by the growth of theses submitted to different universities

(ii) Gender differences.

(iii) Area of emphasis.

(iv) Identify the leading Indian Scholars (Male - Vs.- Female)

(v) Identify the activities of scholars after their Ph.D.s

**12. Results And Discussions**

Table 1 indicates that there is a growth in the field of psychology during 1976-1985, as, only 66 Ph.D.s, were there during 1976-77 and the number increased to 177 during 1984-85 and it decreased to 134 during 1985-86. There were only 18 female research scholars who got Ph.D.s during 1976-77 while 68 female research scholars were there during 1984-85. However, the number again decreased to 47 in 1985-86. The table indicates continuous growth of Ph.D.s during the period. It also indicates about the increasing number of female is a vis male

research scholars during the time span.

**Table 1. Growth Of Research**

| S. No. | Year | F | % | M | % | Total |
|---|---|---|---|---|---|---|
| 1. | 1976-77 | 18 | 27 | 48 | 73 | 66 |
| 2. | 1977-78 | 27 | 31 | 61 | 69 | 88 |
| 3. | 1978-79 | 25 | 27 | 67 | 73 | 92 |
| 4. | 1979-80 | 28 | 31 | 62 | 69 | 90 |
| 5. | 1980-81 | 49 | 35 | 92 | 65 | 141 |
| 6. | 1981-82 | 61 | 39 | 94 | 61 | 155 |
| 7. | 1982-83 | 65 | 42.5 | 88 | 57.5 | 153 |
| 8. | 1983-84 | 56 | 33 | 112 | 67 | 168 |
| 9. | 1984-85 | 68 | 38 | 109 | 62 | 177 |
| 10. | 1985-86 | 47 | 35 | 87 | 65 | 134 |
| Total | | 444 | 35.13 | 820 | 64.87 | 1264 |

F — Female M — Male

**Table 2. No. of Ph. D.s in the Various Fields**

| Subsdisciplines of Psychology | F | % | M | % | Total |
|---|---|---|---|---|---|
| Psychology, General | 1 | 5 | 19 | 95 | 20 |
| Physiological Psy. | 8 | 21 | 30 | 79 | 38 |
| Intelligence, Intellectual and Conscious mental processes | 41 | 39 | 64 | 61 | 105 |
| Subconscious and alert states and processes (Depty. Psy.) | 0 | - | 7 | 100 | 07 |
| Differential and Genetic Psy. | 126 | 46 | 148 | 54 | 274 |
| Comparative Psy. | 2 | 20 | 8 | 80 | 10 |
| Abnormal and Clinical Psy. | 22 | 34 | 43 | 66 | 65 |
| Applied Psy. | 29 | 31.5 | 63 | 68.5 | 92 |
| *Other aspects* : | | | | | |
| Criminal Psy. | 0 | - | 3 | 100 | 03 |
| Women Psy. | 5 | 55.5 | 4 | 44.5 | 09 |
| Religious Group | 1 | 50 | 1 | 50 | 02 |
| Educational Psy. | 217 | 35 | 398 | 65 | 615 |
| Social Psy. | 11 | 37 | 19 | 63 | 30 |
| Total | 463 | 36.45 | 807 | 63.5 | 1270 |

Note : The difference between total of Table 1 and Table 2 is due to the interdisciplinary researches of some scholars, e.g. Dr V V Bharathi's research is considered under Differential and Genetic Psychology : child psychology as well as under Educational Psychology.

The subdisciplines of psychology have been divided according to Dewey Decimal Classification Scheme. The scheme divides the subject in 9 sections and these 9 sections are further divided into various subsections.

The scheme consists of educational psychology with education discipline (classification number as 370.15) and social psychology is with sociology discipline (classification No. as 301.1).

The table indicates that only one 5% female research scholar has done Ph. D. in general psychology while 95% (19) male research scholars

have done Ph. D. in the same field. However maximum Ph. D.s have been done in various sections and subsections of psychology.

Educational psychology is the richest section of psychology as there are 35% (217) female research scholars and 65% (398) male research scholars.

Differential and Genetic psychology is the second richest section of psychology with 46% (126) female research scholars and 54% (148) male research scholars.

Intelligence, intellectual and conscious mental processes is the third section of psychology with 39% (41) females and 61 (64) males.

Applied psychology is the fourth section of psychology with 31.5% (29) females and 68.5% (63) males.

Abnormal and clinical psychology is the fifth section with 34% (22) females and 66% (43) males and social psychology is the next in the list with 37% (11) females and 63% (19) males. After it comes physiological psychology with 21% (8) females and 19% (30) males.

There are new areas with less number of Ph. D.s e.g. comparative psychology with 20% (2) females 80% (8) male research scholars while women psychology is with 55.5% (5) females and 44.5% (4) males.

However, religious group is the only section of psychology where equal number of Ph. D.s have been done by both males and females 50% each.

There are fields like Depth-psychology and criminal psychology in which no Ph. D. has been done by female research scholars.

The study indicates that number of women is less as compared to male number in each and every section and subsection of psychology.

Table 3 indicates that only 48 female research scholars out of 147 have published the papers (126 in number) in various journals of psychology. The abstracts of the papers have been published in the Indian Psychological Abstracts from 1976 to 1985. However, out of 330 male research scholars 124 have published 398 abstracts in Indian Psychological Abstracts during the same period. It shows that only 32.7% females have published their papers in the same journals during the same period.

**Table 3. Distribution of Publications**

| Year | F | No. of Abstracts | M | No. of Abstracts |
|---|---|---|---|---|
| 1976-77 | 7 | 28 | 20 | 102 |
| 1977-78 | 11 | 30 | 23 | 93 |
| 1978-79 | 7 | 22 | 27 | 108 |
| 1979-80 | 13 | 26 | 29 | 18 |
| 1980-81 | 10 | 20 | 25 | 67 |
| Total | 48 | 126 | 124 | 398 |

## 13.Conclusion

The study reveals that number of Ph. D.s done by female research scholar is less as compared to male research scholars in the various fields of psychology. The number of research papers is also dominated by male research scholars. With passage of time the literacy rate among the females has been increasing. Now, women require not only awareness among themselves but also more cooperation and support from the menfolk to improve their lifestyle.

*An earlier version was presented at the UGC National Seminar on Progress in Bibliometric Indicators, 28-29 February, 1996 at Annamalai University, Annamalai, India*

## References


1. Anand, J. K. (1980-1989). Journal of Documentation, : A analytical study. *Library Herald*, **30(2-4)**, July 1991 - January 1992, p. 181-211.
2. Arora, Jagdish and Sharan Pal Kaur. (1994). Bibliometric analysis of core journals on immunology : A study based on the Annual Review of Immunology, *Annals of Library Science and Documentation* **41(3)**, p. 81-94.
3. Bailey, J. M. and others. (1994). Effects of gender and sexual orientation on evolutionarily relevant aspects of human matching psychology. *Journal of Perso. Soc. Psychol.*, June, **66**. p. 91-92.
4. Broverman, I. K. and Others. (1972). Sex role stereotype, a current appraisal. *Journal of Social Issues*, **28(2)**, p. 59-78.
5. Buddhapriya, Sanghamitra and Preetam Khandelwar. (1995). Sex role stereotypes and women managers : A reality. *Indian Journal of Industrial Relations*, July, **31(1)**, p. 72-89.
6. Garg, K. C. and Dutt, B. (1992). Bibliometrics of agricultural research in India. Quarterly Bulletin of the International Association of Agricultural Information Specialits, 37(3), p. 133-9, 11 tabs, ref.
7. Goel, Kamlesh and K. C. Karg. (1993). Social science research in India : A bibliometric study. *Collection Managemen*, **17(4)**, p. 95-104.
8. Jacobson, Linda A. (1992). Physical appearance and gender; soco-biological and socio-cultural perspectives. State University of New York, p. 225-95.
9. Kolbert, Elizabeth. (1985). Scientific ideas : Women's - Vs.- Man's. The New York Times, 17 October, p. C1:1:
10. Kotte, Dieter. (1992). Gender differences in science achievement in 10 countries; 1970-71 to 1983-84. Long P, p. 259-70.
11. Mangal, S. K. (1990). Psychology : An introduction to human behaviour. New Delhi, Sterlin, p. 221.
12. Sangam, S. L. (1989). Information was pattern of researches in the field of psychology : A citation study. *IASLIC Bulletin*, **34(2)**, p. 55-63.
13. Sen, Subir K. and Sunil Kumar Chatterjee. (1990). An introduction to researches in bibliometrics. IASLIC Bulletin, 35(3), September, p. 108-17.
14. Sengupta, I. N. and S. Nagarjan. (1988). Scientific communication bibliometrics/informatics. New Delhi, Batra Book Service, v.p.
15. Singh, M. K. (1987). Participaion of women in economic development : Trends and issues. Indian *Journal of Industrial Relations*, 25(1), July, p. 93.
16. Sutton, E. D and others. (1995). Bibliographic instruction in psychology : *A review of the literature. Reference Services Review* (US), **23(3)**. p. 13-22.
17. Tyagi, K. G. (1995). Information sources in social sciences. IN Raju, A. A. N. New vistas in library and information science. New Delhi, Vikas, p. 29-41.

# *Scientometrics/Bibliometrics in India : An Overview of Studies During 1970-1994**

**Aparna Basu and K C Garg****
National Institute of Science Technology & Development Studies
Dr K S Krishnan Marg, New Delhi 110 012, India.

**Scientometrics is an area with considerable potential for use in quantifying the growth of scientific disciplines and estimation of impact and linkages. India's output of 423 papers presented at conferences or published in Indian and foreign journals during 1970-1994 indicates that India's share, which was greater than one fourth of the world output of papers from 1980-90, has declined slightly, by about 5% since 1990. Due to historical reasons, scientometric/bibliometric studies performed in the first two decades related tolibrary and information science. Many of these were case studies that have appeared in Indian journals of library and information science. Emphasis has been on library related issues such as the scattering of literature of specific disciplines in journals (or ranking of journals) and the extent of citation of journals in reference lists of articles. Since 1990, the situation appears to be changing, with a number of theoretical studies, the emerging use of bibliometrics for decision making, and national and policy issues being reflected more oftern. The application of new mathematical and statistical techniques and the use of computerized databases has been initiated, although a comprehensive mapping of national scientific output has yet to emerge. There appears to be a need for a coordinated program on scientometrics/ bibliomtrics studies to be performed in India in order to be of use to decision making at the national level.**

## 1. Nature And Scope Of Bibliometric Studies

The term bibliometrics was first used by Pritchard [1] in 1969. At that time it referred to certain quantitative inferences that could be made from the record of published scholarly papers. The area that has emerged from these early studies in the 25 years since the term was first coined, has the characteristics of both basic and applied research. For example, the quantitative relationships that were noticed in the early stages followed from empirical work and related to patterns in the productivity of journals, productivity of individual scientists, and word frequencies in the literature. These relationships that go by the names of Bradford [2], Lotka [3] and Zipf [4] all collectively come under what is now generically referred to as the 80-20 Rule. In other words, what was observed is that productivity patterns are usually highly skewed such that approximately 80% of the published papers in a field are located is 20% of the relevant journal set, or, 80% papers are authored by 20% of all authors, and that some words are much more frequently used than others. The empirical relationship of Bradford found immediate use in the economics of library management, such as shelving, access, issuing of books, etc. since the first practitioners of bibliometrics were those in the profession of Library Science. This initial historical bias remains

*Presented at the National Seminar on Progress in Bibliometric Indicatos, Department of Library and Information Science, Annamalai University, Annamalainagar-68002, Tamilnadu, India, Feb. 28-29, 1996.

**Author for correspondence.

to this day. Subsequently, these patterns were theoretically analyzed in terms of mathematical models to identify and explain the complex underlying processes leading to the specific behaviour [5].

In early period, bibliometric studies were sporadic and often conducted by individuals who wished to project the state of a particular discipline or group of performers. The first results of bibliometric studies of a particular discipline were published either in journals belonging to the discipline in question, or in journals of library science. Thus, a cohesive group of practitioners did not emerge with a common identity and their own journals and discussion forums. With the establishment of the journal Scientometrics from Hungary in 1978, and the initiation of a regular biennial international conference since 1987, the situation has been somewhat rectified, and a cohesive research community may be said to have emerged.

Within the last decade, the nature and scope of bibliometric studies has evolved considerably. A definite direction to the discipline was given by the seminal work of de Solla Price [6] in 1963, which may be considered as the precursor of what are called scientometric studies today. The possibility offered by the record of publications to formulate indicators that could give measures of performance and activity in the complex syste of knowledge production and exchange that constitutes scientific activity was first reviewed in 1978 [7]. In these quantitative studies of science, sophisticated data collection and data-handling techniques played a significant role. The possibility of developing *maps* emerged, using techniques such as co-citation or co-word analysis which exploit the similarities or differences between disciplines, country profiles, etc. Networks of authors and papers could be traced, in addition to tracking of emergent disciplines. The field developed under dual pressure (i) from the science management system to evolve objective quantitative measures that could be used in policy formulation, and (ii) by the creation of new possibilities offered by computerized data and computer based analytical tools [8]. Certain fundamental issues in the use of indicators, ethical aspects, problems of assessing science in general, and the use of bibliometric indicators in developing countries in particular have also been addressed [9]. Methods and techniques have been developed to study the cognitive processes in the development of scientific disciplines, and used in historical and sociological studies, for example, age-productivity relationships [10]. As the issues moved away from those specifically related to library issues, new names, i.e., Scientometrics and Informetrics were suggested for this area of studies. These terms are used at present in an equivalent and interchangeable sense. Various bibliometric aspects of this emerging area have been studied recently by Kabir [11].

## 2. Bibliometrics In India

In India there had been a trend in quantitative thinking in the study of information science due to the influence and early thrust provided by S. R. Ranganathan. The term librametry, a forerunner of bibliometrics, was coined by S. R. Ranganathan in a discussion during a conference conductedby ASLIB in 1948 [12]. A fairly long tradition of quantitative studies was initiated as early as 1958, with increased activity following a seminar organized by Documentation Research and Training Center (DRTC), of the Indian Statistical Institute in Bangalore in 1963. During the last two and half decades there has been a growing awareness of this field as witnessed by the introduction of courses in bibliometrics in schools of Library Science, and the establishment of the National Centre for Bibliometrics at Indian National Scientific Documentation Centre (INSDOC), New Delhi in 1988, and a number of conferences (Table 2). India hosted the third International Conference on Scientometrics Bibliometrics and Informetrics at Bangalore in 1991, in which there were 11 Indian contributions. At the fourth Conference at Berlin, there were 23 Indian delegates out of a total of 138 participants from 32 countries. This was the highest participation from any country barring U.S.A. and Germany, the host country. At the last conference in Chicago in 1995, there were 11 Indian participants out of a total of 120 delegates from 24 countries. In addition, reports on the scientific output of specific agencies, relative impact factor of journals used etc. are prepared by the concerned agencies such as Indian Council of Medical Research (ICMR) and Council of Scientific and Industrial Research (CSIR) [13].

**Table 2. Indian Contributions In National/International Conferences**

| Conference | Venue | Date | Papers |
|---|---|---|---|
| Fifth International Conference on Scientometrics Bibliometrics and Informetrics | Chicago, USA | 1995 | 18 |
| Fourth International Conference on Scientometrics Bibliometrics and Informetrics | Berlin, Germany | 1993 | 21 |
| Third International Conference on Scientometrics Bibliometrics and Informetrics | Bangalore, India | 1991 | 11 |
| Second International Conference on Scientometrics Bibliometrics and Informetrics | Ontario, Canada | 1989 | - |
| Proceedings of Seventh SIS Annual Convention | Calcutta, India | 1988 | 9 |
| First International Conference on Scientometrics Bibliometrics and Informetric | Brussels, Belgium | 1987 | 2 |
| Science & Technology Indicators for Development | New Delhi, India | 1987 | 10 |
| Bibliometric Studies and Current Information | Calcutta, India | 1985 | 11 |
| Primary Communication in Science and Technology | New Delhi, India | 1981 | 15 |
| Other 15 conferences | | | 22 |
| Total | | | 119 |

The above numbers and publication trends indicate active participation in scientometric and bibliometric research by the community of workers in this area. The number of publications has grown from 31 publications in 1970-74 to 120 in 1990-94, in the span of 20 years, indicating a doubling time about 10years. This is comparable to the growth rate of publications in this area in the world. With the number of researchers approaching critical mass for a viable field, the time may be right for examining new goals and priorities, as well as research norms essential for healthy development in the field. This has prompted the authors to take stock of the status of scientometric research performed in India during the period 1970-1994, in terms of growth, focus, and channels used for communication. The papers prior to 1970, i.e., in the 12 year period 1958-69, have not been included in the content analysis as their number is small compared to the papers published during 1970-94, and their focus and relevance to the present study fairly limited. Earlier reviews of bibliometric/scientometric studies in general, and with reference to India may be seen in Sengupta [14], Sen and Narendra Kumar [15], Sen and Chatterjee [16], and Ravichandra Rao and Neelameghan [17].

## 3. Scientometric Research in the period 1970-1994

We examine the following characteristics of bibliometric studies in India :

(i) growth rate of Indian literature in scientometrics/bibliometrics in comparison to international or world publications in this area;

(ii) disciplines or sub-disciplines of science and/or technology in which the scientometric/bibliometric studies have been conducted;

(iii) specific scientometric/bibliometric aspects studied, e.g. out-put analysis of individuals and nations, scattering of articles in journals, their ranking and citations etc.;

(iv) content, methodology, nature of contributions, in six categories, for every block of 5 years, i.e.,

- (a) Library issues
- (b) Theory/modelling
- (c) National Level or substantive (i.e., at the level of a discipline or sub-discipline)
- (d) Policy related
- (e) Case studies
- (f) General studies

(v) choice of journal for publication.

The aim of exercise is to obtain a general overview of the area and, in particular, to see if there has been a change in the pattern of research to orient itself from a mere application of the methodologies and techniques used in scientometric/bibliometric studies to obtaining either new theoretical knowledge, or using empirical information and analyses to address specific issues with regard to the conduct of science at say the national level, or studies that might indicate policy decisions.

Data for the period 1970-1990 was taken from a bibliography on Indian contributions in scientometrics/bibliometrics made by Gupta [18]. This is supplemented by data for current scientometric/bibliometric studies (1991-94) from *Library and Information Science Abstracts* (LISA). Some journals not covered by LISA such as the *Journal of Scientific & Industrial Research* (JSIR), are also included as a number of scientometric studies have been published in these journals. In addition, papers presented at the International Conference on Scientometrics, Bibliometrics and Informetrics were included. Studies dealing with social sciences were not included in the present study.

### *3.1 Growth Of Papers : (Table 1)*

The ditribution of papers for the period 1970-94, published in journals or presented at conferences (in blocks of 5 years) for India and world is given in Table 1. Indian output has increased from 31 in 1970-74 to 120 in 1990-94, a period of 20 years. On an average, the Indian output to world output which has hovered around 20% (with a marked low of 10% in 1975-79 and a rise to 27% in 1980-84) has fallen since 1990 to less than 20%. The fall is more marked for journal publications, though conference papers have increased.

**Table 1. World Output -Vs- Indian Output In Bibliometrics/Scientometrics**

| Year | World Output# | Indian Output Journal | Indian Output Conferences | Indian Output Other | Indian Output Total | % of World Output |
|---|---|---|---|---|---|---|
| 1970-74 | 147 | 28 | 2 | 1 | 31 | 21.08 |
| 1975-79 | 287 | 26 | 1 | 1 | 28 | 9.75 |
| 1980-84 | 412 | 81 | 26 | 3 | 110 | 26.69 |
| 1985-89 | 520 | 91 | 40 | 3 | 134 | 25.76 |
| 1990-94 | 603 | 70 | 50 | | 120 | 19.90 |
| Total | 1969 | 296 | 119 | 8 | 423 | 21.53 |

# Source : Kabir, S. H., ILA Bulletin, 28(3-4), 1993, 87-94

### *3.2 Subject Analysis Of Papers (Appendix)*

All the publications including conference papers and book-chapters were grouped into 9 scientific disciplines and an additional miscellaneous category. Identification of the sub-disciplines and the bibliometric aspect dealt with in the paper was made by using keywords within the title of the paper or from the abstracts. The results of the subject-wise analysis of papers dealing with different disciplines or sub-disciplines in Science and Technology (Appendix 1) indicates that the maximum number of scientometric/ bibliometric studies was undertaken in the field of Biological Sciences (51) followed by Space and Earth Sciences (39) and Agricultural Science & Food Sciences (38). The areas of Environmental Sciences, Energy and Engineering & Technology have received less attention as compared to other disciplines as reflected by the number of scientometric studies relating to these disciplines. The areas of Physical Sciences, Chemical Sciences and Medical Sciences have received equal attention as reflected by an almost equal number of papers in these areas. In all disciplines except Chemical

Sciences, sub disciplines have received more attention than the entire discipline, whereas in the Chemical Sciences, chemistry as a broad discipline has received as much attention as the micro disciplines.

The bibliometric aspects studied in the 400 odd publications were than divided into 9 sub-categories:

country-wise analysis of output,
analysis of disciplines,
scattering of articles in journals or journal ranking,
choice of journals for publications,
citation of journals,
author productivity
collaboration patterns, and,
growth of scientific publications.

The data (Appendix 2) indicates a predominance of scattering or ranking studies and citation analyses. This reflects the bias of library schools where the majority of these studies were conducted.

### *3.3 Choice Of Journals For Publication*

Of the 423 papers included in the study, 296 were published in journals, 119 were presented at 21 conferences held in India, and 52 papers presented at the five biennial International Conferences in Bibliometrics, Scientometrics and Informetrics, since its inception in 1989. The remaining 8 are book chapters Major conferences are listed in Table 2.

Of the papers published in journals 212 were published in 13 Indian journals, with the largest number of papers appearing in *Annals of Library Science and Documentation* published by INSDOC, followed by IASLIC Bulletin. The remaining 84 papers appeared in 17 journals published in countries outside India, mainly in *Scientometrics*, *International Library Review* and *Journals of Information Science* (Table 3). The list of other journals along with their impact factors, wherever available, are given in Table 3.

**Table 3. Indian Contributions To Scientometrics 1970-1994 (In Indian and Foreign Journals)**

| | Journal | Country of Publication | Impact Factor | Papers |
|---|---|---|---|---|
| 1. | Annals of Library Science and Documentation | India | Nil | 71 |
| 2. | IASLIC Bulletin | India | Nil | 41 |
| 3. | Scientometrics | Netherlands | 0.593 | 35 |
| 4. | Library Science with a Slant to Documentation | India | Nil | 18 |
| 5. | I. L. A. Bulletin | India | Nil | 14 |
| 6. | Journal of Scientific & Industrial Research | India | 0.237 | 13 |
| 7. | Journal of Library and Information Science | India | Nil | 12 |
| 8. | Library Herald | India | Nil | 12 |
| 9. | International Information, Communication & Education | India | Nil | 10 |
| 10. | Herald of Library Science | India | Nil | 9 |
| 11. | Journal of Information Science | UK | 0.224 | 8 |
| 12. | International Library Review | UK | 0.224 | 7 |
| 13. | Other 5 Indian Journals (with < 5 papers) | | | 12 |
| | Current Science | India | 0.205 | <5 |
| | Indian Journal of Agricultural Lib & Inforation Sci | India | Nil | " |
| | Lucknow Librarian | India | Nil | " |
| | Science Age | India | Nil | " |
| | Science Today | India | Nil | " |
| 14. | Other 14 foreign journals (with <5 papers) | | | 34 |

| | | | |
|---|---|---|---|
| Bulletin of IAALD | USA | Nil | <5 |
| Czech Journal of Physics | Czech | 0.330 | " |
| Environmental Science and Technology | USA | 2.603 | " |
| IEEE Transaction on Engineering Management | USA | 0.0386 | " |
| IEEE Transactions on Professional Communication | USA | Nil | " |
| Information Processing and Management | USA | 0.670 | " |
| International Forum for Information & Documentation | Netherlands | 0.246 | " |
| Journal of American Society of Information Science | USA | 1.074 | " |
| Journal of Documentation | UK | 1.033 | " |
| Nature | UK | 25.466 | " |
| R&D Management | UK | 0.043 | " |
| Special Libraries | UK | 0.535 | " |
| Unesco Bulletin for Libraries | USA | Nil | " |
| World Patent Information | UK | Nil | " |
| Total | | | 296 |
| 15. Book chapters | | | 8 |
| 16. Conferences Papers (Details given in Table 3) | | | 119 |
| Grand Total | | | 423 |

Majority of the papers were published in journals devoted to library and information science. Non library Indian and foreign journals, where articles published were *Journal of Scientific and Industrial Research, Current Science, Czech J. of Physics, Nature, IEEE Transactions on Prof. Communications, R&D Management, IEEE Transactions on Engineering Management, Environmental Science & Technology.* It appears that other than the 35 papers in Scientometrics, Indian publications in high impact journals publishing papers in this area is less than 10 in any journal. (the *Journal of Scientific and Industrial Research* is not of the same type as other journals and is therefore excluded).

**Table 4. Distribution of Papers in Indian and Foreign Journal**

| Years | Papers in Indian Journals | Papers in foreign | Total | Average annual publications | % of papers in Indian Journals |
|---|---|---|---|---|---|
| 1970-74 | 17 | 11 | 28 | 5.6 (3.4) | 60.7 |
| 1975-79 | 20 | 6 | 26 | 5.2 (4.0) | 76.9 |
| 1980-84 | 67 | 14 | 81 | 16.2 (13.4) | 82.7 |
| 1985-89 | 61 | 30 | 91 | 19.2 (12.2) | 67.0 |
| 1990-94 | 46 | 24 | 70 | 14.0 (9.2) | 65.7 |
| Total | 212 | 84 | 296 | 11.8 (8.5) | 71.6 |

### *3.4 Distribution Of Papers In Indian And Foreign Journals*

Table 4 indicates that the percentage of papers in bibliometrics published in Indian journals was 60%. This grew to 82% in the period 1980-84, declining in the following decade to 65%. Whereas this by itself may not be meaningful, when considered in juxtaposition with the results of the

content analysis in Section *3.5*, we find that this may be attributed to the initial proliferation of illustrative case studies, together with an increasing number of studies related to library issues and national level studies for which the more appropriate forum may have been the local journals. Some of the case studies published in the local journals would have resulted from student exercises or dissertations rather than more professional research studies. This is also partially corroborated by the pattern of publication of the more productive authors who may be identified as the continuing professionals in the field, who publish significantly more in international journals.(Table 6).

**Table 5. Distribution of Papers in Terms of Content and Type of Study**

| | 1970-74 | 1975-79 | 1980-84 | 1985-89 | 1990-94 | Total |
|---|---|---|---|---|---|---|
| 1. Library Issues | 6 | 8 | 23 | 12 | 10 | 59 |
| 2. Policy/Evaluation | - | - | 4 | 10 | 7 | 21 |
| 3. Theory/Modelling | 4 | 2 | 13 | 16 | 25 | 60 |
| 4. National Issues | 4 | 4 | 30 | 24 | 38 | 100 |
| 5. Case Studies | 13 | 8 | 32 | 65 | 34 | 152 |
| 6. General Studies | 4 | 6 | 8 | 7 | 6 | 31 |
| Total | 31 | 28 | 110 | 134 | 120 | 423 |

**Table 6. Publication of all Authors -Vs- Profile Authors in Indian and Foreign Journals**

| | Papers in Indian journals | Papers in foreign | Total | % papers in Indian journals |
|---|---|---|---|---|
| Profile authors | 96 | 57 | 153 | 62.7 |
| Others | 116 | 27 | 143 | 81.1 |
| Total | 212 | 84 | 296 | 71.6 |

A break up of the annual output of papers in terms of their general theme into the following categories :

Library Issues
Theory/modelling
National Level sudies
Policy related studies
Case studies
and General studies

indicates that the scene of scientometrics/ bibliometrics studies in India is changing (Table 5). For instance the number of papers on library issues which grew in the initial years has been declining since 1984, while thenumber of theoretical studies is steadily increasing. Similarly macro level studies on national issues hardly reflected in the initial decade, also appear to be increasing. This is similar to international trends with an increasing thrust towards use of bibliometrics as a tool for formulating national productivity indicators, and for decision making in areas other than library science. Even though the total output from India is dominated by the number of case studies, these appear to be declining since 1990, after an initial proliferation. The 10 year period prior to this (1980-89) is also the period when India's total output in this area was more than a quarter of the world output. We may infer that though our contribution is decreasing in quantitative terms, there has been a qualitative change with more studies being addressed to national or theoretical issues, rather than isolated case studies. The number of studies related to policy questions is still fairly small, but may be expected to increase if there is a greater demand for scientometric inputs

from decision making bodies. Other inferences from our examination of the themes are that sociological studies relating to research environment, productivity etc. using bibliometric inputs are still very few in number. The use of large scale computerized databases on publication output, and computer based multivariate analysis and sophisticated statistical concepts has only just been initiated. Although isolated national issues have been addressed, a comprehensive mapping of national level scientific output is still awaited.

## 4. Conclusions

Scientometric/bibliometric studies in India during 1970-1990 has generally been in the context of or within the realm of library and information science with the largest number of studies on scattering of literature in journals, or ranking and extent of citation of journals. Much of the work done in this field represents individual initiative, with the result that there is no perceptible thrust or direction. Very often, the data sets used for the studies are so meagre and analytical tools used are so nominal, that they may only be considered as illustrative exercises, rather than representative case studies, and the results cannot be used for any real life decision making. Comprehensive studies at the national level are few [19].

Since 1990, the nature of the studies has been changing with fewer studies addressed to library issues and theoretical studies, and national and policy questions being reflected more often. This shows a growing awareness in the research community of the potential use of scientometric studies, in line with world trends. Many countries such as Mexico, Hungary, Spain, Australia have expended considerable effort in mapping scientific profiles of their respective countries, along with assessment of performance. In India, there is a need to begin a similar coordinated program. While the trend toward the use of comprehensive computerized databases has been initiated, their use is not widespread. At the same time, the efforts in inducting new techniques and statistical methodologies into our studies have to be stepped up.

Most of the articles published during 1970-90 have appeared in journals related to library and information science. The restricted nature of the readership of library science journals would certainly have influenced the visibility of these studies in a wider scientific community. As a result, the role of scientometric studies for obtaining an overview of a discipline, and as an index of health, productivity and impact of a research area has not been brought out until quite recently. With the introduction of the journal JISSI (Journal of the International Society of Scientometrics and Informetrics) in 1995, published from India, there are now two journals (Scientometrics being the other) devoted exclusively to this area.

However, it is the responsibility of the research community to ensure that the potential of scientometrics is fully utilised by giving these studies adequate exposure in general fora, rather than letting them remain within the pages of journals as academic studies. There are several issues with regard to using bibliometric performance indicators that need to be debated by scientists along with those who generate these indicators. It is typical of an underdeveloped country either to reject wholesale or to accept uncritically the indicators, now being used more and more critical examination of the indicators, as well as a public debate on the extent of their reliability, usefulness and acceptability to the scientific community will not only help to generate more viable indicators but also integrate scientometric studies into the mainstream of scientific activities, and create a demand for scientometric studies from the scientific community itself.

## Appendix - 1

### Scientometric/Bibliometric Studies In India ( 1970-1994) : By Discipline

***Agricultural Sciences and Food Sciences (38)***

| | |
|---|---|
| Agricultural Extension | Identification of core journals using citation analysis. |
| Agricultural Sciences (9) | Identification of core periodicals; collaboration in research, institutional productivity and ranking of journals; publication trend of thesis articles in journals; choice of journals for publications; Agriculture Science Research in India; Recency of citations in Agriculture Sciences. |
| Food Science & Technology (8) | Application of Bradford's Law to CFTRI library; Indian contributions to international food science information service, significant journals in the field, their country of origin; selection of channels for publication by food scientists; analysis of CFTRI publications; most preferred journals in nutrition research and evaluation of Indian nutrition research using citations. |
| Genetics and Plant Breeding (1) | Citation analysis of Literature contributions in Indian J of Genetics and Plant Breeding. |
| Grape Research (1) | Identification of areas of research, sources where research results are published and trend of authorship. |
| Horticulture Science (1) | Comparative study of research contributions in leading Horticulture journals. |
| Marine Fish Research (2) | Analysis of CMFRI publications for identification of disciplines, nature of collaborations and journals where research results are published, Productivity in Marine Sciences and its evaluation. |
| Neem Research (1) | Neem research in India. |
| Plantation Crops (1) | Analysis of CPCRI publications for identification of disciplines, nature of collaborations and journals where research results are published. |
| Potato Research (1) | Identification of core Indian periodicals using different bibliometric techniques. |
| Poultry Science (1) | Identification of core periodicals by using citation analysis. |
| Rice Fish Culture (1) | Choice of literature used for publishing, growth of literature, trend of authorship, and ranking of journals. |
| Soil Sciences (3) | Ranking of periodicals, Analysis of papers in Journal of Indian Society for Soil Sciences for publication output, time lag in publication of articles, citation study, geographical distribution, authorship pattern and institutional distribution; Identification of most important journals from the view point of Indian soil scientists. |
| Sugarcane Breeding (2) | Communication activities of Sugarcane Breeding Research Inst.; collaboration trend in sugarcane research. |
| Veterinary Sciences (2) | Coverage of veterinary literature in Indian Science Abstracts; distributions of University of Agricultural Science publications in journals. |
| Weed Research (1) | Areas of research and growth of literature. |
| Wildlife (1) | Authorship trends in Indian wild life and fish literature. |

***Biological Science (51)***

| | |
|---|---|
| Biochemistry (11) | Ranking of biochemistry periodicals using citations; evaluation of Indian contributions in the field of Biochemistry and evaluation of Prof B K Bachhawat's work; growth of literature and research collaboration. |
| Biological Sciences (3) | Study of Indian contributions; citatin analysis of the journal collection in the library; use of information in Biological Sciences. |

| | |
|---|---|
| Biology Mathematical (1) | Study of the growth pattern, productivity of journals, geographical and language-wise distribution of literature. |
| Biomedical Literature (7) | Growth of literature and ranking of periodicals; comparative study of US & Japanese authors in Biomedical publications; study of speed - vs. - quality for biomedical journals. |
| Biophysical Literature (4) | Growth of literature and ranking of periodicals. |
| Biotechnology (2) | Indian output - vs. - world output, areas of research and contributing institutions; Biotechnology patents in India. |
| Botany (2) | Citation analysis of doctoral dissertations, self citations by Indian botanists. |
| Cytology-Experimental (1) | Contributing Institutions and ranking of periodicals. |
| Entomology (1) | Author productivity and areas of research for Nigerian entomologists. |
| Genetics (3) | Growth of literature, collaboration and scatter of literature. |
| Lactin Literature (3) | Growth of literature and pattern of collaboration. |
| Medicine And Aromatic Plants (3) | Areas of research, country-wise distribution of output, study of plant/genera, and scattering of literature; collaborative research patterns. |
| Microbiology (3) | Growth of literature and ranking of periodicals. |
| Mycotoxins (1) | Growth of literature and its spread into different disciplines. |
| Phytochemistry (1) | Growth of literature, journals where results have been published, contributing institutions and validity of Bradford's law. |
| Phytomorphology (2) | Contributing institutions, journals where research results are published and pattern of collaboration; frequently cited journals, geographical distribution of journals and authors, geographical distribution of journals and authors, obsolescence and scattering of literature. |
| Phytomorphology (2) | Contributing institutions, journals where research results are published and pattern of collaboration; frequently cited journals, geographical distribution of journals and authors, obsolescence and scattering of literature. |
| Phytopathology (4) | Study of collaboration, publishing institutions and countries, areas of research; most frequently cited journals and ranking of journals. |
| Plant Pathology (1) | Assessment of an individual's work. |
| Plant Physiology (1) | Growth of literature and most frequently cited periodicals. |
| Zoology (1) | Citation analysis of doctoral dissertations. |
| ***Chemical Sciences (29)*** | |
| Alcohol Fuels (3) | Publishing countries, forms of literature, correlation between productivity and breadth of research interest; mathematical model for scatter of literature. |
| Chemistry (15) | Growth and obsolescence of chemical literature; growth and development of chemistry periodicals; citation analysis of doctoral dissertations; Ranking of Chemistry periodicals; Subject-wise and country-wise growth of chemical patents; citation study of Indian chemical journals; most preferred journals in chemical & physical sciences; chemistry research in India, International cooperation in chemical siences, cros national assessment of literature in chemical sciences. |
| Desalination (1) | Validity of Bradford's law. |
| Electrochemistry (1) | Indian contributions to world literature and its impact on world literature. |
| Hetrocyclic Chemistry (1) | Growth of literature. |
| Hydrazones (1) | Growth of literature. |
| Materials Science (2) | Study of Indian contributions; validity of Bradford's law. |

| | |
|---|---|
| Organic Chemistry (3) | Assessment of organic chemistry research in India and ranking of periodicals based on citations. |
| Synthetic Chemistry (2) | Ranking of periodicals; assessment of synthetic chemistry research in Inda. |

***Engineering And Technology (22)***

| | |
|---|---|
| Armament Technology (1) | Validity of Bradford's law for Armament technology. |
| Building Science (2) | Validity of Bradford's law; analysis of publication data for Central Building Research Institute. |
| Civil Engineering (1) | Comparative study of two abstracting services. |
| Computer Science (6) | Validity of Lotka's law; pattern of collaboration; identification of core journals; research productivity and breadth of research interest, bibliometric analysis of IEEE Trans. on Computers. |
| Electronics & Electrical Engineering (2) | Citation pattern in IEEE Trans. on Microwave Theory and Techniques; Citation analysis of Proc. Inst. of Electrical Eng.. |
| Glass and Ceramics (3) | Research trends inglass technology based on papers, research reports and patents; ranking of periodicals. |
| Hydrology (1) | Contribution of Indian researchers toJ. of Hydrology. |
| Leathers Science (3) | Analysis of patents filed in leather and allied sciences; citation analysis; validity of Bradford's law. |
| Metallurgy (2) | Growth of literature; choice of journals for publications. |
| Off shore Technology (1) | Evaluation of Off shore Technology Conference Proceedings. |

***Medical Sciences (30)***

| | |
|---|---|
| Aids (2) | Growth of Aids literature, study of authorship pattern and identification of core journals. |
| Cancer Research (1) | Citation analysis of Cancer research literature. |
| Clinical Research (1) | Pattern of authorship. |
| Immunolgy (9) | Coverage of immunology journals in different data bases, high impact and highly cited core journals; authorship patterns; scientific productivity and most cited primary authors, citation analysis/networks in immunology; most cited articles in immunology. |
| In Vitro-Fertilization (1) | Growth of literature and authorship pattern. |
| Medicine (4) | Ranking of medicine journals; citation analysis of Ethiopian medical literature; medical research in India. |
| Neuroscience (4) | Growth of periodicals, pattern of distribution of literature, preference for citing literature, geographical distribution of cited periodicals, their ranking and evaluation. |
| Pharmacology (3) | Growth of literature and ranking of periodicals. |
| Physiology (3) | Ranking of periodicals. |
| Reproductive Endocrinology (1) | Indian contributions - vs.- world contributions and impact of Indian research on world research. |
| Surgery (1) | Information use by surgeons. |

***Physical Sciences And Mathematics (30)***

| | |
|---|---|
| Antennas (1) | Growth of literature, geographical and language-wise distribution of articles and patern of authorship. |

| | |
|---|---|
| Condensed Matter (1) | Citation links of journals. |
| Fibre Optics (1) | Publication output in fibre optics. |
| Fluid Mechanics (1) | Identification of most frequently cited journals. |
| High Energy Physics (1) | Identification of most preferred journals in physics using citation analysis. |
| Holography (1) | Bibliometric analysis of middle level countries. |
| Laser Research (2) | Indian contribution - vs.- world contribution, impact assessment of laser research performed in India by using citations and mathematical model for growth of literature; profile of laser research in middle level countries. |
| Mathematics (1) | Publishing trend of Indian mathematics in different countries and journals. |
| Mossbauer Effect (3) | Growth of literature, country-wise distribution of output and choice of journals for publication |
| Nuclear Science (1) | Identification of contributors and core journals. |
| Optical Engineering (1) | Citation pattern of literature. |
| Others (4) | Citation study of Indian Journal of Physics and Journal of Astronomy and Astrophysics, Citation Study of Prof S Chandrasekhar and evaluation of Scientific productivity of an Indian physics Laboratory. |
| Physics (9) | Choice of journals for publication, Physics research in developing countries, citation analysis of Indian physics, Indian physics articles in foreign journals, Trend of authroship, institutional productivity, identification of field of specialization, journals where the articles have been published, most frequently cited journals by Indian physicists, Growth of primary periodicals in India since independence; cross national assessment of physics literature. |
| Statistics (1) | Cross national assessment of national priorities. |
| Super Conductivity (1) | Publication and citation pattern of articles, their distribution according to language and geography; cross reference analysis of literature. |
| Supergravity (1) | Citation analysis. |
| Systems & Cybernetics (1) | Cross national assessment of national performance. |

***Space and Earth Sciences (39)***

| | |
|---|---|
| Aeronautics & Space Sciences (8) | Indian contributions in aeronautical sciences; pattern of collaboration, publishing and citing pattern of Indian space scientists; Use of ranking studies to adopt selective acquisition policy; Publication activity of National Aeronautical Laboratory (NAL) scientists and citation analysis of documents cited by them. |
| Earth Sciences (8) | Serial literature in the field of earth sciences, citation analysis of Indian Journal of Earth Sciences and Journal of Geological Society of India; title analysis of earth science periodical literature; identification of core journals and citation analysis of earth science literature; growth pattern and scientific productivity of authors. |
| Exploration geophysics (7) | Citation analysis, growth and obsolescence of literature; collaborative research trend; growth pattern of literature and scientific productivity of nations. |
| Geology (5) | Use of journals; pattern of authorship; scatter of Indian geological documents in Indian and foreign journals; validity of Bradford's law. |
| Geophysics (1) | Communication behaviour of geophysicist. |
| Geosciences (2) | Collaborative trend of research among nations; author productivity pattern in Nigeria. |
| Marine Geology (1) | Citation analysis. |

| | |
|---|---|
| Petroleum (3) | Growth of literature and citation analysis. |
| Plate tectonics and Sea Floor Spreading (5) | Transmission of ideas; identification of peers and specialists; citation analysis of single author. |

***Energy and Environmental Sciences (17)***

| | |
|---|---|
| Bioenergy (3) | Identification of areas of research, type of equipment developed and their nature of use; concentration, dispersion of literature and identification of core journals; growth of literature; nature and number of journals carrying information on water hyacinth and study of authorship pattern. |
| Environmental Sciences (7) | Selection of core journals; bibliometric analysis for identification of areas of research, collaboration and authorship pattern, institutional output, and scholarship of Indian work; Channels of communication used by Indian environmental scientists; contribution of NEERI Scientists; contributions in Journal of Environmental Biology, with regardd to nature, time lag of citing research papers, scattering of papers according to country and institutions; evaluation of Indian ecology Journals in relation to mainstream ecology journals; citation analysis of Ph D Thesis in environmental sciences and engineering used by NEERI Scientists. |
| Renewable Energy Sources (1) | Identification of areas of research, contributing institutions, authors and ranking of journals. |
| Solar Energy (6) | Indian contributions - vs.- world contributions and growth of literature; solar energy research in USSR, citation analysis of Geotekhnika, factor responsible for growth of solar energy literature; Solar energy research in India. Applicability of Bradford's law in relation to time for solar power research. |

***Miscellaneous (145)***

| | |
|---|---|
| Aspects of Science and Technology (6) | Evaluation of Indian science and Indian Scientific Journals using different techniques (36).<br>Research Planning and S&T indicators (14)<br>Collaborative research in Indian Science & Technology (7)<br>Evaluation of third world science (9). |
| Bibliometrics/Scientometrics (79) | Bibliometrics : Theory, Practice, Concepts and Applications (30); Bradford's law (19).<br>Science Citation Index : Structure, utility, implication and significance (16)<br>Citation Analysis : Application to Library Management (6)<br>Others (8). |

## Appendix - 2

## Scientometric/Bibliometric Studies In India (1970-1994) : By Nature Of Study

**1. Country-wise analysis of output**

Horticulture Sciences, Soil Sciences,

Mathematical ciology, biotechnology, medicinal/Aromatic plants, Phytomorphology, Phytopathology Chemical Sciences, alcohol fuels.

Immunology, neuroscience,

Physics, antennae, Mossbauer effect, Nuclear Science, Superconductivity.

Exploration geophysics, environmental sciences.

Solar energy, systems & cybernetics.

Mathematics & Statistics.

**2. Analysis of Indian Contributions**

Agricultural sciences, marine fish, Neem research, food science and technology,

Biological sciences, biochemistry, biotechnology, biotechnology patents.

Chemistry, organic chemistry, electrochemistry, material science, systhetic chemistry

Hydrology

Medical sciences, reproductive endocrinology,

Laser

Aeronautics & Space science, solar energy.

**3. Analysis of Institutional Contributions**

Agricultural sciences, soil science, veterinary science, food science and technology.

Biotechnology, experimental cytology, Phytochemistry, Phytomorphology, Phytopathology,

National Aeronautical Laboratory, environmental sciences,

NEERI, Renewable energy sources.

**4. Analysis of disciplines**

Grape research, marine fish, plantation crops, week research.

Medicinal/aromatic plants, entomology, Phytopathology

Chemistry patents,

Computer science

Physics

Earth sciences, environmental sciences, renewable energy sources, bio-energy.

**5. Scattering/Ranking Studies**

Agricultural sciences, Agricultue extension, potato research, poultry science, rice fish culture, soil science, food science and technology.

Biochemistry, biomedical, biophysics, experimental cytology, gnetics, microbiology, medicinal/ aromatic plants, Phytochemistry, Phytomorphology, Phytopathology.

Chemistry, organic chemistry, alcohol fuels, desalination, material science, synthetic chemistry.

Armament technology, building science, computer science, leather science, medicine, AIDS, neuroscience; pharmacology, physiology,

Nuclear science

Aeronautics/space science, earth sciences, geology, environmental sciences, renewable energy sources, solar energy.

**6. Choice Of Journals For Publication**

Agricultural sciences, grape research, marine fish, plantation crops, rice fishculture, sugar cane

breeding, food science and technology, nutrition research,
Phytochemistry, phytomorphology,
Chemical & physical science,
Metallurgy,
Physics, high energy physics, mossbauer effect, geophysics,
Environmental sciences.

**7. Citation Analysis**

Agriculture sciences, genetics and plant breeding, soil sciences, nutrition research.
Biological sciences, biochemistry, botany, phytomorphology, phytopathology, plant physiology, zoology,
Chemical journals.
Computer science, eletronics, leather science, off-shore technology
Medical literature, immunology, neuroscience, cancer research, reproductive endocrinology
Physics, condensed matter physics, fluid mechanics, laser, optical engineering, radio engineering, supergravity, super conductivity
Aeronautics/space science, earth sciences, exploration, geophysics, marine geology, petroleum, plate tectonics.

**8. Author Productivity/Collaboration Studies**

Agricultural sciences, grapes research, marine fisheries, plantation crops, rice fish culture, soil science, wild life, sugar cane breeding
Biochemistry, genetics, laciness, medicinal/aromatic plants, entomology, phytomorphology, Phytopathology
Chemical sciences
Computer sciences
AIDS, immunology, invitro fertilization, clinical research
Physics, antennae
Aeronautics/space science, exploration geophysics, geology, geoscience,
Environmental sciences, water hyacinth.

**9. Growth Studies**

Rice fish culture, weed research, biochemistry, mathematical biology, biomedical literature, biophysical literature
Mycotoxins, genetics, lactin literature, microbiology, phytochemistry, plant physiology,
Chemical science, chemical periodicals, chemistry patents, heterocyclic chemistry, hydrazones.
Metallurgy
AIDS, in vitro fertilization, neuroscience, pharmacology, physics periodicals, antennas, laser research.
Mossbauer effect
Exploration Geophysics, Petroleum
Bio-energy, solar energy.

## References


1. Pritchard, A. (1969). Statistical bibliography or bibliometrics? *Journal of Documentation*, **25**, 348-349.
2. Bradford, S.C. (1948). Documentation, Crosby Lockwood, London.
3. Lotka, A. S. (1926). The frequency distribution of scientific productivity, *Journal of Washington Academy of Science*, **12**, 317-323.
4. Zipf, G. K. (1949). *Human behaviour and the principle of least effort*, Hafner, New York.
5. Some models are reviewed in Basu, A. Hierarchical distributions and Bradford's Law, *Journal of American Society for Information Science*, (JASIS) **43**(7), 1992, 494-500.
6. De Solla Price. (1963). *Little Science, Big Science*, Columbia University Press, New York.
7. Elkana, Y., J. Lederman, R. K. Merton, A. Thackery, H. Zuckerman. (1978). Toward a mtric of Science : The Advent of Science Indicators, Wiley, New York.
8. Van Raan, A. F. J.(ed.). (1991). *Handbook of Quantitative Studies of Science and Technology*, North-Holland, Amsterdam.
9. Moravcsik, M. J. (1991). Some contextual problems of science indicators, in Handbook of quantitative studies of Science and Technology, North-Holland, Amsterdam, ed. A. F. J. Van Raan.
10. Stephan, P. E., S. G. Levin. (1991). Measures of scientific output and the age-productivity relationship, in Handbook of quantitative studies of Science and Technology, North-Holland, Amsterdam, ed., A. F. J. Van Raan.
11. Kabir, S. H. (1993). World Literature on Bibliometrics : Authorship and growth patterns, *ILA Bulletin, XXVIII,* **3-4**, 87.
12. Ranganathan, S. R. (1995). Librametry and its scope, JISSI : *The International Journal of Scientometrics and Informetrics* **1**, 15-21 (Reprinted from DRTC seminar 7 (1969) Paper DA, DRTC, Bangalore, India.
13. CSIR Research Output : Biblometric Analysis of Research Papers, 1987, 1988, 1994, Council of Scientific and Industrial Research, New Delhi.
14. Sengupta, I. N. (1985). Bibliometrics : A bird's eye view, *IASLIC Bulletin,* **30(4)**, 167-174.
15. Sen, B. K. and Narendra Kumar. (1986). Indian contributions in bibliometrics, 1958-1984 : *A review, Annals of Library Science and Documentation*, **33(3)**, 85-103.
16. Sen, S. K. and Chatterjee, S. K. (1990). An introduction to researches in bibliometrics Part I : Background and perspective, IASLIC Bulletin, 35(3), 105-117.
17. Ravichandra Rao, I. K. and Neelameghan, A. (1992). From Librametry to Informetrics : An overview and Ranganathan's contributions, *LIBRI*, **42(3)**, 242-257.
18. Gupta, B. K.(ed.). (1990). Handbook of Libraries, Archives and Information Centres in India : An annotated bibliography, Aditya Prakashan, New Delhi.
19. Garg, K. C. and Tutt, B. (1992). Bibliometrics of Indian Science as reflected through Science Citation Index, *Journal of Scientific and Industrial Research,* **51(4)**, 329-340.

# *Theoretical Issues in Citation Process : A Review*

**Subir K Sen**
Department of Library & Information Science, University of Calcutta
Asutosh Building, College Street Campus, Calcutta - 700 073, India*

**Referencing is an age-old process. For a long time there was no standard practice for this process. It was Eugene Garfield's ingenuity to invert the references (or bibliographic footnotes and endnotes) and produce the inverted files in the form of citation indexes as a marketable commodity. With this, referencing has become a ritual. The references are basically tokens of recognition to earlier related work. References are selected by authors through various types of motivations. The whole chain of activities from citer motivation to actual referencing and ultimately to the production of citation indexes is now known as citation process. Citation indexes are intensively applied for many purposes. There have been numerous empirical studies on many aspects of different components of the citation process. But there are very little theoretical studies in understanding the process. The present review recounts the evolution of referencing from a very ancient time both in the West and in the East. It defines citation indexing as a device of academic competition and distinguishes reference, citation, reference analysis, citation analysis, citation strategy, citation etiquette, citation behaviour, citation practice, citation cycle etc. Literature of a subject can be categorized either as accumulative or as cumulative, citational profile of a subject can be indicator of its cumulativeness and scientific status. The average size of references, to citations appear to remain constant over time as also the ratio of citation and reference. The review discusses editorial and peer influence on referencing, self citedness and uncitedness, Ortega hypothesis and elite hypothesis etc. It is shown that citational theories can be grouped as pre-event and post-event theories : those which attempt at understanding the background activities and those which are more formal theories modeling phenomenon of citation. The various theoretical approaches considered in this review are 'order out of chaos', 'implicit or normative theory', 'evolutionary theory', 'logical dimension of references', 'holographic and maximum speed principles', 'referencing typology and cites motivation' etc. Then a scheme for citing process complex and Vinkler's quasi-quantitative model for citation scenario have been presented. About twelve different theorems from Kochen, Sen, Egghe and Rousseau, Krauze and Hillinger have been discussed together with models of Glanzel (response time), Glanzel and Schubert (stochastic birth), Nakamoto (obsolescence), Gomperts(constant citation), Sharma et al (electron lattice scattering ),Van Raan(Beer's law analogue) have been briefly discussed on the formal approaches. Throughout the review critical appraisals of past work have been made and avenues and problems for future research have been indicated. An over all synthesis has been attempted on the basis of the idea of idea genes or *memes* and meme pools and a compromise between the Ortega and the elite hypotheses has been shown to exist.**

* Address for coresponding : S. K. Sen, 85 (old) Devinivas Road., Calcutta 700074, India Fax : +91 33 5512180

## Introduction

A reference is a token of recognition to an earlier work. But the act of referencing has become a fashionable ritual. References depend upon many subjective and environmental factors. References generate citations and citation indexes. Citation studies have been popular due to the availability of a readymade data base of citation indexes which allows multiple linkages among several useful bibliographic parameters. Many empirical observations are being produced on the basis of them.

All such observations may be manifestation of a profounder inherent process to be at play in all human socio-intellectual behaviour. Whether this process is typical of bibliographic phenomena or of socio-intellectual behaviour of man or of all human social activities, and whether this processs can be explained off by already known models of certain natural phenomena is a big question. This question can only be answered by future researches. As a first step towards that, we need to examine the pheonomenon *citation* more closely and pedagogically. Citation analyses and citation studies are mostly qualitative, phenomenological and observational. Whenever such studies have been quantitative, they have rarely been more than elementary or simplistic in the sense of using analytical tools. Indeed, we have not been able to go beyond empiricism in this area. We have no sound deductive tools for arriving at the results in relation to empirical applications of citation analyses. The present work attempts at finding out the statuses of theoretical approaches and clarifying conceptual ambiguities and pave the way for future research on theoretical understanding of citation phenomena and citation processes.

As may be seen, referencing practice is of ancient origin but utilizing references to produce a social tool is a twentieth century (American) innovation. Like any other such tools we can utilize citation indexes and citation analyses without caring for understanding the underlying process. The transitions from reference based science activities to citation based science activities and information culture need serious attention.

## Referencing or Citing in Ancient World

Acknowledging intellectual debt or mentioning the works of the predecessors is an ancient practice atleast as old as formal scholarship and writing of books.

There could be many reasons of doing this. In ancient India one main reason was to establish one's views or ideas as superior to others. The discourses followed a dialectical method or debate. Putting arguments and counter-arguments in such a discourse there had to be two parties. The proponent would first mention the opponents' views (called *Purbapaksha* or the first party) then we would refer, contradict the views of the first party and establish his own as superior or true interpretation or valid theory (*Uttarapaksha* or the second party). To do this the scholar had to mention the names of the scholars and their works in the purbapaksha as also the names and works of uttarapaksha and of those from whom he would draw supports and authority.

In ancient Indian tradition *Shabda* was considered one form of proof or logical support. *Shabda* may be translated as *testimony.* This means seeking support from some authority. If that be the case then the persons from whose works support was to be sought, had to be named with appropriate reference as to the context and content. The Indian logicians held the view that the testimony must also be personal i.e. based on the words of some trustworthy person, human or divine. "In respect of truth or validity there is no difference between the trustworthy assertions of an ordinary person, a saint, a prophet and the scriptures (e.g. the vedas) as revealed by God". But status of *Shabda* or testimony was not the same for all the sub-disciplines of Indian philosophy and learning. Indeed the "Charvaka", "Buddha" and "Vaisheshika" systems did not recognize testimony as a distinct proof (*pramana*) or source of knowledge. (Satish Chandra Chatterjee, 1978, p.317-319).

In other cases of scholarly presentation where the dialectical system was not adopted and specially, in case of digests or the books where existing knowledge was gathered and systematically reproduced it was a common practice to mention the previous authors and also to show their incompleteness and achievements. We may mention such examples as *`Artha Shastra'* of Kautilya, *`Siddhanta Siromoni'* of Bhaskaracharya, *Charaka Samhita', `Sushruta Samhita',* etc. The references were always within the text and by name of the

author or by the name of the work or by both. Sometimes exact references were also provided by mentioning the chapter and the *'Sloka'(the verse)*.

A study and analysis of referencing in ancient texts in India may prove interesting.

McInnis and Symes (1988) mention that in the Western world the references in the ancient texts were by the title of the work and not by the name of the author. But their observations about non Western world need revision. They write - "The Western concept of author entry came from the Greeks. In the non-Western world and also in the medieval world of Christian Europe, the title identified a book, not the author. ...Similar to art and architecture a book was considered to come from God and the author was only a conductor for God's word. Mirroring this belief bibliographies put a work's title before the author's name or even ignored it..."

Whatever the ancient practices had been and whatever had been the reasons, it is apparent that there existed some norms of bibliographic referencing or citing precedent works. But the methods, the recording or details and the normative influences were not simple. It would probably require a full length research in studying referencing or citing practices in the ancient world especially those dealing with secular scholarship.

We can agree that bibliographic referencing is an age-old affair. It is an in-built mechanism of human culture. The norms, the practices, and the component fields and sub-fields of references as records have evolved throughout the last two or three thousand years.

**Citation Indexing and Citation Indexes**

No one would probably have bothered for analyzing citation analyses, much less so for theoretical understanding of citation process unless there was a device or tool called citation index. Citation index as a social artifact was invention of Eugene Garfield (1955). The idea of inverting references and possibility of showing linkages of relevant information through citations(references) was around for some time. But Garfield could produce it as a marketable commodity and influence the sociology of science towards a revolution. Referencing is an age-old custom or norm of scholarship. Busy scientists have always scanned and implicitly used the list of bibliographic references appended as footnotes or endnotes for re-adjusting knowledge and future and further inquiry and also for judging the value of the content in the citing item.

The Shepard's Citations (Shepard's Index) is in existence from 1873, in which legal citations to a previous case or judgement are shown in a chain. Before 1950's other attempts at producing indexes by taking references as access points and showing the related citing items were made which were sporadic and less than an optimum maintainable (working) size (Egghe & Rousseau 1990).

The citation indexes conceived by Garfield (Garfield 1955) and produced from 1963 had been readily accepted. Its immediate success proves that the communities of working scientists and sociologists or historians of science (atleast in America) were in need of such a tool.

Publication of Citation Indexes (SCI®, SSCI®, A & HCI®) by the Institute for Scientific Information (ISI) has completely changed the practice of producing scientific literature. These indexes have given historians and sociologists of science a tool to 'observe' and 'analyze' the changes, growth, influences etc. of scientific enterprise or scientific culture. Citation indexing has also provided a number of 'indicators' for measuring and ranking productivity or quality or influence of researches, of documents, of institutions, of nations.

Garfield (1964) wrote that citation indexing is essentially an algorithmic process. In other words it is a mechanical device and here lies its potential merits and demerits. Margolis (1967) wrote that citation indexing has a self-organising nature. "The very existence of Science Citation Index® [and other two indexes] will almost certainly have various feed back influences on the writing and citing habits of future authors. This would, inturn, be rejected in the contents of the Index..."

Margolis' words have come true. Referencing is becoming more elaborate and standardized. But along with the progressive usefulness, that he envisaged, some associated hindrances have also developed.

By citation index we should primarily understand SCI®, SSCI® or A & HCI®. But any

index produced form a set of references showing the linking from cited items to citing items produced by anyone even in a form quite different from the model indexes of ISI should be called citation index. In fact we now have two kinds of indexes to scholarly literature. One is the group of Traditional Indexes (TI) or Traditional Indexing Services including abstracting services (TIS) and the other of Citation Indexes (CI) and Citation Indexing Services(CIS).

It should be clearly noted that without some kind of citation indexes of more than an optimum size, much of researches and developments in science of science, in bibliometrics and in science policy studies would be doomed.

**Citation Indexing as a Device of Academic Competition**

Long back in an American novel I first came across the explicit demonstration of the American ideal of Americanism : *Competition (and winning)*. Competition as an American way of life has been highlighted by more serious authors too (e.g. Toffler, 1970). It is obvious that where there is competition there must be wins and losses. People compete for achievements and rewards just not for fun. To give away prizes you require winners or selections. To select you need competition. To have competition you must have a device, an instrument, a game to generate and monitor the competition. If competitions are keen, rewards are covetable, prestigious and valued, then there may be competition even for participation; sanity and sanctity may at times be overlooked.

Citation index appeared as a means, a device, a tool for monitoring competition and selecting winners in the academic game field. Higher the citation score, better is the work, better is the scientist, better is the journal, and better is the institution. Thus has been the messages of this toll.

Citation indexing was top to bottom American. But at the same time international. It appeared where there was a vacuum in the American academic scene or in the competitive arena of research and scholarly publication. Its coverage was global - all the countries and all the publishing languages. But majority of source documents were American. This prominent American bias generated a citation winning cycle for American science. As there are sufficiently many American journals, many among them have better citations and higher impact factors through top seed ones. Simultaneously, journals from peripheral countries being very few in number show lowly impacts. Eventually it is proved that American journals are very good. So, increase their numbers in the data base and decrease the number of journals from other countries with low impact factors. This in turn would give more visibility and prestige to those journals already with high impact scores and they would attract more and better contributions which would again boost up citations. The Mathew effect has thus set in on this game.

**References and Citations**

In this discourse we shall distinguish between a reference and a citation. A citation is the citing item, a reference is the cited item. References are bibliographic footnotes or bibliographic endnotes. Citations are items which can be connected to an earlier item (a reference or a cited item) derived from an inverted file made of references as access points or sought terms and citations or citing items as linked terms.

This point of inversion is crucial. The bibliographic references appended to a scholarly item act as parents or ancestors to the citing item. In a simple listing of bibliographic items (as is done in a TIS) the references are simply over looked. In a Citation Index (CI) the citing items become the offsprings or progenies.

We also use referencing to mean citing as an act of selecting and putting references.

**Reference Analysis and Citation Analysis**

One can study the sets of references appended to articles or books or any kind of documents and derive a number of results in respect of the topic or the document or the authors whatever be the aim of the study. Sometimes such analyses are called citation analyses, which is a wrong notion.

Citation analyses can only be done through specially prepared meticulously from the references of a number of documents taken as a sample population.

In reference analyses we usually deal with diachronous items which are older than the publications from which they are taken. In reference

analysis, samples are to be drawn directly from the referencce lists of the documents. In case of citation analysis, samples are to be drawn from citation index.

Citation analyses provide more insights than reference analyses.

Studies of citation behaviour or habits relate directly to references. But studies of citational linkages look at the citation chains or citation cycles as autonomous.

**Citation Strategy and Citation Etiquette**

Similarly the terms citation strategy and citation etiquette should be distinguished. Citation etiquette is not just the citation practice, nor is the citation strategy equal to the citation behaviour. Citation etiquette is a part of overall system of social control among scientists. It is in fact following the normative principles of 'story book' (idealized) science. Citation strategy on the other hand is a decisional problem-solving process. The strategy differs from author to author, situation to situation. With this a citer decides what to cite and what not to.

The strategy is more subjective and situational, etiquette is more normative and depends on the subject discipline. There may therefore arise at times some conflict between strategy and etiquette.

**Citation Behaviour and Citation Practice**

Citation behaviour is a term which is sometimes used to mean referencing behaviour. Referencing behaviour is the behaviour of a citer which leads to an incidence of citing or reference. In another sense it is the behavioural pattern or set of behavioural patterns which may be discerned of the citers or a group of citers taken together. Citation practice on the other hand is the act of citing or what is discernible from the actual references. Practice of citation is the outcome of citation behaviour. This is the practice of citing. Referencing or citation behaviour is intimately related to strategy. The practice or pattern is more related to the etiquette and the norms.

**Data Elements in Citational Records**

In the citation process, citation or a reference is a bibliographic record; such a record contains only the essential fields or identifying tags. Each record can be considered as a datum. Each component (field) of the record can also be taken as datum. Citation can be viewed as a binary relation between two such records or data and also as establishing some sort of connection between the data elements (components or fields) of one record with those of another. At a closure look the elements or components of a citational record are the author (single authored article), authors or one of the authors (in case of multiple authors), the address of the document or the source, the location where the item is to be found in the source document, the time at which the item is published, with or without the addition of the title of the item. These data elements or components of a citational record act as parameters in the citation process. Another implicit parameter or data element can be heuristically labelled. That is the subject area in which the item may be classified. The extension and intention of the subject field can be defined arbitrarily depending on the context and purpose of analysis. Definition of a subject field is somewhat elastic in such analyses. As the address(es) and affiliation(s) of the author(s) are available from source documents or from the source indexes of commercially available citation indexes, the address or the institution or organisation or the country or city can be taken as a field or data element in citational studies and can be used to establish linkages with such fields (e.g., address) of different sets of data or to establish linkages through these parameters between cited and citing records.

A citational record can be represented as $D(t,p,a,s,l)$ where $D$ denotes the particular item or document and $t,p,a,s,l$ represent time of publication (usually year); publication medium as journal name with volume and issue; author(s); traditional subject field or topic and the institutional affiliation or geographical location respectively.

**Mathematical Status of the Citation Relation (Sen 1990)**

The citational process is a time-dependent, dynamic one and the relation citation is a binary relation (c). The relation can be treated as a functional as well (C).

For two documents $d_1$ and $d_2$, a statement (proposition) $d_2$ c $d_1$ would mean that document $d_2$ cites $d_1$ which implies $d_1$ $c^{-1}$ $d_2$, meaning $d_1$ is cited by $d_2$. This statement should be written more

explicitly as :

$$d_2(t_2) c d_1(t_1) \in d_1(t_1)\, c^{-1}\, d_2(t_2),\ t_1 < t_2$$

Here $t_1 < t_2 \in T$, T is a completely ordered projection on a positive real line (we tacitly avoid stating T as the time parameter explicitly)

We have, therefore, for the relation c,

i) $d_1$ c $d_2$ and $d_2$ c $d_3$ do not necessarily imply $d_1$ c $d_3$ : Transitivity does not hold.

ii) $d_1$ c $d_2$ does not imply $d_2$ c $d_1$ and vice-versa: Reciprocity does not hold.

iii) $d_j$ c $d_j$ cannot happen : Reflexity does not hold.

The transitive cases may occasionally occur. We say that c as a binary relation is non-transitive, non reciprocal and irreflexive.

When we consider citation for a single citing item $d_1$, however $d_1$ cites a set and not a single item.

Thus $C : d_1 \Rightarrow C(d_1) = \{d_j \mid d_1 \text{ c } d_j,\ j \in N\}$.

**Cumulation Versus Accumulation**

Science is said to be cumulative in nature. It means literature of science obsolesces rapidly. Older knowledge base is replaced by new knowledge. Not everything of this literature or knowledge becomes obsolete or replaced at a time. A part of the information is incorporated in the current body of public knowledge (Wilson, 1977).

That is also the current status of valid knowledge. The social order and culture norms require that the part which is cumulated should be specifically indicated. To do that, the part of the information acquired from other sources need to be indicated. The parts where other claims are refuted and debated are also to be indicated. There is a third part which support, vindicate and illustrate. All these three parts require referencing. As cumulation of science is a continuous process the references need not go back much in time.

But there are subjects which are accumulative in nature. Most of the humanities and soft social sciences like philosophy or history are accumulative. In philosophy or in religion almost nothing in the form of knowledge or information becomes back dated. Referencing criteria have to be more complex in such cases of accumulative literature. Age of references can be very old in these subjects. Or else, there may not be any articulated reference at all. As the authors do not feel the need of explicit citing scriptures and classical writings.

There has been many attempts to classify subjects in broad categories and characterise their literature by certain features. Some of these features are internal and are related to methodologies, content and presentation. Other features are extraneous which include bibliometric parameters. Among the bibliometric parameters almost all attentions have been directed to references and citations. Price (1986) divided subject fields as hard science, soft science, technology and non-scientific. He tried to show empirically that in case of hard sciences citation age or age of references is much less in proportion than in case of soft sciences and non-sciences. Literature characteristics of the technologies are rather different from the other three. He also showed that the number of references on an average is higher in the hard sciences than in others.

Apparently there should be a link between the cumulative character and hardness as defined by Price. M.K. Buckland (1972) used the term compactness for studying the hardness. There can be a gradation list of compactness as well. For quantifying compactness Buckland used two parameters. One is the proportional age of references used in the literature of the subject and another is a scattering coefficient in terms of references in most productive portion of journals or core journals by Bradfordian analysis. In Buckland's study physics is more compact than chemistry.but mathematics is less compact than either chemistry or physiology.

There have been many studies on the size of references (size distribution) and the age distribution of references. Yet much remains to be done. We do not have a coherent picture or a complete spectrum of the age scattering and the number (size distribution) of references in different types of articles (say, full length research articles, reviews, short notes) over different subject fields and of different types within the same subject field. The same comment is applicable for citational scattering. One should also want to relate the degrees of collaboration or the nature of multiple-authorship with these two. Nederhof (1988) has shown that if one distinguishes between the object documents or study documents and subject documents then the differences of age structures of references between hard

sciences and non-sciences are narrowed down to a great extent. By study or object documents we mean those documents which are treated as objects of research and study or commentary. By subject documents we mean those documents whose assistances are sought for studying the object documents or the topic of research. In natural science, one rarely deals with an object document.

It may thus be taken for granted that referencing practices do not have large scale variations over subject fields.

No topic or subject is however totally accumulative (100% accumulation) or totally cumulative (100% cumulation). Every topic has some accumulative and some cumulative components or elements. Cumulativeness is not a discrete parameter. We can observe continuous shades of cumulativeness from highly cumulative to highly accumulative. We should be able to understand quantitatively the relative proportion of cumulation and accumulation (of information) in a subject field, in a sub-field, among the types of documents (viz., monographs, journal articles, reviews, reports, etc.), typology of research characteristics (e.g. theoretical, experimental, conceptual, observational, methodological etc.). We should be able to differentiate between accumulative references and cumulative references.

One main task before the citation process theory is to meaningfully correlate and link-up cumulation - accumulation characteristics and trends with the divisions of traditional subject fields, media distribution or document typology of literature in a subject field, literature type or article type (review, research article, short-communication etc.), content typology. (methodological, theoreti-cal, experimental etc.), collaborativeness and citational characteristics such as total number of references per document or per literature unit, normalised referencing indexes or density ratio, immediacy index, Price index, recency scores, object document references versus supporting document references, concentration and distribution of references over the text, taxonomic types of references and citational scattering (we need to be very clear about the citational parameters and reference parameters; for example, immediacy index is a citational parameter whereas Price's index and recency scores are reference parameters). These characteristics and parameters may also be correlated with ranked distribution of journals, institutions, countries, regions, from core to periphery and from high metabolic and front line research topics to cold and archaic research types. But if one compares the cumulative references of one subject field with another and accumulative references of one with another separately, there would probably not be much difference. The context analyses of references and taxonomic studies of references have overlooked any such comparisons. There has not been any taxonomic classification of references based on cumulative and accumulative features.

**Citation Cycle (Price, 1980,1986)**

Citation cycle is the foremost manifestation of citation phenomenon. The citation cycle exhibits an inter-locking metabolic complex of informetric parameters in a comprehensive and integrated structure. This may be thought in parallelism to nitrogen cycle. This cycle delineates the citation process (atleast one part of the process) and the recurrent construction of citation indexes. The cycle begins with the selection of source journals and the items contained in them. The number of source items are then connected with the authors of those items. Thus there is an integral connection of source-media, citable source items, authors of those items. These components of a record or an item are then connected to the components or data elements of citations on one hand and to references on the other. A reference is a potential citable item. Those among the references which are cited have two way links, one with one or more ancestors and the other with one or more progenies. Whenever this happens the cited item is called source and the citing item an item. However this nomenclature is not strictly observed. The two way linkage of a cited reference may be repeated. Some of the citing items become cited items later on and the chain recurs. This generates the cycle.

Citation cycle is the autonomous manifestation of citation process. Following one or more citation cycles (minor cycles) one can follow the core flow of scientific progress on a topic or subject. Similarly, the whole chain of important scientific activities can be traced following the total cycle as a whole. This citation cycliation shows a manifestation of the so-called elite hypothesis (a term coined

by the present author to contrast the Ortega hypothesis).

**Constancy of the Size of Reference**

Until and unless a bibliographic item appears as a reference, no citation occurs. It is the primary condition of an item to become a reference that the existence of this item must be known to the citer. There many social, environmental, local and personal reasons for ignorance about the existence of an item. Even when the items are known to exist there are number of constraints and motivational factors (discussed later) for selection of certain items among them for the purpose of references. On the other hand, the full or detailed bibliographic references may not be known to the citer but some incomplete or vague idea about the item may be there. In such cases the information or the idea contained in the item may be implicitly used or even represented in the citing item but no reference can be possible. Yet, in another level, there may be some deliberate omissions of references. Still others may be omitted or over-looked due to the fact that the information were received via some secondary sources or intermediaries. There is then the question of the size of the list of references or the space available for giving the references. Although there is usually no dictation or specification by either the authorities of the publication media i.e. the journals (an exception is *Science* vide Kidd (1990)) or by any other authority in scientific or academic establishments, there is a sort of self control mechanism for almost all the producers of research papers imposing a soft limit upon the number of references to be added to a paper. It is very difficult to suggest any explanation of the nature of this controlling mechanism and the optimum number of references. Price's estimation has no strong logical basis except some empirical associations. The optimum number should vary from subject to subject, from country to country, from language to language and from institution to institution and person to person. Yet within the frame-work of a region, language, subject-field and some temporal context, an optimisation of the size of references can be discerned or observed. This situation has a parallelism which is called Synnott's Law (Synnott and Gruber, 1981). This law states that all group meetings in management context has an optimum running time which is approximately one and half hours. If such a meeting runs more than this optimum period, participants get bored, and if it is less, the participants stretch on talking nonsense. Again, Synnott has no explanation but only personal observations. In many other human activities or social behaviour we find that an inherent control mechanism is in silent operation giving an optimum size to the operational output, in terms of time, space or number. A theory of citation process should in principle produce an explanation of this control for optimisation of the number of references and also the most probable number in a certain context.

G.A. Miller (1956) in his classical paper titled 'The magical number seven plus or minus two : some limits on our capacity for processing information' has shown that human mind cannot deal with more than a limited number of concepts at a time. Has this any relation to reference size? This may have some relation with *meme pool* and *memetis* (q.v.).

This status of an optimum size of references to a citation gives rise to a competition among the potential items of reference to get a place in a reference list. This has led to the investigation into the referencing behaviour and citer motivation. It is however doubtful whether researches into citer motivation and referencing behaviour can supply a theory of citation process although these three aspects are intimately related. Citation process should be viewed on its own right. Very recently Wouters (1993) has argued in the same line. His argument is that citation and reference are semiotically different as two distinct signs. Most of the studies until very recently have intermingled these two signs and have used them synonymously. Because of these the citation debate has continued and the theories produced so far have either missed the point or have failed to achieve result.

Thus, we can distinguish among various types of influences in a research paper and also the various types of references. Hicks and Potter consider the explicit references which go on for inclusion in citation indexes as a very small slice of potential references. They say that there are three distinct layers or groups of items of potentional references.

The first group comprise those sources which are bit far or obliterated for explicit referencing. The second group is those which appear as explicit references and the third group is most recent informal influence which are too close to be acknowledged as explicit references. But some from this third group may appear in the text in the acknowledgement part.

MacRoberts and MacRoberts (1986) made a content analysis of some papers and showed how much is explicitly covered and how much is not covered. They have also enumerated the cases of reference miss-citation i.e. incorrect references and lack of references. Citations and references are not mere data sets, they represent events as well. In fact, a citational record is a record of an event; that event is of being cited or of citing. MacRoberts and MacRoberts (1990) argued that their findings called into question the practice of culling references from bibliographies and using these as data. A quotation of some length from their paper may highlight the position : the mere presence of a reference is not a marker of influence, nor is the absence of a reference evidence that it is un-influential. References are simple obvious historical leads and evidence of influence only when they have been demonstrated to be so... . Our study raises questions about the relationships among the events an investigator desires to study, the data he used to substitute for those events, and the conclusions that he draws from the data. In the present case, the events to be examined were `influence' or `communication' among scientists and the relative `quality' of scientific work". The data that were substituted for those events were numbers of references in bibliographies or in SCI® database. The conclusions were drawn from the data, not from the events. At no point was the `event — data — interpretation' continuum checked to determine whether or not the data actually reflected or could be substituted for - the events.

"After completing this paper [of ours], we analysed it. Not more than 20 percent of the influence is captured by the references. Also, the work that most influenced us is not mentioned because it is unknown to the intended audience and thus would not be persuasively valuable or symbolically meaningful. In fact, because it is now known, it might detract from the paper's persuasiveness"

With all such case studies the conclusion remains the same, that there is an optimum size allowable or observable in case of bibliographic references to research articles.

It is usually thought that the technological advances should have made it possible for a researcher to access more information sources and more easily than was possible otherwise. It is also argued that only awareness of existing sources but also collection of actual documents or information pieces has become easier and cheaper.

Moreover, following Price's hypothesis that references are products of all the information sources in store of humankind (or civilization) at a specific time all these factors would then create a steady steep growth in the number of references in all types of publications. The other factors that would contribute to such a growth are gradual increase in collaborating research and team research; better communication among researchers in national and international meetings; inter-disciplinary research etc.

But in reality, the growth has not been so, it is only nominal except for the influence that was generated by the publication of citation indexes, which influence has again stabilised after some time.

Lancaster and Li (1988/1989) put forward an explanation. They hypothesised a law of constant accessibility of information. They assert that as the literature on a topic grows, it becomes increasingly scattered as has been enunciated in Sengupta's Law (1973). Sengupta's Law states that 'during phases of rapid and vigorous growth of knowledge in a scientific discipline, articles of interest to that discipline appear in increasing number in periodicals distant from that field'. Yet even here, a multi-dimensional 'APUPA' principle works (Sen et.al. 1991). A-stands for alien, P-stands for penumbral and U-stands for umbral. A one dimensional or linear principle of grouping of documents in subject classification and for arrangement on shelves was enunciated by Ranganathan and he called it 'APUPA' (Ranganathan 1967 *Prologomena*). What Sen et.al. suggested was that the distribution of items among sources also follow the 'APUPA' principle. There is a core group of umbral sources which have most of the relevant items. Then there are sources which

may carry marginally important items and only some times relevant items. The alien group seldom has any sources containing a relevant item. The searcher almost always concentrate on the sources in the umbral group. It is only a chance event or a strong persuation through a peer advice or sought citational link that a searcher comes across a source of the penumbral group. Even then though both umbral and penumbral sources are used only when they are easily available.

It may be added here that the strongest citational link occurs among the items in the umbral group. Items in penumbral group almost always refer to the items in the umbral group rather than other items in the penumbral group.

The Bradfordian scattering and Sengupta's law, both can be exaplained by a multi-dimensional dynamic 'APUPA' principle. In the beginning when papers on a new topic starts coming up, the umbral and penumbral groups or zones are small. With increase or growth in publications both umbral and penumbral zones increase. But the size of the penumbral zone increases much more rapidly than in the umbral zone. It has been observed that nucleation and scattering of items on sources start at an early phase of growth of literature on a topic (unpublished observation by Sen). The nucleus or the umbral zone is apparent from the beginning. After some time it reaches a quasi-static size. The point is that a few specialised journals start publishing in the subject field if it becomes important, and/or some of the journals dealing with a broader subject area devote much more space on such a high metabolic or current interest area. Only the penumbral region goes on increasing.

Accessibility and availability of penumbral items become more and more difficult, even though better techniques are adopted and reading over-load, search-time and procuring time are ignored.

Martyn (1987) compared the results of two studies : the second one done two decades after the first one. They show enhanced awareness of the importance of information and greater use of information gathering methods but no corresponding increase in the use of information sources. The number of cases of late discovery or procuring of useful relevant information has increased.

This is not unusual or abnormal. A research has time bound programme and total information search time has not increased, rather decreased. Reading overload and preparing and pursuing for 'grants' and other facilities have also restrictive influence to information gathering.

Kidd (1990) has noted that the journal *Science* has already an editorial instruction on the number of reference for thematic articles, although this is not adhered to on many occasions. His contention is that a time may come when such restrictions on references size will be a general norm. Whatever may happen in the future, there is already a social, normative and physical restriction on the size of the reference list.

Law of constant accessibility is a fact. This leads to a law of optimum referencing size. But this remains an open question whether this size approximates to Price's law of average references (1976).

**Price's Law of Average Referencing**

Price's law of average referencing (Price 1976, 1980, 1986) is a proposition or hypothesis. There is no direct theoretical proof of it. But it is highly plausible that the law is atleast approximately correct. It has indirect supports in many empirical studies and from the laws of Krauze & Hillinger (q.v.), Gomperts (q.v.) and others.

The law states that the average number of references per research paper (in science) is a small positive constant less than one plus the natural logarithm of the size of the available archive of literature (that is the total number of citable items already available).

Price argued that the estimated number of items comprising the archive of science was approximately one million (upto early nineteen seventies). The natural logarithm of one million is about 14. Hence average number of references per research paper in science should be around 14. Price, however, revised the estimate later on and suggested the value of $16 \pm 6$ (that is between 10 and 22) for the average number of references. This definitely does not apply to review articles.

Price himself stated that he had no hard and fast answer as to why the norm of scholarship limits the number of references per paper from ca 10

to 22. But he also observed that the average number of references in all the subject fields are growing over time although not steadily.

**Editorial and Peer Influence on Referencing and Citation**

Peers, editors, colleagues, listeners in a conference or seminar influence references. It may be interesting to study the change that occur in the list of references to a paper between the first version (or draft) and the ultimate published one. Although these peers make influence, the ultimate act of referencing or including a particular item in the reference list is author's own decisional problem. I do not know of any serious research done in this area. Such changes in references influenced by the colleagues, referees, peers and editors would change the ultimate citation scenario also.

On the other hand, editorial activity may influence and modify citation pattern directly. This can only happen through publication of a certain type of articles or not. Sievert & Haughawout (1989) in a paper showed that with change of editors, the impact factor or immediacy index could change for an educational journal (*ESJ*). But their findings can not be generalised, their conclusions are only policy influences and not any direct influencing to referencing practices. But the point is that editorial policies can definitely influence the referencing pattern and citation process directly and indirectly. These influences are very difficult to be incorporated in a theory of citational phenomena.

The only point of relief is that on ultimate score such influences become randomised on actual size of references and citations.

**Self Citedness**

For delineating a theory of citation process one should be concerned with the amount of self citations. In case of using simple counting of citations to account for the influences or impact, self citations can distort the realities. For this reason some authors are of opinion that the number of citations should be replaced by the number of citers i.e. citing authors. In this case the accountable value for self citations would be one irrespective of the number of self citations. This may create some difficulties in tackling papers by many authors (more than two) and while assessing status of journals through citation counts.

A theory for citation process should then be able to show a relationship between the number of citers and the number of citations as also the self citations. We may probably apply Lotka's law or Bradford's law or both for making an estimate of citer-citation relationship.

Let us have a closer look at self-citation. Individual self-citation is possible only for an author who has atleast two publications or atleast two documents to his credit. In case there are more than one author of a document, only one author usually becomes the cited author. If we take a counting of cited and citing documents as well as cited authors and citers, we come across an interesting situation, a multi-authored citing document has a number of citers whereas in the case of a multi-authored cited document a single cited author (the first author) can be accounted for. It is difficult to accommodate multi-authored papers for proper author — document relationship. Thus, the context of self-citation becomes a little bit blurred while dealing with multi-authored items.

According to Lotka's law number of authors producing one paper only in an epoch, is about 61% of total authors or total documents. Within such an epoch then, most of the papers can not have self citation. Again, by application of Bradford's law it can be safely said that more than 50% of all the papers (either taking a random sample or taking whole of the document space) remain uncited. Compounding these two we find that only less than one-sixth of all the citable documents can have possibility of self-citation.

There are empirical evidences that multi-authored papers are cited more as also the papers which carry more bibliographic references. Rousseeuw (1992) has given a simple argument following Rousseau (1991) based on Bayesian formulae.

He assumes that 5% of all the potential relevant items are included in the list of references of a work. 95% of items would therefore go uncited in a particular publication.

He also assumes that the selection of this 5% of potential items depends on the knowledge of the items or the literature. An author is best knowledgable about his own work. Then his knowledge would be better about papers written by a friend or

a colleague or an author who is personally known than about papers by other unknown authors. The least possibility is of knowing and using a paper by a completely unknown author writing in an alien language. Rousseau arrives at values for inclusion in reference list for papers written by the author, by a colleague and by others as 1.0, 0.33 and 0.077 respectively. Therefore, chance of self-authored papers more often belong to two categories with a 100% or a 33% chance of being included in a reference list. The percentage of multi-authored papers vary from one subject field to another.

**Uncitedness**

It is again a question whether the issue of uncitedness should be taken into consideration of a theoretical approach to citation process. The answer is yes. The issues of uncitedness and reasons and motivations and typologies of various kinds of citedness or referencing are related to one author. In the formal theories of citation process, the citational scattering should take into account the uncited items, at least the proportion of uncited items (Zero item sources of Sen, 1989).

Sen estimated more than 50% uncitedness on theoretical grounds (Sen 1989, Sen et. al. 1992).

Empirical studies on uncitedness are very difficult to perform. In the citation index or in any bibliography prepared through references, one would get those items only which have been cited. For studying uncitedness one has to take a sample of items at any point of time in the past and then to find out appearance of each and every item through subsequent years through a citation index or in references of items published in a representative set of journals (or better still in all the journals in all the languages published anywhere). However, one may study the publications in a single year after a gap of a certain number of years from the year of publication of the sample set. It is apparent that it is improbable to study uncitedness for each and every item published on a global scale. The results of the uncitedness studies done so far vary widely.

A study by Koenig reports 50% uncitedness in the pharmaceutical literature. But in other studies they vary from two to seven percent. A study by Stern (1990) reports about 21% uncited articles. Seglen (1992) in his detailed examination has found a value of more than 50% of uncitedness.

**Two Sides of the Citation Phenomenon**

The phenomenon of citation has several facets. It is combination of atleast two distinct processes. Distinction of these processes are marked by the two distinct approches to this phenomenon. The first is the process of referencing or the process of citing which culminates in the event of citation. The second is the process of citation or the citation process which starts regenerating the citation cycle, which in its turn induces the recycling of the first process. Referencing as an event is the culmination of a socio-academic process. And at the same time starting point of all that go in the name of citation indexing and citation analyses. The first part is a product of a control mechanism of scholarship and the enterprise of research and publication. The second is a process which is largely autonomous. After the references are made, they generate citations. As soon as this happens, there is no control over the distributions or linkages or any other outcome or manifestation of citations except for the selection (mostly arbitrary) of sample population of references. This second process which starts with the event of referencing and, better still, with the preparation of inverted file of citations and goes on generating the next cycle should be formally called the citation process.

The literature on theoretical aspects of citation are clearly divided into two groups.

In the first group there are attempts in understanding the process or processes (backgrounds and motivations) leading to referencing. They include theories or analyses of referencing practices and referencing behaviour; taxonomies or analyses of citer motivations; sociological background of references, i.e., references as socially controlled activity, etc.

In the second group there are the formal theories : attempts for theorizing the citation process or the process which is evident in citation-reference-citation linkages. These formalizations are generally more mathematical and stochastical in approach. They treat the references or citations or components of them or sets of them as parameters which are then linked with each other and such other parameters as time, ratio of citations and ref-

erences, the number of citers etc.

We may call these two categories as pre-event citational process theories or theories of citing and the post-event citational process theories or theories of citation.

**Classification of Theoretical Approaches**

Till now most of the researches on citation are empirical in nature. Like most of the crafts, reference analysis (by preparing a database of references from any number of documents by anyone, with or without any explicit rationale for selection and collection of the sample population) and citation analysis (through commercial citation index databases) are crafts which do not have sound theoretical background. Indeed, Leydesdorff has rightly stated that, "Citation can be used empirically without the benefit of a theory of citation" (Leydesdorff 1987).

The few theoretical attempts to understand citation process can be categorized in different groups as have already been explained in the preceding section. These groups can again be subdivided into sub groups.

The first group (pre-event) theories may be subdivided as sociological theories and semiotic theories. Most of the sociological theories are critical to elite hypothesis and favours Ortega hypothesis, explicitly or implicitly. Most of the semiotic theories favour an elite hypothesis.

Most of these (especially the semiotic approaches) do not consider cognitive relations in citational process. But there are a class of theoretical approaches which try for content analysis and cognitive connections of cited documents and the citation.

Among the sociological theories the main divisions which can be found are (i) those dealing with motivations, (ii) those dealing with referencing typology and reference practices, (iii) those dealing with content analyses and cognitive connections.

The second category of post-event theories may be called formal theories.

The first important subgroup of these formal theories is the theories of citational scattering. Most of these theories have been produced for explaining journal scattering, productivity or bibliographic scattering as a whole. These scattering theories will not be reviewed.

The second subgroup may be called heuristic theories. These theoretical models try to enunciate theorems and prove them in a formal manner. Cognitive linkages are prima facie ignored in these approaches.

The third subgroup attempts at an overall stochastic formalism. There is almost no model in this subgroup. Let us call this group of theories as cumulational theories or citational cumulation process theories. In this approach, citations are taken as tokens or records of cognitive flow in a time series and the events occur as random variables in a Markov chain. Such theories may attempt at linking the number of citers with citations and accounting for various shades of citations with proportional impacts. But no such theory is in view except some attempts by Schubert and Glänzel (1986).

Artus (1993) distinguishes between methodological approach and theoretical approach. According to him, methodological studies pretend to collect objective data which can serve as valid indicators for the temporal structure of a scientific discourse. This is one way of correlating different sets of empirical data and is a case of empirical operationalization without having theoretical ingredients.

Such an approach is however effective and useful in many cases. It can connect one class of empirical data with another. It can be used to demonstrate through empirical illustrations potential validity of some hypothesis and premises. Methodological approaches can lead to some understanding of the processes or operations and may indicate possible lines of attack for a future theory. In many cases they are the only available models in absence of any other theoretical model.

It will be seen in the following that many of the theoretical approaches to citation theory are in fact methodological. Some other theories are simple explanations or speculations and few others are empirical rather than inductive.

**Citations : Order out of Chaos Theory**

Starr (1983) viewed scientific work as the representation of chaos in an orderly fashion. Cronin (1983) observes that if this interpretation is accepted then the citations play the role of a control mechanism in confering orderliness and acceptibility on published research.

These issues are linked with the debate of historical development of scientific (and technological) enterprise and scientific activities. Is science autonomous? The answer is both no and yes. Even the most abstract and esoteric research is somehow related to social influence and social perspective. But once the scientific activity advances it is pursued by many workers on its own merit. There are interplay of a social motivation or social pressure (necessity is the mother of invention), an egoistic motivation, a utilitarian motivation (applying for some practical purpose), an achievement or success motivation and a hobbyist (or, arts for arts sake) motivation.

Citation is the cultural and ethical outcome of this interplay of social and personal motivations at four different levels.

Citation can also be viewed as a process of assimilation of an idea or intellectual contribution (descriptions or written records of *mentefact* and *artefact*) in the body of common knowledge. It is also recognition of intellectual property rights. Ravetz (1971) developed the idea of intellectual property and its rights. It has been rightly said that in science (or any scholarship) it is giving away (or publicize) for establishing a claim or propriety (Merton, 1973, Cozzens, 1989). But publishing means adding to or contributing to information as well as bibliographic chaos. There is a traditional method of organising this chaotic condition and establishing order. Citation links can provide an alternative means for controlling disorder and establishing a particular type of order.

If we accept this view, then citations (or references) produce a chain through citational cycle which help suppressing natural chaotic situation in the universe of knowledge through random human activities - thus helps decreasing entropy and increasing order or negentropy, hence information.

**Implicit or Normative Theory of Citing**

The proponents and supporters of citation analysis look at the citation process in an ideal and abstract setting corresponding to the Renascent ideal of objective science. To them, citation is a part of the formal accounting process of science. References are formal records of acknowledgement and intellectual debt. There exists, therefore, a cognitive relationship between a citing item and a cited item. Only those items are cited which have some influence or relevance to the content of the citing item. Citations, therefore, act as indicators of influence. The citation process is atleast as objective and secular as the doctrine of the formal science. Science is said to be cumulative. This cumulative nature is manifested through citational links.

In this 'story book' or 'copy book' ideal of science (or knowledge), the authors are also ideal rational being taking part in an idealised pool of information communication process. With the (apparent) success of citation analyses many sociologists and authors presume that largely, if not wholly, the omnipotent existence of the ideal of objectivity acknowledging debt in science. This ideal structural and operational representation of citation has been called 'implicit theory' by Mulkay (1974) and 'normative theory of citing' by MacRoberts and Macroberts (1987).

The opponents and critics of citation anlysis who adhere to peer review process for all judgements about scientific endeavours and the process of science do also work on the basis of an idealised version of science.

The point is, science or so to say all of man's scholarly enterprise has at its core a precisely preserved ideal of story book science. The peripheral is filled with more down to earth stuff.

**Four Cognitive Functions of References**

When there is a clear cognitive role of a reference, one should be interested in classification or typology of this role. Citation context analyses try to find these roles empirically. Amsterdamska and Leydesdorff (1989) claim that the relative impact of articles over time is not the only way in which their significance can be differentiated. They could distinguish four relevant functions of the cognitive references :

(1) modification, acceptance or rejection of knowledge claim; (2) use of a knowledge claim in inferential arguments; (3) challenging new claims to further knowledge and theory developments (agenda building); (4) codification at the level of the development of the specialty.

Thus cognitive impacts may be felt at four levels : counter claim or claim adjustment; support or utilization; agenda building or filtration; and composition or contextualization or codification

(extra-terms are supplied by SKS). These four functions together form cumulation of scholarly information.

**Evolutionary Theory of Referencing**

McInnis and Symes (1988) have presented a historical view of changing status of citation and references following a theory of historical development of human character proposed by David Riesman (1988). Riesman suggested that there have been three stages of influences on the development of a person in the Western civilization. In the first stage the major influence was of tradition so the personal characters were tradition directed. In the second stage it is inner motivation; he named such characters or personalities as inner-directed. In the modern period the influence is from outside and of others which he called other-directed. In the present time the other-directed individual is the characteristics of American society and appeared in the mid-twentieth century. Production, a significant feature of the social order of the inner-directed person became secondary to consumption, a feature of the social order of the other directed person. The other-directed character signals a broad shift from nature to society, from competition to co-operation.

McInnis and Symes (1988) argue that in the early period of scholarship till the time of Renaissance the citation and referencing practice was also tradition directed. In this period the references were used mostly as evidence and for debts in controversial issues. In the 1940's, they claim "dramatic changes in the concept of bibliographic citation began to occur. Reference became an integral part of this logic of this discipline's rhetoric. It is a link in this claim of evidence associated with a topic. The notion of citation as an integral part of scholarship was articulated". It is also felt that a reference must also outwardly conform to the established format of the particular discipline in which a scholar writes. From the 60's because of citation indexes, citations or authors' names assumed the role of subject headings as well. Citations have replaced the concepts of ideas that are contained in this work. In other words document description becomes a symbolic representation of subject description. References in a scholarly discourse signify the originality of the authors of the discourse and at the same time recognition of credit to others. But this is formalised and influenced by other-directed social order.

In the inner-directed situation it was the urge of the author and influence of his immediate surrounding for giving references or footnotes. In most cases the name(s) of the author(s) (mostly with the first author's name with an et al added for multi-authored papers) and the name of the document, volume or date and page (usually a specific page number) were given in these bibliographical footnotes. In most cases, the title of the paper was not provided (it is still the practice in the journal *Nature*)

The immediate surrounding included the family and the institution where the author belonged or the immediate friend circle. The strong minded teacher or peer or superior had much to influence. But the foremost was the author's own urge and choice. For ths reason the references and referencing practices differed widely even within a discipline and within a single journal issue. But there were less chance of perfunctory, illustrative, operational and casual references.

In other-directed situation references become more standardized and formalized in representation and context but not so in content. They are influenced by peers, the norms of the discipline, the journal editors and the style manuals, the secondary services and last but not the least by the citation indexes.

In the first case, i.e. inner-directed references it was the production, the need of the paper wirtten which influenced the references. In the second case (other-directed) it is the consumption - the anticipation or thought of the potential citers and of utmost visibility and recognition that influences the referencing practice.

**Logical Dimensions of References**

Artus (1993, 1992) like others as well (Sen,1990; Sen & Gan, 1983; Vinkler, 1987) recognizes that the taxonomic analyses of references types or citing are mostly methodological approaches and they arise because of the tacit assumption or the hidden premise that references represent cognitive links and these links are of equal value (This is the major motive force for productions of citation indexes and using citation analyses). Artus argues that one should recognize the existence of cognitive - social ambivalence in referencing practice.

On his part, he also devices a taxonomic classification of references. But here he takes care of the ambivalence. Although many others have illustrated the interplay of social and cognitive motivations and functions (e.g., Vinkler, 1987), Artus takes it to a new level. To him references occur at one of four logical dimensions or levels while fulfilling both cognitive as well as social functions. These four dimensions are indicative of statuses or states of the items of reference (the terms or names of these dimensions are here coined by SKS). They are :

1. State and Status of current discourse : the actual current discourses or information as are referred to.
2. State and status of store of information and knowledge.
3. State and status of paradigmatic confirmation; prehistory of the discourse; (basis or background state)
4. State and status as sources, materials etc. (object of study and status vide...)

Nearly all functions of citations mentioned in the literature i.e., the functions ascribed to references and nearly all properties or features or processes or events ascribed to the author's activities in scientific community and scientific communication take place in the first dimension. All classification of references and citational relations refer to this dimension. All the references are representation of the function of discourse without necessarily becoming part of the discoure themselves.

Citations themselves refer to an accumulated store of information. The references cited represent the active or recognized store of information. According to Artus the huge portion of uncited items represent the store of virtual or passive (scientific) information. Either group of cited and uncited documents can tell a lot about the real discourse of the scientific community. This is Artus second dimension.

The third dimension of Artus refers to those references which Cano (1989) found to occur in the introductory sections of an article. They can also occur at other places as well in an article. Their numbers may vary according to the status of various subject areas or scientific discipline. This dimension includes self citations, background references and paradigmatic or doctrinal lineage. Artus says references with this third dimentional function are characteristics of social sciences mainly.

Fourth dimentional references are not for typical citational linkage. They do not link discourses, rather they refer to the documents as objects of study. These references are typical of humanities and social science research. In natural sciences the objects of study are natural phenomena or natural objects. In technology objects of study are artefacts. But in social science and humanities objects of study can be other bibliographic documents. These objects or documents are normally cited among bibliographic references as a rule. This is in fact a distinguishing feature of literature of humanities and social sciences from the natural sciences and the technologies.

**Holographic & Maximum Speed Principles**

Bonitz (1991) formulated two principles of science communication :

*Holography principle.* Scientific information so behaves that it is eventually stored everywhere. Scientists so behave that they gain access to their information from everywhere.

*Maximum speed principle.* Scientific information so behaves that it reaches its destination in the shortest possible time. Scientists so behave that they acquire their information in the shortest possible time.

These principles have apparently two parts. The first part is that of dissemination and diffusion of information which are source related. The second part is user or citer related. Their concern is accessioning or acquisition of information.

The principles can be recomposed in a slightly modified manner. Information dissemination is not autonomous. This process is operated and channelled through social institutions and the authors themselves. We can rephrase this part as - the publishing authors so behave and carriers of information attempts to so behave that they can reach everywhere with an optimum speed. The users of information (most of whom seek information for poducing further information) so behave that they can gain access to and acquire requisite information in shortest possible time from anywhere and everywhere.

Both these processes and functions (or principles) have tremendous influence on referencing behaviour and referencing pattern. The scientists and authors try to choose publication media which would give them highest visibility and maximum chance of being used and cited e.g,, high impact and quick publication journals and international conferences. On the other side of the scene scientists and authors try to take note of as much as and as many as relevant and potentially citable items. They also try to know how much of the work of others (both projected and completed) have already anticipated their projected work, how much of the work of others may give intellectual support and how much of the work of themselves remains really new!

Both these attempts are thoroughly incomplete. Complete holographic situation or complete diffusion never happens. It takes a very long time for complete diffusion as the speed of diffusion slows down with time and through encounter of other ideas, information and personalities. There is not a single channel for dissemination; speeds of all the channels are not the same. Not all authors can take the fast channels (quick journals, seminars, press conference etc.). Similarly there are barriers to speedy diffusion and easy accessibility. No single channel or vehicle (the publishing document e.g., disc, book, proceedings, journal) can not reach every space point.There is again a skewed distribution of circulation. Faster vehicles or channels may not have the highest circulations and vice-versa. Accessing and acquiring and using information and information sources then are not what Bonitz's principles demand, rather the actual situations may be caricatures to the Bonitz's principles.

One can take typical examples of a publication in a small third world journal — which takes years to publish an item - after a manuscript is received, always lagging behind schedule and reaches almost nobody outside a very limited subscribers within the country of origin. If it is in a language other than English, German, French or Russian, it would get none of the target readership i.e., the potential citers. On the other hand, the author in a similar third world situation would not gain access to most of the important items. This author would know most of the information via secondary sources. His database for using in his writing and hence citing (referencing to) would be a hotch potch conglomeration of very old and new, relevant and marginaly relevant, important and unimportant items. In many cases he would know about the existence of certain relevant or potentially relevant items but would not be able to 'see' and 'consult' them. Hence he can not cite any of them. Even if he cites that would be at random with fair chance of being incomplete and error-prone.

In case of an author working in a highly advanced situation, problem is of selecting, choosing and discarding to keep the size of the reference-set to an optimum. In case of an author in backward situation it is the problem of citing whatever at hand (no choice).

Bonitz's principles also represent the idealised situation, not only of the science as a social enterrprise but also of the social academic system as a whole. In a third world or a backward environment the potential author or researcher does not have means for quick access or getting into an international invisible college. The information or rather the information carrying documents would not reach him or his immediate environment at all. There may not be fund for foreign journal subscription. Not much have been written about these barriers and channels in connection with citation and reference studies.

***Ultimately we have a picture which is something like this :***

Authors and media of dissemination behave in such a way so that the infomration product can become most visible within minimum time. There is however a whole range of time scatter; the speed of dissemination and reaching out varies with the medium, the author's environment and with time. On the other hand there is a wide variation in accessibility of potentil items. Those carriers or media are most easily accessible or available which are generated within the social space, time, and environment. But in many cases most easily accessible items may not be the most useful ones. Thus we have holographic pockets. We may enunciate a model. The process of science publication or information production process of science aims at maximum speed holographic model. Due to several barriers, this ideal is never attained. Thus we get instead some rich pockets or rich

environments where all information and documents become available (can reach) in shortest time. A citer or author in this environment has a problem of choice — what to keep and what to discard (let us call this *choice problem*). On the other hand the citers in poor environments or por pockets are victims of slow access and non-availability. (We may call this *no-choice problem*).

Ultimately modified principles of Bonitz show us that in case of referencing and generating citation cycles a Mathew effect is evident. There might be large international citation cycles with participation of rich pockets and there might be some small satellite citation cycles with middle level and poor pockets. This aspect would definitely need a large amount of empirical research.

**Referencing Typology and Citer Motivation**

There are potentially many citable items. Only a few among them are really cited. There are some items which are taken as references from outside the set of potentialy citable items. Not all the citable items are known to the citer. Some items may only be known from intermediate sources. Again, some usable information can be known via a documentary or non-documentary source for which the original source may not be known at least to the extent of making a specific reference. Even after all these, there may be a disproportionately large amount of potentially citable items known to the citer for inclusion as possible references. But the accommodation is limited. There is therefore, an inherent competition among the citable items. The situation has prompted studies in referncing behaviour, typology of references and citer motivation (Cronin 1983). There are many classification lists of citation types or taxonomy of citer motives.

We have been able to locate more than a score of such lists. Some of them are Lipetz (1965), Weinstock (1971), Chubin and Moitra (1975), Moravcsik and Murugesan (1975), Spiegel-Rosing (1977), Garfield (1977), Thorne (1977), Hodges (1978), Oppenheim and Renn (1978), Frost (1979), Finney (1979), Duncan et al (1981), Bonzi (1981), Peritz (1983), Brooks (1985), Amsterdamska and Leydesdorff (1989), Sen (1990).

Almost all of them are covered in either of Cronin (1983), Egghe & Rousseau (1990) and Brooks (1987). Sen's list (being published at a later date) is not included in any of them. But this list covers all the aspects in a simplified manner and is being reproduced here. The list in this review would be referred to in occasions and theoretical studies would be seen to be fitting with its categories. There are three main categories of items used or cited in a document :

a) used and cited;

b) used but not cited;

c) cited but not used.

The items in the first category may be for various purposes, such as —

i) for furtherance of ideas or generating newer ideas related to the ideas in the cited item. If gradings are attached to references, such items would get high ranking.

ii) For experimental vindication.

iii) For theoretical justification.

iv) For refutation and criticism.

v) For suggesting alternatives

vi) For a historical or background review.

vii) For achieving importance or support via authority.

viii) For listing similar or related works.

ix) For making passing remarks

x) For using result(s), illustration(s), example(s), processes(s), design(s), etc.

xi) For nominal discussions in very general terms, especially in the case of books or reviews

In the second category are those items that

i) contain well-known, quite old and generally accepted items, already incorporated in the common body of knowledge. Nobody specifically refers to 1905 paper on special relativity by Einstein, except in the case of historical research;

ii) may be deliberate omission;

iii) may be conditional omission : of three documents having contents of equivalent nature, the least important one being omitted; or during the preparation of a bibliography exact reference was not at hand; or the ideas were known via secondary sources.

The items in the third category are those cited for

i) civility

ii) creating corresponding visibility

iii) acquiring importance and this occurs mostly in the case of self citations or citations of important peers

(d) There is another category or situation where there can not be any corresponding cognitive reference because of the very genuine fact that the idea presented is so novel that there exists no known earlier reference. This may also happen in case of a re-discovery or use of some idea that might have been published long ago or in an obscure source not knwon to the author. In some papers ideas may be so novel and as the background information utilised may be taken for granted as common knowledge, the papers may not carry any reference at all. One classical example is Einstein's 1905 paper on special relativity.

If one looks through any or all of the taxonomic lists of citational motivation and referencing typology one would find that only a small proportion of cited items is really worth the status .

MacRoberts and MacRoberts (1986,1989) have shown that the items that are used and cited are of insignificant proportion to the implicit and informal influences that make the text of a research paper. It is however to be understood that only formal documents or publications can get a place in the list of references. Consequently, the citations are also formal documents. There can not be a non-documentary or informal citation or reference to be taken into the citation analysis or for a formal theory of citation process.

Cano (1989) has made an analysis of locations of reference in the body of a research paper. He has found that references are most heavily concentrated in the introductory sections and many of them are of historical and perfunctory nature. Most of the perfunctory type of citations have low utility level. It becomes apparent that the notion of citations or references as discrete measurable information units of roughly comparable cognitive value becomes shattered. References "do not reflect influences that belong to the intellectual baggage accumulated by years of intellectual pursuit". Rather, there is high mismatch between the intellectual baggage or the spectrum of real influence and the set of references or the cited pack. Hence the citing pack does not betray (expose) the actual extent of impact or influence (cognitive or informational).

One may support Cano and Bertram that more research is needed to find out the relations between locations and influences of references for a text and to discern the elements of information cited to get a ranked distribution with relative cognitive connection (or influence) from the cited items to the citing item. We can think of a research programme where the influences can be put on a scale (say from +1 to -1) and also the locations of the references on a scale and put them on a correlational map. Many such studies can only give us an idea, somewhat mechanically, about relative influences of references (so of citations). Even then no such scaling can be incorporated in a citation index. Whatever be the situation, in a CIS, every citation would appear to have the same influence or impact value. It is an open question, however, what can a theoretical model accommodate.

Citer motivation comes mainly from persuasiveness rather than any other reason. This means the motivation comes from the anxiety of getting the paper pass through the filter of peer-reviewing process and reaching the potential user. Brooks (1985,1987) listed other types of motivations such as currency, negative credit, positive credit, operational information, reader alert and social consensus. These may be considered as articulate or idealised motivations. But we have already seen that references can occur because of other inarticulate reasons and extra-academic motivations.

With all this background we can now reformulate the basic assumptions for citation process, citation phenomenon and citation analysis :

*Ignoring marginalities, abnormalities, and peripherals, the practice of bibliographic references and the process of citation imply in essence an irreversible time dependent cumulation of cognitive relation among the articulated items of literature i.e. between the cited and the citing documents.*

Citation process ignores the inarticulate influences on research. Citation process captures

only a small segment of potential citable items available in the literature. The size of the citational pack is limited by recency, availability and literary warrant (by literary warrant is meant a status of public knowledge or common knowledge incorporated in the accepted body of formalised knowledge).

We do not survey the literature on citer motivation, referencing behaviour and taxonomies of reference types, motivational influences etc. Nor do we present lists of such taxonomic classifications. Instead some coherent pictures of theoretical structure and possibilities are prepared from many sources and presented in the following sections. There is however a need for making a synthesis of all the lists and present a holistic analytic picture. This is beyond the scope of this discourse.

**Scheme for Citing Process Complex**

Referencing involves knowledge or awareness about existence of information sources and also about the contents of those information sources. Knowledge of information sources is required for locating and identifying the items one has to search. The motivation for searching information and information sources arises from a desire or an informational need. Unless the need is strong enough no search occurs. Lancaster has schematically presented the information seeking, information gathering and information use behaviour. He then enumerated different types of activities and influencing factors of the motivations for seeking information and selecting particular sources. When we are concerned with citation process, the type of activity that generates information need is research. The influencing factors are the extent of collaboration, interdisciplinary associations, involvement, research result etc.

When one starts seeking for information the seeker may or may not be aware whether the information already exists or not, whether the information is available in recorded, comprehensible form. Usually, the search starts only because the seeker presumes existence of obtainable information sources. The search procedure and information gathering depend upon a number of factors. The most important factor is the experience or background of the searcher. The other factors are previous experience of searching similar type of information i.e., the knowledge of different channels of information flow, availability of searching tools and agencies capable of information supplying (e.g., good library network, SDI and current awareness services), secondary services, nature of information sought (whether supporting information or contradictory information); the value of information and the cost of obtaining information; personality factors such as influences of peers, members of the research teams, colleagues, extent of invisible colleges with which the members of the research team may be involved etc.

After the possible addresses of potential information sources are known (after being aware of the existing sources) an action of selecting a particular information source from among those known has to take place. Only those sources are selected which are most easily available (principle of least effort). In some cases, however, much effort may be given to collect some information source on the face of extreme difficulty. A text in an alien language in an obscure document existing in a far-off minor institution or library may be copied, collected and translated spending large amount of money, time, manpower etc.

All these factors ultimately influence references in a publication and contribute to the citation process. Taking the cue from Lancaster, a scheme is presented here :

The complexity of citation process leading to the decision of citing is a sequence of activities, awareness generation, felt need and expressed requirements, urge for getting hold of the information, searching and securing, using information, producing an information product or intellectual product, getting it published.

A. Activities are

- : Research (wherefrom the need for information mainly arises)
- : Collaborating
- : Discussing with others
- : Literature searching
- : Procuring and acquiring information and document
- : Reading, comprehending
- : Keeping ready for recognising relevant

information
: Reporting, writing
: Citing
: Modifying, changing (upon feedback or upon new development)
: Final referencing

B. Influences on reference size, typology and the actual references are -
: Past experiences (past activities)
: Availability of particular type of document e.g. a current bibliography or review article
: Colleagues (through discussion)
: Invisible college (preprints, reprints, private letters)
: Formal meetings, conferences etc.
: Informal and chance meetings
: Peer review, peer suggestions
: Editorial suggestions
: Egotism

C. Effectiveness of search and acquiring information depend on
: Involvement in general
: Time available for completing the study and publication schedule
: Other work or assignment than the research topic
: Interdisciplinary association
: Extent of collaboration
: Available support both human and technological
: Personal urge
: The action as a task (it applies to a person in employment or in a funded research project).

D. Source related factors are -
: Physical, intellectual and psychological accessibility (These are also related to negative influences and personal and environmental factors) of the sources
: Perceptions of accessibility
: Quality of the source items, which include authority, reliability, objectivity, trust, currency and treatment.
: Format or content of presentation (this is related to comprehensibility)

E. Barriers and negative influences may be -
: Low impact journals
: Less prestigious or spurious seminars, conferences etc.
: Unattended meetings
: Obscure or rare sources
: Alien language
: Contradicting information
: Cost of obtaining information
: Negative environmental factors e.g., remote region, no good library, no peer help, no eficient tools etc.

F. Strategic decisions depend on the factors which may be categorised in the following manner

(a) Personal or Citer Oriented :

(i) The position or the designation of the citer (researcher). A high position gives social and financial power and hence better access. The strategic decision would depend upon using that power (of position and designation) in harnessing information.

(ii) Membership to various regional, national and international bodies or organisations. These may be utilized for accessing information. Sometimes, documents are as a rule made available to such members. Editors and referees have early access to some of the current literature.

(iii) Time. The most important factor in strategic decision of acquiring and using information is time This is linked with cost, effort, source, type etc.

(iv) Egotism, prejudices, accidental causes.

(v) Selection of the actual items to put as references is the ultimate strategic decision.

(b) Environmental

(i) Available financial resource or

organisational support.

(ii) Trade-off between reward and cost. This strategic decision depends upon the former one that is of availability of funds to speed or organisational support to take care of expenses. Decision needs to be taken whether accessing or acquiring a document would be worth the expenditure to be incurred.

(iii) Extent or amount of effort needed. There is always a threshold to this. Least effort path is naturally followed. Extra effort needs to be within threshold. But this threshold is somewhat elastic depending on importance and other factors.

(iv) Access. Access means here access to good libraries, personal collections (may be of reprints); access to databases and secondary services; access to new technologies, which again involves cost such as FAX, Online, E-Mail, CD-ROM etc.; and access to networks of preprint and reprint circulations within invisible colleges.

Accepting the conflict between principles of pain avoidance and least effort on the one hand and personal urge, reward motivation and/or task on the other, the complex of citer motivation or atleast the complex of referencing can be enumerated. The scheme presented here does not try to classify motivations or motivational types. Rather it shows the different stages through which the ultimate reference is made. This shows that —

(1) a reference is a random variable, so is a citation.

(2) referencing or citing is integral part and parcel of human social behaviour of knowledge seeking and acknowledging the source of information.

**Vinkler's Quasi-quantative Model**

Sen (S K Sen, unpublished paper, 1982) defined applied information science or the work of information professionals as the process of producing raw materials out of finished products. If scientific research is 'information producing process'(Vinkler, 1987), then the finished product is the research paper or citation. When it is used, in its turn, it is used as a raw material, as information input. It then appears as a reference to a new information product.

Vinkler suggested a quasi-quantitative model of citing. This again dealt with motivational aspects. He also recognizes following Solla Price that "There are no institutionalized, accepted and widespread norms or rules of behaviour for publication and citing activities". Yet there is a "relatively well working self-controlled publication system".

Vinkler's model "was not to quantify the effect of motivations of such a complex intellectual process, ending in a decision as citing". However, the model attempted at a proportional categorization of various effects in citing. In this model "citation is regarded as one of the products of scientific research as information producing process"

Vinkler's approach for informational influences in a paper and their reflections in references fit in the categorization made by Sen (1990);

(i) used and cited;

(ii) used but not cited:

(iii) cited but not used;

Relative (proportional) significance of references in the third category was also attempted to be estimated quasi-quantitatively.

Vinkler recognized only three major types of motivation : Professional motivation (P), Connectional motivation (C) and otivations for neglecting or missing citations (N). Professional motivations are cognitive motivations which result in both relevant use of information and citing (Sen's first category). The connectional motivations result in "halo-effect" citing (Sen's third category). N represents the cases of using but not citing(Sen's second category). He enlists a number of sub-categories of motivations in each of P, C and N. He also defined notions of cognition threshold, citation threshold and reference threshold and relevance limit.

Vinkler showed the citation scenario diagrammatically on a rectangular area. He presented the strength of motivation on a 0 - 3 scale. He represented his model (atleast the proportional figures) on the basis of an empirical study made by himself.

At the first stage there are the professional motivations (P) and connectional motivations (C). Influence of P on referencing is about 9 and 10 times that of C. In the next stage in the model there is an analyses of the use of past information. About 20-22% of information result in N. Half of them relate to the incorporated information that is the text bok information or common knowledge type. The other half (i.e. about 10% of the total) is the hidden information mainly due to *review effect*. The information used were to be shown through references but are not because they are received via a secondary source or intermediary.

The other part is shown as the portion of connectional references and professionally irrelevant information. Relevant and irrelevant information vis-a-vis references ($R_p/R_c$) should be in the ratio of 6:1. However all these proportions may vary from sample to sample.

Vinkler defined citation threshold as the lowest value of the cognitive pressure. The strength of the cognitive pressure depends on the use of the information or recognition of the extant information in past publications. There is therefore a cognitive threshold leading to citation threshold and ultimately to reference threshold. Reference threshold is different from the citation threshold because references due to connectional reasons do not have to cross the cognitive threshold for being referred to.

In Vinkler's scale from cognitive threshold at zero level to citation or reference threshold at three, irrelevant information in professional motivation was upto level 0.3. Extent of negative effect of motivations was between levels 0.3 and 0.6. Extent of review effect (via information) was between level 0.6 and 1.0. Extent of professional incorporation was between 0.6 and 1.5. The extent of connectional effect was between levels 1.5 and 2.0 of the threshold scale.

This mosaic, Vinkler hoped, could lead to a theoretical model of citing. He was however more successful in devising a methodology for empirical study of these aspects or features which could be used as indicators.

The three stages of motivations, use of information and actual referencing set against extents of cognitive-citing threshold scale may be taken as a useful methodological model rather than a theory.

**Formal Theories of Citation**

The formal theories start with some propositions taking citation as an event and then attempts at proving theorems.The proofs of the theorems collected here will not be given in full.

**Citation as an Event**

Citation as an event manifests simultanously two outcomes or events namely the event of citing (referencing) and the event of being cited. We have already seen that this event is a result or outcome of a complex sequence of causes and influences, the so called motivations, operations and situations. The event also links the citation with the references and changes the net citation map of earlier items. The following process is autonomous to some extent. This process requires formalization.

The verb, 'cites' is a two.place predicate in the sense of mathematical logic indicating a relation on all possible pairs of document tallies (Kochen, 1974). The verb 'cites' has its reverse or inverse expressed in the phrase 'is cited by', e.g., *a* cites *b* implies *b* is cited by *a*.

These two manifest the binary relations between the cited and citing items. In the following, the relation 'cites' will be denoted by $C$ and its inverse is cited by $C^{-1}$. $C$ is the matrix $C_{ij}$ linking items $i$ to $j$ therefore $C^{-1} = C_{ji}$.

$C$ is a time dependent relation and the process of citations is a one way irreversible.

**Propositions (following Kochen 1974, Sen 1990, Egghe and Rousseau 1990)of Citation Event**

1. Let $C(d)$ denote the set of all documents that a document $d$ cites. In other words $C(d)$ is the set of all references in $d$.

Then $C^{-1}(d)$ is the set of all documents from which d receives a citation.

2. Let also $d^1 \in C(d)$, $d^1$ is a document in the set of documents $C(d)$; then $C^{-1}(d) = \{ d_i : d_i$ is a citing document$\}$.

This means $d^1$ as cited item has the $d_i$'s as the set of citing items or citations.

3. $C$ is a time dependent relation. If $d_k$ is a randomly selected document at the time $t_k$, then among all the documents added during a unit period of time, $t_1$, after $t_k$, there would be a set of documents $C(t_1/d_k)$ would vary with $d_k$ as a random variable. The set of the random variables $C(t_1/d_k)$

can also be taken as a random variable which may be denoted by $C(t_l/t_k)$.

4. Mean of the random variable $C(t_l/t_k)$ would represent a coupling constant in item-item linkage.

5. Any item $d_k$ can be considered as a function of time $t$ such that $d_k = f(t_k)$ and $t_k \in T$ where T is a completely ordered projection on positive real line. Thus $t_k$ and hence $d_k$ is a time series.

6. Let the number of documents published at a time $t_k$ ( $t_k$ is not just a point but is a small interval of time, e.g., a month, even a year) be denoted by $N(t_k)$ and the set of items having citation linkage between each pair of items in the set at an instant, $t$ (or small interval) be denoted by $D(t)$.

Then $D$ is the set of all items that are citationally linked.

**Theorem 1 (Kochen 1974)**

$$d \in \bigcup_{d' \in C(d)} C^{-1}(d^l)$$

This simply means that if $d^l$ is one of cited items of $d$, then $d$ is a member of all the citing items of $d^l$. That is, sets of citing items of $d^l$ includes $d$ as a member, or $d$ is one of the citing items.

**Theorem 2 (Kochen 1974, Egghe and Rousseau, 1990)**

$$d_l \in \bigcap_{d \in C(d_l)} C^{-1}(d)$$

Where $d_l$ is a fixed document of $\{d_i\}$, the set of citing items.

The result comes directly from proposition 2.

Egghe and Rousseau (1990) describe it as a proposition. It merely states that $d_l$ belongs to set of all documents that cite $d$.

**Corollary**

Egghe provides a corollary to the theorem 2 (Egghe and Rousseau 1990) :

For every document $d_l$

$$\bigcap_{d \in C(d_l)} C^{-1}(d)$$

is the set of all documents $d_i$ such that the reference list of $d_i$ includes the reference list of $d_l$. (For proof see Egghe and Rousseau 1990).

**Theorem 3 (Kochen 1974)**

$$D = \bigcup_{j=1}^{\infty} C_j$$

Where $C_j = \bigcup_{di \in C_j^{-1}} [C(d_j) \cup C^{-1}(d_j)]$ and $C_0 = C(d_0) \cup C^{-1}(d_0)$ for some $d_0 \in C_j^{-1}$

This is direct outcome of proposition 6.

**Corollary**

Egghe and Rousseau (1990) give a modified form using a finite union instead of an infinite sequence of unions. This modified form can be taken as a corollary which states :

If the citation graph of a non-empty, finite set $D$ of $n$ documents is weakly connected (that means members of $D$ have citation linkages with one another either directly or indirectly, i.e., through one or more members in between), then, for any $d_0$ of $D$ —

where $D = \bigcup_{j=0}^{N-1} C_j$

$$C_j = \bigcup_{di \in C_j^{-1}} [C(d_j) \cup C^{-1}(d_i)], \qquad i > 0$$

and

$$C_0 = \{d_0\} = C(d_0) \cup C^{-1}(d_0),\ d_0 \in D,\ N(D) = n$$

It may be seen that in Kochen's case, union is infinite which means it covers all the items in the history having weakly connected citational linkage. In case of Egghe and Rousseau's modification we can consider a finite cluster through an arbitrarily chosen finite number of citation cycles. In Egghe and Rousseau's words "this theorem yields an algorithm for obtaining all the documents of a given collection, provided the collection is reasonably homogeneous, so that the citation graph is weakly connected. Moreover, if $D$ is a large computer file, then the algorithm gives a procedure for exploring the core of a topic (take $d_0$ to be a case document) and moving further and further towards the boundary" (periphery).

The proof is straight forward. Following Egghe and Rousseau it may be argued that any set $C_l$ can be the set of all documents which either cites $d_0$ or are cited by $d_0$ . $D$ is union of $C_i$'s, therefore, $C_l$ is not empty unless $D$ is a singleton $\{d_0\}$. On

the other hand if it is supposed that some $d \in D$ does not belong to $\bigcup_{j=0}^{N-1} C_j$, a contradiction arises. Because of weak connec-tedness there is path (or link), necessarily finite, joining $d$ to $d_0$ whose number would be between $j$ and $N-1$. This directly implies $d \in C_j$. The proof actually presupposes a weak citational link of every item to every other in $D$.

**Theorem 4 (Kochen 1974, see also Egghe and Rousseau 1990)**

The average number of references per item times the number of items being considered (or published at a time $t_1$) is equal to the number of cita-tions per item (a fixed document) times the total number of citing items or references (at $t_2$).

The appearance of time as parameter is not necessary in the proof, $t_1$ and $t_2$ can be considered as representing earlier and later documents. $C(t_2/t_1)$ just represents citation link from $t_2$ to $t_1$, i.e., $t_2$ cites $t_1$.

The statement of the theorem can be symbolically written as —

$$\bar{C}(t_2/t_1)\, N(t_1) = \bar{C}^{-1}(t_1/t_2)\, N(t_2)$$

(For proof see Egghe and Rousseau 1990)

Egghe and Rousseau give an example of its application to find out the average number of citations per item from a pack of items (may be randomly chosen or may be for some specific set of items). Let the total number of citing items be $N$ and the average number of references per document be $R$ and the total number of references be $S$ then the average number of citations would be $N.R/S$.

The universal value of this ratio should come out near to 1.7.

The easiest way of finding this is by producing citation matrices. The number of columns would give the number of citing items and the number of rows – the number of cited items. The number of references can easily be found from the matrix.

**Theorem 5 (Kochen, 1974)**

The citation linkage relations or functions or citation matrices as random variables are independent of each other. In other words, probability that an item $t_1$, is cited by another item $t_2$ does not depend upon the event that $t_2$ being cited by $t_3$. $t_2$'s appearance as reference is independent of $t_2$'s appearance as a citation.

This can be symbolically written as —

$$Pr\,[C(t_2/t_1) \cap C(t_3/t_2)]$$
$$= Pr\,[C(t_2/t_1)] \cdot Pr\,[C(t_3/t_2)]$$
$$and\ Pr\,[C(t_2/t_1) \cup C(t_3/t_2)]$$
$$= Pr\,[C(t_2/t_1)] + Pr\,[C(t_3/t_2)]$$

This can be generalised for sets of items. Appearance of a set of documents as references is independent of their appearance as citations.

**Theorem 6 (Kochen 1974)**

Average number of citations per article multiplied by the rate of publications is the number of citation pairs.

**Theorem 7 (Kochen 1974)**

If the average number of citations $\bar{C}(t)$ remains the same over time, then the citation pairs grow in the same manner of growth of literature.

**Corollary**

If self citations are neglected or suppressed, the resultant parameter for citational growth would give the proliferation of information.

Proofs of these theorems depend directly upon certain definitions and assumptions.

**Assumption**

The citing items represent the growth of literature, because, among the total number of items published in any epoch there is only a negligible portion of items which do not carry a single reference ( i.e., is not a citing item).

**Definition**

Let the rate of growth of citing documents be denoted by

$$D'(t) = \frac{d}{dt} D(t) = \lim_{\tau \to 0} \left| \frac{D(t+\tau) - D(t)}{\tau} \right| = \left| \frac{A_\tau\, D(t)}{\tau} \right|$$

Here $A_\tau$ denotes the number of documents added during the period $\tau$ after the time. $D(t)$ is the number of citing documents at time $t$.

Theorem 4 can now be generalised as —

$$D'(t_1) \cdot \bar{C}(t_2/t_1) = D'(t_2/t_1) = D'(t_2) \cdot \bar{C}^{-1}(t_1/t_2)$$

where $D'(t_2/t_1)$ gives the number of citation pairs generated due to $A_\tau$.

From this, the theorems, 5, 6, 7 are straight forward conclusions. In fact, these theorems can be taken as corollaries to theorem 4, if the definition and assumption are taken as lemmas.

The assumption made above has a connection with the theorems stated later.

**Theorem 8 (Kochen, 1974)**

The probability distribution of citations $C(t_2|t_1)$ does not depend on either $t_1$ or $t_2$ but depends on the interval $(t_2 - t_1)$ only. The theorem implies that the average number of citations per article would not change over time.

The proof follows from theorems 4 and 5.

This *stationarity theorem* has its relation with the law ofconstancy of accessibility of information documents (q.v.) and constancy of the size of reference to research articles (q.v.).

**Theorem 9 (Krauze and Hillinger 1971, Garfield 1976)**

In Garfield's words the statement is :

The ratio between the number of citations and items cited is a constant.

Garfield did not enunciate this observation as a theorem nor he supplied any proof.

Krauze and Hillinger provided a derivation for this constancy of ratio of citing and cited items. In their version the theorem states :

The ratio of citations and references is constant and necessarily greater than one.

**The Proof**

Krauze and Hillinger distinguished between the half lives of references and citations. Let $h$ be the half life of references that is the time lag (from the publication date of the citing document) during which, half of the references are made.

Similarly, let $k$ be the half life during which half of all the citations are expected to occur. The values of $h$ and $k$ can be derived easily from the well known formulas of decay, growth and half life (as are found in the physical science).

They showed that the ratio of citation and references at a time spot $t$ may be given by $C(t)/R(t) = k/h = \text{constant} > 1$.

This ratio is actually a quasi-constant rather than an absolute constant. (The proof is not given here).

The above theorem should however be modified as follows :

**Theorem 9'**

The ratio of the number of citations and references at a given time and for all the items is virtually a constant.

*No one so far estimated this constant ratio for empircal values.*

It is somewhat interesting to note that Garfield did not cite Krauze and Hillinger.

**Theorem 10 (Sen, 1990)**

The cited pack (the set of references) of all the items published in a period are the only items whose citing pack (the set of citations) accounts for all the publications in that later specified period which carry bibliographic reference.

This is self explanatory.

The items which are cited are the generators of the citing items.

The linkage between the set of cited items and the set of citing items is unique. (For a detailed proof see Sen, 1990).

From the foregoing it becomes clear that the growth of literature or growth of information is synonymous or equivalent to the growth of citing items. Ratio of citing items and cited items is greater than unity.

We can also state a few more corollaries or theorems with the help of above theorems.

**Theorem 11 (Corollary to theorem 10)**

Any single item of the past can be cited many times or not cited a single time. But as the citing pack is much larger than the cited pack and as the cited pack is only a small portion of all the citable items, a citational scatter occurs.

**Theorem 12 (Corollary)**

The growth of knowledge of useful information is cumulated through a core flow of items which are citationally linked.

This supports the elite hypothesis and speaks against the Ortega hypothesis.

**Theorem 13**

Only a small representative subset of the set of all the references appended to items accounts for all the items that have been cited.

**Response Time Mode of Glänzel**

Glänzel has attempted a stochastic model of the citation process and has come up with some verifiable indicators based on the concept of stopping time. Schubert and Glänzel in 1986 defined an indicator called *mean response time* (MRT). The response time (RT) was defined as the time elapsed between the date of publication of a paper and the date of first citation received. RT could be infinity at one extreme (case of uncitedness) or very near to zero at the other extreme. They assumed that a five year citation period would be sufficient to cover almost all the cited papers (those uncited within five years would remain so for all future time). MRT was then estimated as an exponential average over five years for a set (sample population) of articles, e.g., all articles published in a journal during a year. Thus —

$$MTP = - \ln \sum_{i}^{4} f_i \, (exp - i)$$

$$= - \ln \{f_0 + f_1 \, exp \, -1 + f_2 \, exp \, -2 + f_3 \, exp - 3 + f_4 \, exp - 4\}$$

$f_t$ is the fraction of papers receiving their first citation in the $i$th year after publication.

Glänzel (1992) defined another indicator : harmonic mean response time (HMRT). To do that $f_t$ is replaced by $f_t = G_1(t) \, (exp(-t) - exp \, (-t + 1))$

Now if one wants to account for recurrent citations then one has to consider not only $G_1$, the factor of first citation but also $G_2$, $G_3$, $G_4$ ... etc. the factors for second, third, etc. citation spanning a finite time period say $n$ years. The HMRT for the $i$th response would be

$$HMRT \, (i) = \sum_{i=1}^{n} \{Gi(t) - Gi(t - 1)/ (t + 1) + Gi(0)\}^{-1} - 1$$

This Glänzel asserts to estimate the expectation ratio

$$E \, (Ti \, / \, (1 + Ti)) \, / \, E \, (1/(1 + Ti))$$

He then shows semi-empirically that HMRT $(i)$ can be approximated by a linear formula

$$\text{HMRT}(i) \approx a. \, i, \; i = 0, \, 1, \, 2, \, .....$$

and $a$ is a positive real constant which characterises the response successions as reflected by citations. The reciprocal $1/a$ expresses the speed of successive responses. This index $(1/a)$ was named as *average rapidness of citation successions* (ARCS). He argued that MRT, HMRT, the average citation success time (ACST or, $a$) and ARCS can act as immediacy indexes and can show the citation successions. The linearity behaviour was checked with papers in two journals (*JACS* and *Lancet*) and a group of papers on a topic (papers on condensed matter during 1980). He concludes that the first citation is always a significant factor and indicator despite its being self citation or not.

**Stochastic Birth Model of Citation Process**

Glänzel and Schubert (1992) used the dynamic Waring Model to generate a stochastic model of citation process. This is a particular version of pure birth process with or without immigration. These two conditions give rise to two different versions. In citation process, during a certain period a number of articles are published or generated. Successively citations are produced to those papers. The size of the first set is fixed *a priori*. No new immigration or emigration is possible. The second version of the model is suitable here. The annual citation rates (for both the versions through dynamic Waring model) should obey negative binomial distribution (NBD). There would be two parameters for this NBD, one is a constant parameter N and another, Q(t) is a function of time t. The authors tried to verify the predictive aspect of both versions of the model. Estimated parameter values and values of successive annual frequencies of citation rates were in agreement with 'average' cases but not with 'elite' cases or highly cited items.

**Nakamoto's Model**

Nakamoto in two papers published long back in Japanese (1964, 1965) produced a theoretical model of obsolescence of literature which he applied to the rates of changes in citations from SCI® and also to references. In other words his mathematical model could account for both diachronous and synchronous cases. The decrease in citation age could be shown to be exponential and independent of the number of source items. In his model diachronous obsolescence coincides with synchronous obsolescence. This again indicates the constancy of average reference and average citations per item over the years. Hence the ratio of average reference and average citations should also be constant.

**Law of Constant Citation (Gomperts, 1968)**

Gomperts (1968) in a semi-empirical, semi-theoretical paper found that the product of the citation factor for a particular year (say *j*-th year) and the total number of articles that have been published upto and including that year (i.e. the *j*-th year) is a constant. He defined citation factor $R_j$ as the ratio :

$R_j$ = (real number of citations upto and including year, j) / (maximum possible number of citations upto and including year, j).

He studied all the available papers in English, German and French on 'vibrating plates' in acoustics for 150 years since the publication of Chladni's findings in 1787. He took $R_j$ as a function of $a_j$ number of articles published upto and including the year j. The graph of $R_j$ against $a_j$ was hyperbolic for his sample. He found $R_j a_j$ = constant = 5.6.

He also showed theoretically that the mean number of relevant citations per article per year remained constant at a value of 2.8 which is half of 5.6 inspite of the fact that three times as many articles were published on the subject after 1950 than were published before 1950. Gomperts suggests that such a law of constant citation should prevail over all the subject fields or topics. The averge number of citations per paper in SCI® data base as we know is about 1.7.

**Analogy to Electron Lattice Scattering**

Sharma, Singhal and Gupta (1980) attempted to visualise the citation phenomenon in analogy to electron lattice scattering. In this paper Sharma et al made a novel approach. Instead of taking citing or cited items they took citers as the main agent for citation process. A group of potential citers were considered as analogous to a group of electrons, and a lattice to a system of research papers which may collide (or come across) a citer in a stochastically random manner. They followed the theoretical model of electron lattice scattering and came out with a formula for mean life within which the average number of potential citers reduces to (1/e)th part of its original value. They ultimately produced a new impact index which they called *Absorption Index Activity* (AIA).

I feel that this model could be modified and generalized and could be utilized for understanding referencing and citation process from a very different angle of view.

**Analogy with Beer's Law in Optics**

Van Raan (1988) made a very similar approach to that of Sharma et al, but with some obvious differences. In analogy to Beer's law in optics the impact of a publication (research paper) can be compared with radiation (radiation may be photons, electrons or other such sub-atomic particles) passing through an absorbing layer. In Beer's law the amount of absorbed radiation is proportional to the incoming radiation intensity and the thickness of the absorbing layer. Thus the eventual impact of a paper would depend upon the initial impact it can make and the number of citers it can reach. Van Raan discussed three consequences of such a frame-work : (i) the colleague-scientists i.e. the potential citers judge the values of publications (relevance ?); (ii) the 'inter-subjective' or 'collective' judging of usefulness is represented in a statistically significant way by the number of received citations; (iii) it must be assumed, as a consequence of the above two situations, publications compete with each other for usefulness (this is the thickness of the absorbing layer).

Van Raan's attempt was more empirical and heuristic. It is however apparent that taking potential citers as randomly moving particles, so also the potentially citable items — one can build up a model. Probably the Goffmann's model of spread of ideas as contagious disease can also be utilized by suitably modifying it.

**Citational Inequality and CAP**

Allison et al. (1982) attempted a model of cumulative advantage and inequality count for understanding scientific productivity of publications. They also claimed that this should apply to citations. All the same, they suggested that there are several reasons to expect a more rapid increase in inequality for citations than for publications. First, the number of citations per paper is relatively constant for a particular author; change in the number of publications will change citational inequality directly. That means citation score of an author depends upon the number of publications to his credit. Second, there should be increasing inequality over time in the average number of citations per paper since the work of a prestigious author is more widely read. Third, there is a limitation on the production of research papers. There is a productive life for

research and no one can work more than 24 hours a day. But citations are generated by many persons over the years. So, there are no obvious limitations for generation of citations. They, however, show that there is much greater inequality of citations than that of publications; but there is no evidence of a relative increase in inequality over time. This is to say that citational pattern remains stable or of the same form over the years. The empirical results found by them signify a very important property of citational distribution or inequality. Citational inequality is not built up gradually over a long time, but becomes well established within a few years. *Other empirical studies have shown that this time-span extends from two to seven years.* They also suggested that older publications are cited with less inequality than more recent ones. Increasing inequality of citations to recent work is counter balanced by decreasing inequality of citations to older works.

They worked on two elements of the hypotheses of CAP which are crucially important in their view. These are : (i) each publication increases a scientist's propensity for future publications; similarly each citation increases a scientist's propensity for future citations; (ii) the occurrence of publications or citations is at least partially governed by random processes. Thus fortutious citations (and publications) are converted into lasting advantages.

We have already noted the significance of these two hypotheses or principles. They may therefore be taken as integral part of citational processes.

Allison et al (1982) used the contagious Poisson process as the basis of their model. One implication of the model is that the cumulative number of citations has negative binomial distribution. The number of publications in any given interval of time has also a negative binomial distribution. Negative binomial distribution has been shown to provide good fit to observed citational distributions and obsolescence model. The model used by Allison et al in its modified version does not specify an exact probability distribution of the cumulative number of citations that a scientist receives in a specific time interval. It gives a significant result that a modified co-efficient of variation specifies the measure of inequality. Value of this co-efficient is shown to be a constant over time. Thus, their conclusion was that the CAP contributes to the scattering in various levels that does not change the character of inequality but proportion of inequality over time.

This result is confirming to the observations and theoretical results of the law of constant accessibility of information and as such of a constant distributional feature of citations over time.

They gave the modified co-efficient of variation, $C$ as the ratio of the difference between variation of the random variable $X(t)$ of the cumulative number of citations and the expectation or mean of $X(t)$, to the square of this expectation or mean, i.e.,

$$C = (\text{variation} - \text{mean}) / (\text{mean})^2$$
$$= b/a \text{ (say)}$$

$C$ gives the degree of inequality. They also showed that the *Reterogeneity* in the rate of cumulative advantage implies increasing inequality.

We have not shown the mathematical details involved. We have not also grouped this with other models of Zipfian or skewed distributions and empirical formulations.

This work is one of the earliest without referring to any empirical law and which derives something directly related to the properties of citational inequality but not citational scattering distribution persè. This model stresses on the process itself.

In the next we shall just mention the scattering models without describing any of them in detail.

**Citational Scattering and CAP**

Cumulative advantage process (CAP) (Price, 1976) has been variously called as Mathew-effect (Merton, 1973), Gibrat's law (Ijiri and Simon, 1975). They all represent different types of skewed distributions which Haitun (1982) has named as Zipfian distributions. They are also represented by 80:20 or in general P:(100-P) rules of proportion. All of them represent generically the processes where items are concentrated over a few number of sources. There will be huge number of sources with only one item each (usually more than 50%). These distributions are called scattering in case of bibliometrics. The studies on them usually ignore (or can not account for) the empty or zero

item sources (see for discussion Sen, Chatterjee, Gan 1992).

CAP is almost universally accepted to represent social processes which lead to concentration of items among sources. It is said and empirically observed that productivity of authors, journal scattering, word frequencies and citations to past documents are all manifestations of CAP. A theory of citation process, it is felt, should lead to an empirical vindication of citational scattering. There are a number of attempts for theorising CAP or scattering and many of them claim to be explaining citational scattering as well, but most of them address themselves to other types of scattering phenomena rather than that of citation.

In dealing with the phenomenon of scattering or skewed distributions, the sources are ranked according to productivity i.e. the number of items in each source. Thus scattering distributions are usually rank-frequency distributions (in general the major statistical distributions are class-frequency or size frequency distributions.

Bibliographic scattering is a term which applies to scattering or skewed distribution of bibliographic items among bibliographic sources. The most prominent and extensively studied are the phenomenon of journal scattering. This type of scattering is formally described by two empirical observations called Bradford's laws. In this case journals or periodicals are bibliographic sources and the articles or research papers in the journals are items.

The second class of bibliographic scattering is the citational scattering. Here, the bibliographic sources are the cited items i.e. references and the items are the citing items or the citations. Some times however cited authors, cited journals, cited institutions may be taken as sources.

It is simply assumed that a single explanation or a single theoretical model should be applicable for explaining both journal scattering and citational scattering. But only a few authors have this explicitly stated and rarely the theoretical models are applied for empirical verification of both types of scattering.

There are many theoretical models and explanation of the bibliographic scattering phenomena. Most of them are primarily for journal scattering. Many of them also deal with author productivity or word frequency distributions. Haitun (1982) made a long list of these formulations upto 1981. We may add a few more which have appeared afterwards ( Egghe and Rousseau 1990, Maia and Maia 1984, Burrell 1988, 1988A; Sichel 1985, 1986; Sen 1989; Sen et al 1992; Basu 1992, 1993 etc.).

We have already noted that in this discourse we do not describe the theoretical models of scattering studies unless the model is so generalized as to cover other aspects of citation phenomena. Reviewing the literature on scattering would mean an independent long term research project.

Among others Sen (1989) and Sen et al (1992) specifically mention citational scattering. In Sen (1989), suggestions for empirical vindication of the model in citational scattering has been mooted. In Sen et al (1992) empirical verification of citational data is also included.

None of the authors, to our knowledge, has so far pointed out a distinctive characteristics of citational scattering from the other types of scattering. Citational scattering is much more randomized than any other type is.

**Discussion**

We have discussed in this review, the properties and primary characteristics of references, referencing and citation processes. We have seen that the attempted theories vary widely in status and level of sophistication. The pre-event and post-event classification is a novel finding of this study.The main property of reference size and constancy of reference / citation ratio seems to be vindicated by most of the studies.

Among the approaches presented, order out of chaos, implicit or normative principles, cognitive functions, logical dimensions, holographic and maximum speed are just principles and general looks out. Evolutionary theory and Vinkler's theory are two important approaches. Both are sociological, but Vinkler's theory has also semiotic elements. Typology of referencing and citer motivation is not the topic of this review and has already been represented in more or less comprehensive manner by others. The citing process complex again highlights the whole pre-event process and puts pre-event citation process in the background of general information seeking behaviour of the scientists and production of scientific literature.

The theorems reproduced are all formal attempts but are sporadic attempts. A general framework is yet to be developed here. Proofs of the theorems have not been provided because of the fact that they would have made the presentation more complex and mathematically involved.

Models of Glänzel, Glänzel and Schubert, Gomperts, Sharma et al and Van Raan are semi-empirical. But the approach by Glänzel using the concept of stopping time is a pointer for future formal post-event theoretical models. To do justice to citational distribution or scattering and its relation to CAP, one needs to review the very large mass of literature on bibliographic scattering and bibliometric laws. This was not within the scope of this study. The literature on scattering has been well reviewed time to time.

In the next sections we try to assess the reactions to citation indexing and possible suggestions for alternatives, to resolve the debates and indicate a future course of action.

## Reactions and Alternatives to Citation Analyses

In May 1991, *Science* published a 'news and comment' on the Royal Society's report on reactions of British academicians to citation indexing and its role in value judgement (Anderson, 1991). Some of the reactions were that CI is "pseudoscience", "fallacy, totally mistaken", "the refuge of the Philistines", the citation analysis is flawed to the point of being both misleading and inherently absurd.

These reactions and criticisms to an American enterprise or device which may generate a publication rat race, may malign the whole process by maligning the system. Unless one can understand the theoretical backgrounds and bases, controversies would pile up. And they have piled up already. No effective precautions can be taken in analyses and comparative studies unless the theoretical issues are settled and theoretical infrastructures are provided for.

The British academicians who shun the citation indexing and analyses based on citation indexing vie for their alternatives.

The alternative to competition via mechanised device like citation index is humanized peer reviewing process. Individual peer reviewing is again prone to subjective considerations. But the total enterprise of peer reviewing however acts as a normative system. We will see that either peer reviewing or citation analysis presumes an ideal story book image of science and scientific scholarship. Peer-reviewing is not mechanical and can not give quick readymade results. Citation indexing is completely mechanical, rather mechanically blind - than what is required in assessing scientific endeavours and preparing the infrastructural outline of history of science.

There has been no exhaustive comparisons between the citation indexing, citation analyses on one hand and the peer reviewing system on the other. That can again be a major research project.

The one comparative study with Dutch science (Moed et al, 1985) 'reveals a serious lack of agreement' between the citational indicators and the peer review process for value judgement of scientific research. The study raised many relevant serious questions, answers to which are not known to us.

Another study by Folly et al (1981) on Hungarian science tried to show that the citing authors follow the same or similar criteria of selecting papers for reference as the peers (and editors) use in their refereeing and publication process.

There can be one positive argument in favour of citation indexing. Citation indexes are prepared from the references of the published items (documents). These documents can only be published after they are filtered through refereeing and editorial screening. After publication again the items are subjected to peer review and value judgements. Only after such testing some of them appear as references. Thus citation indexing can be considered as a second or third level outcome of peer-review process. No one has so far explicitly stated this.

Question then arises whether citation analyses based on citation indexes would suffice or whether separate, independent, overall new peer judgements will be needed for value judgement of authors, papers, institutions and for conferring rewards. The processes of refereeing and peer reviewing and their reliabilities have recently been studied in two different contexts (Sen and Chakrabarty, 1993; Daniel 1993). In the second study a comparative analysis

was made of papers published by a reputed journal and those rejected but published elsewhere. Citations to these two groups showed significant difference (Daniel 1993). But we have seen that many of citation classics selected by ISI were primarily rejected by referees and editors (Sen, unpublished work).

This issue cannot be easily resolved. The shortcomings and benefits of citation analyses based on citation indexing have been discussed in detail in literature, but not about peer reviewing. The strength and weakness of citation analyses are due to the same reason — the mechanical way of producing the indexes. Another reason is the selection of source documents. Many observers think that by changing the source document (or source journal) population, the resultant citation scenario would change significantly. An all-American source document population would produce a citational map or scenario completely different from one produced by taking all source documents from Indian origin only. Such a proposition or hypothesis can not be easily proved or disproved. Moreover, the citationists would say that the really best or important works would be definitely and prominently represented in both the cases. Mismatches will occur mostly in average and below average cases.

These discussions just show, despite numerous studies of reference analyses and citation analyses and criticisms and polemics we have not done much worth while work to clear up the real messy corners.

## Ortega Hypothesis versus Elite Hypothesis

Ortega (1932) made a claim that for progress of science, contributions of average (non-elite) scientists have a credible share. In the words of MacRoberts and MacRoberts (Scientometrics, 1986) the research of average scientists contributes substantially to the advancement of science. Cole and Cole (1972) on the other hand concluded that only a small elite group is responsible for progress of science. Through citation studies it becomes evident that the citation chain or citation sequence connects only a small proportion of scientists at any instance. This proposition by Cole and Cole may be called elite hypothesis.

It is a fact that a seminal paper or a revolutionary paper influences generations of papers for a long time to come. Unless the basic premise or main results of the paper is somehow debased, the citations to original papers go on for some time. The citation cycle it generates continues for a pretty long time. The spurt of papers suddenly falls if the main paper is debased. We can cite examples of claimed discovery of poly-water or cold-fusion in this regard.

The papers thus produced on the trail of a blazing one refer to the core papers, but they themselves are rarely cited. Some of them may become more important and attract more citations than others. The proponents of elite hypothesis claim that these papers only produce future seminal papers on their turn. In the cognitive flow of scientific progress, influence of the average and peripheral papers is negligible. Supporters of Ortega hypothesis feel that although these papers are not explicitly cited they play a major role in scientific progress through silent influences.

There are papers which cite others but are not cited themselves. Folly et al (1981) on their study on citation of Hungarian scientists designated three types of citations, namely, self citation; co-operational citation and independent citation. They defined these three concepts respectively as : (i) self citation - if the set of cited authors has elements common with the set of citing authors; (ii) co-operational citation - when the cited author under study and one citing author were co-authors in a previous paper prior to the publication of the cited papers and the criterion of self citation does not hold; (iii) independent citation occurs when no detectable relation between cited and citing authors are found. They consider that real significant influence can only be discerned by counting the number of independent citations only. They took one citation per year of type three as the level of significance, because the estimated value of average citations per cited item taking all the items in the ISI database is 1.7 citations per year. They have also made an assumption that peers or the referees judge the quality of scientific work by the same criteria as citing authors do. They found out that in case of authors with significant level of influence the percentage of type three citation varies from 48 to 78

with an average of 70 among all the three types of citations. Such a study only corroborates the views in favour of elite hypothesis. Self citations and corporate citations come around 30%.

Yet there is a significant perspective which should be considered. In creative literature (there is no system of footnotes or referencing in writing fictions, poems etc) also, there are few elite authors who influence the average authors to copy them or write like them. These average authors are never recognized as major authors themselves. A major or a great author in most cases digest small influences of many average authors on his part and produces a signficant work. It is not easy to find out whether such indirect influences come from the average and nominal authors in science as well. But a strong possibility is there. The most of the uncited works, however, play some other roles in the flow of information. They describe a large number of minor applications of an influential paper. These papers serve also as exercises for disserting, using the ideas and propagading the culture far and wide. These authors also help to create a mass-base of scientific activities.

The suggestion that partcipation in scientific research and in the publication game should be restricted without any chance of hampering the normal progress of science need to be viewed from another side.

Performance (and activity) in science is like a cone or inverted truncated paraboloid with a large base and a narrow tip. The significant authors are in top region. They are supported by the mass below them. If one follows Lotka's law for citation scores also then a parabolic (inverse power law) situation is evident. If the base is smaller (smaller area of the cone) height of the cone will be less and the top will also be smaller and narrower and of less height. With sufficient decrease in the number of grass-root or base level scientific workers and nominal authors ('massses' of Ortega) the whole system would become unstable and significant research would be less and less. It is not difficult to build up a quantitative model in this line but here it is no place to do that (that would be done elsewhere).

The problem of controversy with the issue arises because of looking at the issue from two different angles. If one considers citation chains, citation cycles, institutional traditions, teacher - student continuum — elite hypothesis is upheld. If the whole enterprise is looked at a global scale and the dissemination of scientific knowledge and making of scientists in every culture is considered, Ortega hypothesis is upheld.

Citation process discloses the elite hypothesis and abhors the Ortega hypothesis. It manifests the participaion of masses but shows the flow of inheritance through an elite core (a selected few). This has a parallelism in biology. Among the sparrows and some other species of birds, all the males participate in the pre-mating game, but only a few can ultimately mate and become fathers.

**Memetics and Citation Process**

Richard Dawkins (1976) drew attention to the roles played by some stable or key concepts in cultural evolution similar to those played by genes in biological evolution. Dawkins observed that ideas flow culturally in time and space (flow through space involves time, even if small). He said that ideas have their basic units, and these units, like the units of heredity, "the genes" are carriers of cultural inheritence.

As genes are biologically transmitted, and biological evolution proceeds through gene invariance and gene mutation, so are these idea genes transmitted culturally being invariant mostly and mutating gradually. He calls these "idea genes" as "memes".

"If a meme is to dominate the attention of a human brain, it must do so at the expense of rival memes. Other commodities for which memes compete are radio and television time, bill-board spaces, newspaper column inches and library shelf-space".

Citation counts do not reflect whether the ideas contained in a paper are rejected or accepted or partly utilized. Whenever an idea is universally accepted or becomes embodied in the current thought its generating papers are not cited at all. Whereas a paper whose ideas have been rejected after some debate may also show the same sort of obsolescence on citation counts. In such measures of obsolescence we are ignorant of the actual life of an idea.

The process of the evolution of culture, especially that of the cumulative growth of S&T, is effected through certain stable concepts and ideas. Stabilisation of the concepts (or ideas) happens through use and recognition of those concepts by others. Such culturally stable and potent concepts are memes. Newer ideas are generated from older ones as if by mutation. Evolution of growth of S&T is the process of a flow of ideas (in respect of mentefacts and artefacts) from generation to generation, the process of looking ahead by standing on the shoulders of previous giants, as Newton said. The words "previous giants" are significant. The evoluion of ideas on the process of culture does not depend upon all the ideas or concepts of the past, but almost solely on a small, stable and sturdy part of the ideas or concepts produced or available at a given time (elite hypothesis).

This evolutionary growth is best manifested through the growth of recorded thought, i.e. documents, and is recognised socially, as well as academically, through the tradition of citation. Citations and bibliographic references carry out yet another function; they help provide readers or information seekers with an indication of related ideas and documents.

By considering such an analogy of the evolutionary process at the biological and cultural levels, we may say that the documents may be taken as species populations, and the core ideas or stable concepts as memes. Whenever citation occurs, it is implied in an abstract sense that the ideas in the citing documents are inherited from the "meme pools" of the cited documents. Such inheritance or influence may, however, be of various kinds. Indeed, citations do occur occasionally to represent mutation or regression of the meme pools of the cited documents.

Except in some highly skewed attempts in content analysis, there is no formal or standardised tool or procedure for assessing or enumerating the cognitive content of a document or written or non-written texts of one's expression. Instead of cognitive content, we prefer to use the term semantic content or semantic value. When we talk of ideas or concepts or memes, we are concerned with semantic elements or cognitive content. The influence of ideas (flow of information) depends upon cognition of the content, using the semantic elements. As there is no way, so far, to deal with the units of literature or textual elements of the documents in case of S&T literature, especially research literature, we assume that the relation of the content of a unit or item of literature or text to the item itself is straight forward. So, whenever we think of the relationship between two documents we think of that between the thought contents or semantic contents of the two.

## Conclusion

"The wheel was invented much before there were any theory of rotation. If the wheel moves and the cart moves, no one cares for its theory. But it has been different with citation indexing. From the beginning, every one wanted to understand the underlying theoretical principles. More important, there was suddenly a new social phnomenon — the citation cycle. All these needed explanation and formalization.We have seen, whatever the reactions to citation analyses and underlying citation procss, citational activities follow the same ideal norms as those followd by peer reviewing. Again both are prone to distortions and deviations from these norms. Indeed, referencing may be considered as a third level outcome of peer-reviewing or judgements by peers and colleagues."

The *Ortega-elite debate* is easily resolved, as has been shown. The main outcome of this review is the insight that the core flow of citation chains really represents cognitive links and the cumulation of scientific ideas. We can also say that the citation chains can be followed as meme flows.This puts citation process as a part of the scholarly cultural evolution. Any holistic theory of citation process should probably be a model based on memetics and showing the citational process should probably be a model based on memetics and showing the citational scattering as final results. This can probably be done by considering citation as a Markov chain process.

It should be mentioned here that we have also left out another important aspect. That is the relation between theoretical approaches and the citational indicators. Citational indicators are now many in number and are being used for many pur-

poses, especially for value judgements. These indicators are on the main arbitrary and artificially produced. Except in one or two cases theoretical studies do not or cannot produce any of such indicators. Indicators suggested by a theory may be quite different from any indicator being used. The major problem is that the indicators which may be suggested as outcomes of theoretical studies may be too complex to be easily manipulated or enumerated. On the other hand the indicators in use are in most cases too simplistic and hence may be questionable. This is definitely an issue for future theoretical work on citation processes.

**Acknowledgements**

This review had never been initiated but for a request from Prof. Allen Kent for writing an article on Citation Process Theory for the prestigious Encyclopedia of Library and Information Science (ELIS). I could not do it in time for inevitable reasons. Writing of the review was ultimatly possible through a partial financial support from the National Information System for Science and Technology (NISSAT), Department of Scientific and Industrial Research (DSIR), Government of India through its Grant No. NI/CB/103/92 of March, 1992. Many friends, colleagues, students helped in various ways, listing of individual names is not possible here.

## References

1. AMSTERDAMSKA, O and L. Leydesdorff. (1989),Indicators of significance. *Scientometrics,* **15**(5-6) 449-472.
2. ALLEN, T. J. and P. G. Gerstberger. (1967) Criteria for selection of an information source. Cambridge : Massachusetts Institute of Technology.
3. ALLISON,Paul., D. J. Scott Long, T. K. Krauze.(1982) Cumulative advantage and inequality in science. American Sociological Review, **47**, 615-625.
4. ANDERSON, Alan (1991), No citation analyses please, we are British. *Science* (1 May) 639.
5. ARTUS, H. M. (1993) Wrong interpretation of wrong data, reference analysis and the obsolescence of scientific literature. *4th International Conference on Bibliometrics, Informetrics and Scientometrics*, Berlin, 11-15 Sept. 1993.
6. ARTUS, H. M. (1992) Caraue literatur : Zum Informellen Kommunikations System der Sozial Wissenschaften [grey literature in the informal communication system of the social science] (unpublished research report), **IZS**, Bonn
7. AVRAMSESCU, A. (1979) Actuality and obsolescence of scientific literature. *JASIS*, **30**, 296-303.
8. AVRAMESCU, A.(1973) Science and obsolescence. *St. Cere. Doc,* **15,** 345.-365.
9. BARTHOLOMEW, D. J.(1982) Stochastic models for social processes. New York : Wiley & Sons.
10. BASU, Aparna.(1992) On the theoretical foundations of Bradford's law. (in *Informetrics*) 91 ed by I. K. Ravichandra Rao, Bangalore, Sarada Ranganathan Endowment for Library Science) 51-69.
11. BAYER, A. E. and G. Jahoda.(1981) Effects of online bibliographic searching on scientists' information style. *Online Review*, **5,** 323-333.
12. BONZI, Susan and H. W. Snyder. (1991) Motivations for citation : a comparison of self citation and citation to others. *Scientometrics*, **21**(2), 245-254.
13. BONZI, S. (1982) Characteristics of a literature as predictors o relatedness between cited and citing works. *JASIS*, **33**(4), 208-216.
14. BROOKS, T. A.(1987) Citer motivations. (in : *ELIS*, Vol. 42. Edited by Allen Kent. New York : Marcel Dekker, ) 48-59.
15. BROOKS, T. A.(1985) Private acts and public objects : an investigation of citer motivations. *JASIS*, **36** , 223-229.
16. BUCKLAND, M. K. (1972) Are obsolescence and scattering related? *Journal of Documentation*, **28**(3), 242-246.

17. BURREL, Q. L. (1988) A simple empirical method for predicting library circulation. *Journal of Documentation*, **44**(4), 302-314.

18. BURREL, Q. L. (1989) On the growth of bibliographics with time : an exercise in bibliometric prediction. *Journal of Documentation*, **45**(4), Dec. , 302-317.

19. CAIRNS, R. B. (1979) Social development : the origin and plasticity of interchanges. San Francisco : W. H. Freeman & Co.

20. CANO, V (1989) Citation behaviour : classification, utility and location. *JASIS*, **40**(4), 284-290.

21. CHATTERJEE, Satish Chandra. (1978) The Nyaya theory of knowledge. Calcutta : Calcutta University.

22. CHUBIN, D. E. and S. D. Mitra. (1975) Content analysis of references : adjunct or alternative to citation counting? *Social Studies of Science*, **5** 423-441.

23. COADIC, Y. F. Le. (1979) The heuristic diffusion of scientific ideas : a mathematical and sociological approach. *International Forum of Information and Documentation*, **4**(2), 7-11.

24. COLE, J. and S. Cole. (1972) The Ortega hypothesis. *Science*, **178**, 368.

25. COLEMAN, S. R. (1992) The laboratory as a productivity and citation unit in the publications of an experimental-psychology specialty. *JASIS*, **43**(9), 639-643.

26. COZZENS, S.(1982) Split citation identity : a case study from economics. *JASIS*, **33**(4), 233-236

27. COZZENS, S.(1989) What do citations count? The rhetoric first model. *Scientometrics*, **15**, 437-447.

28. CRONIN, Blaise.(1980) Some reflections on citation habits in psychology. *Journal of Information Science*, **2**, 309-311.

29. CRONIN, Blaise.(1981) Agreement and divergence on referencing practice. *Journal of Information Science*, **3**, 27-33.

30. CRONIN, Blaise.(1991)Let the credits roll : a preliminary examination of the role played by mentors and trusted assessors in disciplinary formation. *Journal of Documentation*, **47**(3), 227-239.

31. CRONIN, Blaise. (1984) The citation process : the role and significance of citatious in scientific communication. London : Taylor Graham.

32. CRONIN, Blaise and Gail Mckenzie.(1992) The trajectory of rejection. *Journal of Documentation*, **48**(3), Sept. 310-317.

33. CULNAN, M. J.(1985) The dimensions of perceived accessibility to information : implications for the delivery of information systems and services. *JASIS*, **36**, 302-308.

34. DANIEL, H. D.(1993) Guardians of science; fairness and reliability of peer review. Weinheim : VCH.

35. DAWKINS, R. (1976) The selfish gene. Oxford : Oxford University Press.

36. DUNCAN, E. B. et al.(1981) Qualified citation indexing : its relevance to educational technology. (In Duncan, E. D. and McAleese, R.(eds.). Information retrieval in educational technology. *Proceedings of the First Symposium on Information Retrieval in Educational Technology*, 1st April 1981, Aberdeen. Aberdeen : University of Aberdeen) 70-79.

37. EDGE, D. O. (1977) Why I am not a co-citationist. *Society for the Social Studies of Science Newsletter*, **2**, 13-19.

38. EDGE, D. O. (1968) Quantitative measures of communication in science : a critical review. *History of Science*, **17**, 102-134.

39. EGGHE, L. and I. K. Ravichandra Rao. (1992) Citation age data and the obsolescence function : fits and explanations. *Information Processing & Management*, **28**(2), 201-217,

40. EGGHE, L. AND R. Rousseau. (1990)Eds. Informetrics 89/90. (*Proc. 2nd International Conference on Bibliometrics, Scientometrics and INFOR-METRICS*. University of Western Ontario, Canada 1989). Amsterdam : Elsevier.

41. EGGHE, L. and R. Rousseau. (1990) Introduction to informetrics. Amsterdam : Elsevier.

42. ENGER, Kathy B., Georgia Quirk and J. Andrew Stewart. (1989) Statistical methods used by authors of library and information science journal articles. *LISR*. **11**, 37-46.

43. FAIRTHORNE, R. A.(1969) Emperical hyperbolic distributions (Bradford-Zepf-Mandelbrot) for bibliometric description and prediction. *Journal of Documentation*, **25**, 319.

44. FERLIGOJ, Anuska and others. (1988) Cluster analysis of citation histories from an institutional setting. *Scientia Yogoslavica* **14**(3-4), 159-169.

45. FINNEY, B. (1979) The reference characteristics of scientific texts. London : Centre for Information Science (MSc Dissertation).

46. FOLLY, G. B., Hajtman, J. I. Nagy and I. Ruff. (1981) Some methodological problems in ranking scientists by citation analysis. *Scientometrics* **3**(2), 135-147.

47. FROST, C.O. (1979 ) The use of citations in literary research : a preliminary classification of citation functions. *Library Quarterly*, **49**, p. 399.

48. GARFIELD, E. (1955) Citation indexes for science. *Science*, **122** (3159), 108-111.

49. GARFIELD, E. (1964) Science citation index : a new dimension in indexing. *Science*, **144** (3619), 649-654.

50. GARFIELD, E. (1980) Is citation indexing in the arts and humanities inherently different from that in science? The effects that ISI's Citation Index for the Arts and Humanities is expected to have on future scholarship. *Library Quarterly*, **50**(1), 40-57.

51. GARFIELD, E. (1979) Citation indexing : its theory and application in science, technology and humanities. New York : Wiley.

52. GARFIELD, E. ( 1977) Can Citation Indexing be automated? (In Essays of an Information Scientist, Vol. 1), Philadelphia : ISI Press.

53. GARVEY, W. G. and others. (1974) The dynamic scientific-information user. *Information Storage and Retrieval*, **10**, 105-131.

54. GARVEY, W. D., N. Lin, and C. E. Nelson. (1970) Communication in the physical and social sciences. *Science*, **170** (December 11), 1166-1173.

55. GELLER, N. L. and J. S. DECANI. (1981) Lifetime-citation rate : a mathematical model to compare scientists work. *JASIS*, **32**, 3-15.

56. GILBERT, G. N. (1977) Referencing as persuasion. *Social Studies of Science*, **7**, 113-122.

57. GOMPERTS, M. C. ( 1968) The law of constant citation for scientific literature. *Journal of Documentation*, **24**(2), June.

58. GORDON, M. D. (1979) Deficiencies of scientific information and access in less developed countries. *JASIS*, **36**(3), 340-342.

59. GRIFFITH, B. C., P. N. Servi, A. L. Anker and M. C. Drott. (1979) The aging of scientific literature : a citation analysis. *Journal of Documentation*, **35**(3), 179-196.

60. HAITUN, S. D. (1982) Stationary Scientific distributions, Parts I, II, III Scientometrics 4(1,2,3) 5 - 25, 89-104, 181-194.

61. HARDY, A. P. (1982) The selection of channels when seeking information : cost-benefit vs. least effort. *Information Processing and Management*, **18**, 289-293

62. HODGES, T. (1978) Citation indexing : its potential for bibliographic control. Berkeley : University of California (Ph. D. thesis).

63. IJIRI, Y. and H. A. Simon. (1975), Some distributions associated with Bose-Einstein statistics. *Proceedings of National Academy of Sciences of the U. S. A.*, **72.** 1654.

64. JOHANSON, C. M. A. ( 1976) Citation behaviour in chemistry. London : City University, Centre for Information Science. (M. Sc. dissertation).

65. KAPLAN, N. (1965) The norms of citation behaviour : prolegomena to the footnote. *American Documentation*, **16**(3), 179-184.

66. KIDD, J. S. (1990) Measuring referencing practices. *JASIS*, **41**(3), 157-163.

67. KOCHEN, M. (1974) Principles of information retrieval. Los Angeles : Melville Pub.

68. KOCHEN, M. (1987) How well do we acknowledge intellectual debts, *Journal of Documentation*, **43**(1), 54-64.

69. KRAUZE, T. K., Hillinger, C (1971). Citations, references and the growth of scientific literature : a model of dynamic interaction. *JASIS*, **21**, 333-336.

70. LANCASTER, F. W. and J. Li. (1989) The law of constant acccessibility of information. *Bulletin of the American Society for Information Science*, **15**(2), 24-25.

71. LAWANI, S. M. (1982) On the heterogeneity and classification of author self-citations, *JASIS*, **33**, 281.

72. LEYDESDORFF, L. (1987) Towards a theory of citation? [A comment on "Testing the Ortega hypothesis : facts and artifacts" by M. H. Macroberts, B. R. Macroberts, *Scientometrics*, **12**, 1987, 293-295]. *Scientometrics*, **12**(5-6), 305-309.

73. LEYDESDORFF, L. (1989) The relations between qualitative theory and scientometric methods in science and technology studies. *Scientometrics*, **15**(5-6), 333-347.

74. LIPERTZ, B. A. (1965) Improvement of the selectivity of citation indexes to science literature through the inclusion of citation relationship indicators. *American Documentation*, **16**(2), 81-90.

75. MacROBERTS, Michael, H. and Barbara R. Macroberts. (1986) Quantitative measures of communication in science : a study of the formal level. *Social Studies of Science*, **16**, 151-172.

76. MacROBERTS, M. H. and B. R. MacRoberts. (1987) Testing the Ortega hypothesis : facts and artifacts. *Scientometrics*, **12**(5-6), 293-295.

77. MacROBERTS, Michact H. and Barbara R. MacRoberts. (1989) Problems of citation analysis : a critical review. *JASIS*, **40**(5), 342-49.

78. MacROBERTS, M. H. and B. R. MacRoberts. (1987) Another test of the normative theory of citing. *JASIS*, **38**, 305-306.

79. MAIA, M. J. F. and M. D. Maia. (1984) On the unity of Bradford's law. *Journal of Documentation*, **40**, 206-216.

80. MARICIC, S. (1983) A fourfold bibliometric evaluation of journals from less developed scientific communities : the pattern of references in medical journals from Yogoslovia. *Annal Lib Sci Doc.* **30**(2), 45-62.

81. MARGOLIS, J. (1967) Citation indexing and evaluation of scientific papers. *Science*, **155** (3767), 1213-19.

82. MARTYN, J. (1987) Literature searching habits and attitudes of research scientists. Boston Spa : British Library.

83. MATRICCIANI, E. (1991) The probability distribution of the age of references in engineering papers. *IEEE Transactions of Professional communication*, **34**, 7-12.

84. McINNIS, Raymond G. and Dal Symes. (1988) David Riesman and the concept of bibliographic citation. Bellingham : Western Washington University.

85. MEADOWS, A. J. (1974) Communication in science. London : Butterworths.

86. MERTON, R. K. (1973) The Mathew effect in science (in Sociology of science. Chicago : University of Chicago), 439-450.

87. MILLER, G. A. (1956) The magical number seven, plus or minus two : some limits on our capacity for processing information. ; *Psychological Review*, 63, 981-7.

88. MOED, H. F., W. J. M. Burger, J. G. Frankfort, A. F. J. Van Raan. (1985) A comparative study of bibliometric past performance analysis and peer judgement. *Scientometrics*, **8**(3/4), 149-159.

89. MORAVCSIK, Michael J. (1988) Citation 'context classification of a citation classic concerning citation context classification. *Social Studies of Science*, **18**, 515-21.

90. MORAVCSIK, M. J. and P. Murugesan.(1975) Some results on the function and quality of citations. *Social Studies of Science*, **5**, 86-92.

91. MOTYLEV, V. M. (1981) Study into the stochastic process of change in the literature citation pattern and possible approaches to literature obsolescence estimation. *International Forum on Information and Documentation*, **6**(2), 3-12.

92. MOTYLEV, V. M. (1989) The main problems of studying literature aging. *Scientometrics*, **15**, 97-109.

93. MULKAY, M. (1974) Methodology in the sociology of science : some reflections on the study of radio astronomy. *Social Science Information*, **13**, 109-119.

94. NAKAMOTO, Hideshiro. (1965) The obsolescence model of scientific literature (in Japanese). *Dokmentesyon-Kenkyu*, **16**(1), 16-20.

95. NAKAMOTO, Hideshiro. (1964) Obsolescence of literature (in Japanese). *Dokumentesyon-Kenkyu*, **14**(6), 123-28.

96. NEDERHOF, A. J. (1988) The validity and reliability of evaluation of scholarly performance (in Hand Book of quantitative studies of science and technology ed. by A. F. J. Van Raan. Amsterdam : Elsevier Science Publisher), Chapter 7.

97. NEDERHOF, A. J. and A. F. J. Van Raan. (1986) Citation theory and Ortega hypothesis. *Scientometrics*, **12**, 325-328.

98. OPPENHEIM, C. and S. P. Renn. (1978) Highly cited old papers and the reasons why they continue to be cited. *JASIS*, **29**(5), 225-231.

99. ORR, R. H. (1970) The scientist as an information processor : a conceptual model illustrated with data on variables related to library utilization. (In : *Communication Among Scientists and Engineer;* ed. by C. Nelson and D. K. Pollack, Lexington, Mass, Health), 143-189.

100. ORTEGA, Y.and J. Gasset. (1932) The result of the masses. New York : W. W. Norton.

101. PERITZ, Bluma C. (1983) A classification of citation roles for the social science and related fields. *Scientometrics*, **5**(5), 303-312.

102. PERITZ, B. C. (1983) Are methodological papers more cited than theoretical or empirical ones? The case of sociology. *Scientometrics*, **5**, 211-218.

103. PERITZ, B. C. (1990) A citation analysis of clinical trials : are definitive studies less cited than others (in Informetrics, 89/90 ed. by L. Egghe and R. Rousseau. Amsterdam : Elsevier) 243-250.

104. PRICE, D. J. Solla. (1976) A generalised theory of bibliometric and other cumulative advantage processes. *JASIS*, **27**(5), 292-306.

105. PRICE, D. J. Solla. (1970) Citation measures of hard science, soft science, technology and nonscience (in Nelson, C. E. and D. K. Pollack. Communication among scientists and engineers, Lexington), Mass : Health, 1-22.

106. PRICE, D. J. Solla. (1986) Little science, big science and beyond. N. Y. : Columbia Univ. Press.

107. POOLE, H. H. (1985) Theories of the middle range. Norwood, N. J. : Ablex Publishing.

108. RAFSNIDER, JEAN (1975). Complementarity between a use study and a citation study in considering publication practices in the physical and socil sciences. *JASIS*, Nov-Dec. 344-345.

109. RAVICHANDRA RAO, I. K. (1988) Probability distributions and inequality measures for analysis of circular data (in L. Egghe and R. Rousseau (eds.), *Informetrics*, 87/88, Proceedings of First International Conference on Bibliometrics and Theoretical Aspects of Information Retrieval, LUC, Belgium, 1987. Amsterdam : Elsevier), 231-248.

110. RICHARDS, J. M. Jr. (1991) Years cited an alternative measure of scientific accomplishment, Scientometrics **20**(3), 427-438.

111. ROSENBERG, V. (1966) The application of psychometric techniques to determine the attitudes of individual toward information seeking. Bethlehem, Pa : Leigh University, Center for Information Science.

112. ROUSSEAU, Ronald. (1992) Why am I not cited or why are multi-authored papers more cited than others? *Journal of Documentation*, **48**(1), 9-80.

113. ROUSSEEUW, P. (1991) Why the wrong paper gets published. *Chance*, **4**(1), 41-43.

114. SEGLEN, P. O. (1992) The skewness of science, *JASIS*, **43**(9), 628-638.

115. SEN, B. K. and K. Shailendra. (1992) Evaluation of recent scientific research output by a bibliometric method. *Scientometrics*, **23**(1), 31-6.

116. SEN, S. K. (1990) A theoretical glance at citation process. *IFID*, **15**(1), 1-7.

117. SEN S. K. (1981) A note on the idea gene and its relation to information science. *Annals Lib. Sci. Doc.*, **28**(1-4).

118. SEN, S. K. ABSCINDEX (1988) : Proposal of a new tool for information retrieval and science indication in regional contexts. *Annals Lib. Sci. Doc.*, **35**, 116-123.

119. SEN, S. K. and S. K. Gan. (1983) A mathematical extension of the idea of bibliographic coupling and its application. *Annals Lib. Sci. Doc.*, **30**, 78-82.

120. SEN, S. K. and Sunil Kumar Chatterjee. (1990) An introduction to researches in bibliometrics : Part II. Glossary. *IASLIC Bulletin*, **35**(4), 155-170.

121. SEN, S. K. and S. K. Chakrabarty. (1993) Quality control in Indian Science : a study on refereeing systems of twenty six Indian journals. *JSIR(India)*, **52**(3), 133-150.

122. SEN, S. K. and Sunil Kumar Chatterjee and Shyamal K. Gan. (1992) Making bibliographic scattering studies useful : significance of zero-item (empty) and single-item sources. Third International Conference on Informetrics, Bangalore, 9-12 Aug, 1991. (In Informetrics '91 ed. by I. K. Ravichandra Rao), 426-446.

123. SHARMA, K. D. L. Singhal and B. M. Gupta. (1980) Electronic lattice scattering technique for citation analysis and evaluating the impact of research papers. *Annals Lib. Sci. Doc.* **27**(1-4), 52-60.

124. SIEVERT, Mary Ellen and Mary Haughawout. (1989) An editor's influence on citation patterns : a case study of elementary school journal. *JASIS*, **40**(5), 334-41.

125. SMALL, H. G. (1978) Cited documents as concept symbols. *Social Studies of Science*, **8**, 327-340.

126. SNYDER, H. Q. and S. Bonzi. (1989) An enquiry into the behaviour of author self citation. ASIS'89 : *Proceedings of the 52nd Annual Meeting of the American Society for Informa-*

*tion Science*, Medford, New Jersey, Learned Information, 147-51.

127. SO, C. Y. K. (1990) Openness index and affinity index : two new citation indicators. *Scientometrics*, **19**(1-2), 25-34.
128. SOPER, M. E. (1976) Characteristics and use of personal collections. *Library Quarterly*, **46**.
129. SPIEGEL-ROSING, I. (1977) Science studies : Bibliometric and content analysis. *Social Studies of Science*, **7**, 9, 7-113.
130. STERN, Richard. E. (1990) Uncitedness in the biomedical literature. *JASIS*, **41**(3), 193-196.
131. STINSON, E. E. and F. W. Lancaster (1987). Synchronous versus diachronous methods in the measurement of obsolescence by citation studies. *Journal of Information Science*, **13**, 65-74.
132. SUMMERS, E. G. and others. (1983) The effect of personal, professional and psychological attributes, and information seeking behavior on the use of information sources by educators. *JASIS*, **34**, 75-85.
133. SWANSON, D. R. (1987) Two medical literatures that are logicaly but not bibliographically connected. *JASIS*, **38**, 228-233.
134. TAGLIACOZZ, R. (1977) Self-citation in scientific literature. *Journal of Documentation*, **33**(4), 251.
135. THORNE, F. C. (1977) The citation index : another case of spurious validity. *Journal of Clinical Psychology*, **33**, 1157-1161.
136. TIJSSEN, R. J. W. and A. F. J. Van Raan. (1990) Net citation balance : a measure of influence between scientific journals. *JASIS*, **41**(4), 298-304.
137. TOFFLER, (1970) Alvin. Future Shock. N.Y. : Random House.
138. VAN RAAN, H. F. J. (1987) Impact of research performance as measured by citations a new model (in Informetrics 87/88 ws. by L. Egghe and R. Rousseau. Amsterdam : Elsevier), 293-299.
139. VINKLER, P. (1987) A quasi-quantitative citation model. *Scientometrics*, **12**, 47-72.
140. WEINSTOCK, M. (1971) Citation indexes. Encyclopedia of Library and Information Science ed. by A. Kent and others, **5**, 16-40.
141. WILSON, Patrick (1977). Public knowledge, private ignorance : towards a library and information policy. Westport (CT) Greenwood Press.
142. WOUTERS, Paul. (1993) Writing histories of scientometrics or what precisely is scientometrics? *4th Internàtional Conference on Bibliometrics, Informetrics and Scientometrics.* Berlin, 11-15 Sep. 1-10.
143. WOUTERS, Paul and L. Leydesdorff. (1993) Has Price's dream come true : is scientometrics a field of science? *4th International Conference on Bibliometrics, Informetrics and Scientometrics.* Berlin, 11-15 Sep.

# *Informing Democracy : A Review of the July 1994 Special Issue of the Journal of the American Society for Information Science*

**Lisa R. Schiff**
School of Information Management & Systems
102 South Hall, University of California at Berkeley
Berkeley, CA 94720 USA

**This article reviews the July 1994 special issue of the *Journal of the American Society for Information Science* on information and democracy. This issue offers a useful set of initial theoretical and empirical investigations of how information and information systems enable and constrain democratic practices. The weaknesses of many, though not all, of the pieces is a tendency towards utopian technological determinism and a focus on Western industrialized nations, particularly the United States. Keeping these informations in mind however, we are still provided with evidence that information and information systems that are *involving* hold the possibility for contributing to the growth and strength of democratic activities. Additionally, these pieces point the way towards further theoretical and empirical research.**

## 1. Introduction

Within the United States, the concepts of information and democracy have a long history as a rhetorical couple. Public libraries have constantly made this link a key part of their public discourse and justification for existence, though related actions to implement these concepts were less energetically taken up, and the realities were somewhat different from the touted ideals. Today, with the expansion of information systems and their increased accessibility to more US citizens, the buzz words "information" and "democracy" are paired and bounced about even more. The July 1994 special issue of the *Journal of the American Society for Information Science (JASIS)* on information resources and democracy (primarily in the context of the United States) is a reflection of both the historical position of this topic in information studies and the recent amount of attention the pairing has garnered in all sectors of US society, especially mass media.

The general position concerning information and democracy in U.S. common lore has been that better informed citizens can participate more effectively in the democratic process; in other words, people need more access to the information that exists "out there," and if this occurs, there will be a better society. While there are many difficulties with this, two are primary. The first is that no generic block of information exists to answer the needs of a given resident or citizen. Different people have different questions and needs, and thus will be satisfied with different types of information. Further, information itself is highly contested and reflects the social positions of those involved in generating it and using it, which may or may not be the same interests of those trying to find information. Second, and perhaps more fundamental, citizens can only fully participate in a democracy when their full participation is structural possible. Within a situation of unequal social, political and economic relations, and with a representative system largely driven by those sectors controlling the most wealth and influence, democratic

processes are greatly compromised. More access to information does not circumvent this central difficulty and does not, necessarily, do anything to alter it. However, access to information (not necessarily more access, but perhaps different access) is an essential ingredient to those wanting to alter such a system in order to support the opportunities for widespread democratic participation.

The articles in this *JASIS* issue scarcely begin to address this latter point, but they do begin to get somewhere on the first, however tentatively or obliquely, offering both theoretical discussions and empirical studies. A compelling vision of communicative information systems enhancing democratic possibilities runs through them all. This vision dangerously skirts technological-determinism, fueled by the communicative aspects embedded in new telecommunications technology and only qualified at the last instant by a recognition of social constraints. The trouble is with this last instant, which warrants much more than the tacit acknowledgment and lack of analysis that it receives. The possibilities for information resources, and especially information systems and telecommunications technology, to support democracy may indeed be endless, but we will only know how to proceed once we begin to clearly examine the road before us, looking at both the free pathways and the obstacles. Despite this weakness, and despite quite a US-centered set of articles, this special issue of *JASIS* does move us at least one step forward in understanding both theoretically and empirically some of the relationships between information resources and democratic practices.

## 2. Summary and Critique

A set of theoretical papers by Lievrouw, Braman and Dervin opens the issue, each reflecting on our foundational thinking processes in how we consider information and information in relationship to democracy, primarily within the context of the United States. All three offer critiques of traditional ways of thinking about information, democracy and the world, and by way of resolution offer alternative conceptions that in their strongest articulations could be only be enhanced by the use of information technology.

In her introductory article "Information Resources and Democracy: Understanding the Paradox," Lievrouw seeks to describe what she sees as a profound irony in the reality of contemporary political life in the United States—while information resources grow in number and accessibility, political participation continues to decline. Thus the commonly accepted notion that information is an all important ingredient of democracy seems to require some rethinking.

The center point for this rethinking, according to Lievrouw, should be the degree to which actual discourse is truly enhanced by new information resources and systems. This in turn can be evaluated by the degree to which an information and communication system is either informing or involving. Informing systems have little if any potential for discourse embedded in them. They are one-way systems, such as television, that send messages to a passive, relatively homogenized consumer. By contrast, involving systems create possibilities for dialogue and engagement. This type of system corresponds to an ideal notion of democracy that Lievrouw relies on called "discourse" democracy," a prerequisite to participatory democracy, which claims some inspiration from Habermas. Yet there are incredible constraints on this medium, as Lievrouw finally must admit at the end of her article, namely continued media ownership concentration, a trend towards product delivery, and costs of constantly changing technology.

Sandra Braman sees almost revolutionary potential in new information and communication technologies. In her article, "The Autopoietic State: Communication and Democratic Potential in the Net," Braman heralds new information and communication technologies as the veritable linchpin for democracy. We can choose to have an ever expanding information infrastructure that brings together "knowledge and decisionmaking powers" and 'local and general ("scientific") knowledge' which will enhance our potential for action.

Alongside this theme, Braman introduces systems theory and chaos theory, making a tight parallel between the natural phenomena various scientists eventually described under the heading chaos theory, and the reasoned yet turbulent state of social organizations and especially nation-states in the modern era. For this latter point there are two thrusts: 1) the revisioning of nation states, from the redrawing of national boundaries (e.g. the former Soviet Bloc) to the dismantling of welfare

states and 2) the dominance of transnational corporations (TNCs) which on one level hold no allegiances to their nations of origin. Braman then embarks on a lengthy, confused and distracting ontological discussion which when married with the nation-state analysis results in her theory that we are entering a world in which turbulence is so pervasive that a given state must constantly regenerate itself (autopoeisis). This turbulence and autopoeisis is a positive phenomena because it provides openings for new social formations. Information enters into this process at several points: 1) complex information and communication systems allowed TNCs to develop; 2) information can be an act of power, as with the collocation of knowledge; 3) information helps transform potential power to actual power, as with education; and 4) information "is" a resource about the past, present and future.

Many problems crop up throughout this article, such as a tendency towards technological determinism—the systems allow social possibilities, when in actuality the converse is true; a general ahistoricism that is particularly revealing in the discussion of TNCs; an implicit assumption that US experiences reflect the experiences of all other nations; a lack of discussion concerning the political and economic constraints on the design, implementation and use of information technology; and finally a lack of connection between the discussion of nation-states and ontology with the much briefer discussion of information systems.

While sharing with Braman an interest in chaos theory and a penchant for ontological perusals, Brenda Dervin's article much more clearly focuses on questions of information.

In "Information Democracy : An Examination of Underlying Assumptions," Dervin deconstructs the dominant assumption that information is inextricably tied to democracy, i.e. that there is a widespread desire to be "correctly" informed via more access to "good" or "correct" information. Dervin views this assumption through six different ontological views of the world, each of which understand reality, nature and human beings differently, especially in regards to their respective degrees of chaos, orderliness, and dependence upon authority. In each she discusses what "good information" means.

Dervin argues that information and communication systems are primarily built around the assumption that the goal is to get "good" information to people and that a "free market place of ideas" will do this, because "only the best information" survives. Such systems assume, and in certain ways, enforce our belief in one reality. This is problematic because it favors an understanding of reality that is based on one cultural norm (that of the sector that established the system) and doesn't recognize the different understandings that co-exist between and within cultures, and the different degrees of access to use and to create information systems. Worst of all, this type of approach maintains hierarchical lines of authority in order to prevent a challenge to and a disintegration of the legitimacy of the single world view that is required to maintain that authority.

Dervin provides us an alternative ontological framework at the end of her discussion and criticism, which she calls communitarianism—a blend of cultural relativism's respect for other sets of values and postmodernism's individual realities, essentially a never ending back and forth process literally creating reality through communication. Within this context, what can be considered "good" information will always depend on who is doing the asking and for what purpose. Dervin offers an example of this by analyzing the debate on types of health care in the United States, showing how authority defines what is correct and good information that actually impedes the process of acquiring good health care that may come from non-Western resources.

As far as democracy is concerned then, information is essential, but only when considered in light of the person and the context of their need, thus neither information nor democracy conform to the dominant narrative in US society. For information systems to go beyond this narrative, they must be recursive and involve users from the outset of design. Diversity will thus be embodied and challenge to authority will be possible. By implication, this is the ideal pairing between information and democracy.

The theoretical pieces discussed above stress the need for multi-perspective, dynamic systems based on communication, if information systems are to support democracy. The remaining four articles present real-world examples of spaces where information systems and democratic potential seem to cross-paths, successfully or not. These works

allow us to begin to assess the real possibilities of implementing information systems that can enhance the possibilities outlined by Lievrouw, Braman and Dervin.

Within this set of empirical works, two pieces most clearly explore the informing versus involving aspects of information systems. The first is John Newhagen's article, "Media Use and Political Efficacy: The Suburbanization of Race and Class" which provides an interesting example of the effects of mass media on people of different ethnic and class backgrounds, in terms of their political expectations. The second is "The PEN Project in Santa Monica: Interactive Communication, Equality, and Political Action" by Everett M. Rogers, Lori Collins-Jarvis, and Joseph Schmitz, which details how design of such systems can support political participation. In each we see evidence for the power of involving information and communication systems to support activities closer to or squarely in the realm of participatory democracy.

Information in the context of the Newhagan piece can be defined as understandings gathered from such sources as television, newspapers and radio, primarily the informing media in Lievrouw's terminology. Democracy is looked at in terms of an individual's estimation of system efficacy—the political/governmental system's ability to achieve its goals—and self-efficacy—the ability of that individual to cope with the political system. Newhagan's research was based on a random telephone survey of a sample stratified by race (African-American or White) and class (based on neighborhood housing values) within one county in the United States. Demographic information such as education level was collected at the time of the survey.

The findings are interesting and support the author's hypotheses, though they are presented in a way that makes it easy to see differences associated with race, but difficult to see differences associated with class or education. Use of newspapers and national television news increased self-efficacy in general, especially for African-Americans but tended to decrease system-efficacy for African-Americans. Primetime television viewing (a non-traditional, yet critical and often overlooked definition of information) sharply decreased both self-efficacy and system-efficacy for African-Americans and caused relatively little change for white respondents. This is not surprising considering the preponderance of negative images of African-Americans on primetime television. Finally, listening to talk-radio shows dramatically increased self-efficacy (system efficacy was not discussed) as did talking with others. These results provide initial support for Lievrouw's contrasting informing and involving systems, making an interesting step towards connecting information seeking and political engagement.

The article on Santa Monica's community network, "The PEN Project in Santa Monica: Interactive Communication, Equality, and Political Action," follows up on Newhagan's final point, taking the next step from a one-to-many system such as a call-in radio talk show, to a many-to-many communications system. Of all of the empirical articles in this issue, this case study of the Santa Monica community network, based on previous research by the authors, most clearly reflects an example of the intersection between democracy and information technology. The authors provide a brief history of the development of Santa Monica's Public Electronic Network (PEN), focusing on the impact of design decisions on use of the system, particularly in terms of "socioeconomic and gender equality aspects of the public's use."

Community networks exemplify the involving systems Lievrouw would like to see and so far have managed to stay away from the corporate imperatives that could potentially constrain them. PEN is probably one of the best known community networks because of its early existence and the impact it had on Santa Monica. It was designed with traditional democratic goals in mind—increased access to public information, city services, communication, electronic forums, and technology itself. Several major design decisions contributed to the democratic nature of the system, the most important of which seem to be having publicly accessible terminals, interactive conferences, and free email for residents. These decisions were supported and pushed through by the liberal office holders in Santa Monica at the time and had many positive effects, including increasing the percentage of users by providing access to residents who did not have computers with modems at home or at work (resulting in a noticeably high number of female users) and providing direct access to local politi-

cal processes for all registrants, but especially for such commonly disenfranchised groups as the homeless.

This article is a welcome introduction to community networks and PEN in particular, however it merely whets the appetite. Described as an analysis of how the design process effects usage, we are given little detail about that process and little analysis of it. For instance, was the importance of considering users in the design process recognized from the start and embraced by all key parties (designers, funders, vocal user groups)? If not, how was this perspective achieved? How were potential users involved? What design issues didn't get implemented and why? These are fundamental questions to the development of an information system, but especially those that hope to support democratic practices. A better understanding of how these and other issues arose and were addressed in such a successful system as PEN would be an invaluable resource for the rest of the community networking field and the information studies field in general.

Michael Martinez's article, "Access to Information Technologies among School-Age Children: Implications for a Democratic Society" is a classic example of one strand of the "haves and have-nots" debate surrounding access to information technologies. This debate expresses the concern that those who have access to telecommunications technology and know how to use it will have power, and that those who do not have the means (primarily financial) to gain this access will be shut out. Many in this debate assume that information and information technology is a vital resource in society today.

In some ways, this piece is the odd-one out in this collection of articles because of its foundational premise: the direct connection between information and democracy, the very same connection that Brenda Dervin deconstructs in her article in this same issue. This piece is much more along the lines of Braman's discussion, seeing a myriad of possibilities and few limitations in technology. Martinez reasons that if information is central to a successful democracy, and computers are becoming the primary mechanism for dissemination of and access to information, then computers are likely to become an additional essential ingredient for democracy. If this is true, and if the ideal of democracy implies participation amongst equals (at least in the sense of having equal access to information), then levels of access to computers is something to be concerned with if we want to have a successful democracy.

Martinez contributes to the examination of this set of relationships by looking at access to computers and computer literacy by children, ranging in ages and ethnic/racial backgrounds, using data on United States elementary and secondary school students from the National Assessment of Educational Progress which measures computer knowledge and skill. He found that access to computers in school was correlated to computer competence, but also strongly related is extracurricular access, at home for example. Martinez's data didn't allow him to examine other variables, such as socioeconomic status, as an important indicator for computer competence. He suggests that this might be considered, but feels that home access to a computer is the strongest explanatory factor.

While all of this is interesting and certainly contributes to the empirical support for the "haves/have-nots" debate, Martinez does not really begin to address issues concerning democracy, information, technology and children. The connection between information and democracy, and technology and democracy is scarcely examined. Unfortunately, Martinez assumes that technology and the realms it can take one to are somehow unaffected by the society in which they exists, which is certainly false. There is no reason to think that the existence of computers and technology will somehow cause a democratic flowering on their own. No doubt computer literacy is useful in today's world, but computer literacy is simply one skill among many that a well prepared citizen will want, and arguably not the most important.

Su-Lien Sun and George A. Barnett, in their article "The International Telephone Network and Democratization" provide the sole discussion of any true international scope in this *JASIS* special issue. Their piece examines the connection between a nation-state's position in the international telephone network and its level of democratization, looked at via a network analysis. The authors attempt to simultaneously reveal the underlying structure of relations between nations, situating their analysis within an analytical framework of global economic and political interdependence, influenced

especially by Wallerstein and centered around the center-periphery model. This model maps the patterns and relations of dominance created by colonialism and imperialism. Thus imperialist nations tend to be in the center, nations that engage in both imperialist- like behavior but which are also the subject of imperialist relations comprise the semi-periphery, and finally periphery nations tend to be primarily subjects of imperialist behavior. A nation's placement in the center-periphery system is usually highly correlated to levels of industrialization. Along side this picture of global relations is a picture of telecommunications and democratization. Sun and Barnett see great potential in communications technology, especially two-way mediums, for supporting the development of democracy, especially because of the potential diffusion of ideas and interactive opportunities.

The authors compared communications levels by combining several measures from AT&T's self-report survey for nations about international telephone traffic in 1982, 1986 and 1990 that together could be thought of as network robustness. They compared levels of democracy by looking at international rankings of political rights and civil liberties as measured by Freedom House starting in 1972, using a standard of participatory democracy. This is an interesting comparison which would have been more interesting if we had been provided with as much of the data concerning measures of democracy as we did around telephone network traffic. Given that political and civil liberties can be define and measured in many ways and against many different standards, it is difficult to accept their scale blindly.

Nevertheless, a rather strong association between levels of democracy and telephone systems was discovered, without, however any causal links, as demonstrated by the lack of a time lag between the development of the telephone network and democratic development. A secondary finding was that economic development and political development are connected, as are telecommunications development and economic development. All of this goes to say that those nations with the most economic development, those in the core, are also those with the highest level of developed communications systems and democratic participation. Finally, the authors found patterns of communication that followed the lines of colonialism, e.g. the United Kingdom is the country called most frequently from India and Hong Kong.

The conclusion of Sun and Barnett seems to be that while telecommunications systems have great potential to expand democratic possibilities, they are constrained by the global core-periphery relations and further enable those relations to continue. Proof of the latter is found in the core/semi-periphery/periphery pattern found in the telephone network data. As the authors state, "The pervasive influence of communication technologies also generates electronic colonialism. The powerful new technologies may widen the gap between the core and the periphery."

## 3. Conclusion

In terms of the topic of information imperialism, there is little global perspective in this issue of *JASIS*, except Sun and Barnett's piece on the international telephone network, and parts of both Braman's and Dervin's theoretical discussions. To use the vocabulary of Sun and Barnett's analytical framework, this collection of articles is very much about the domestic concerns of core nations (or at least one core nation) in the global community. However, it is not improbable that the findings are generalizable.

With that in mind, the obvious conclusion from the set of articles presented in this JASIS issue is that information resources can both constrain and enable participatory democratic practices. Information and communication systems are tools that can be developed and used for a wide-variety of ends, and they shape and are shaped by the social contexts in which exist. Further, the information that gets accessed and generated within these systems both constrains and enables such practices.

From the empirical research, it is seen that the common link between information resources that enable people to communicate, to express and find a diversity of information and to eventually act does seem to reside in Lievrouw's notion of an involving system that engages actors and thus encourages participation. With such proof of the potential value of involving systems to support democratic participation, a rigorous discussion of the economic, political and social impediments to their realization, and possibilities for surpassing those impediments, would have filled a noticeable void in this collection. This however, may be the next step for those concerned with the use of information resources for democratic possibilities.

BOOK REVIEW

# *Yu. T. Sharabchiev. "Communications in Science : Sociometric Aspect" (1995)*

Dr. Sharabchiev's monograph *Communications in Science : Sociometric Aspects* is devoted to the problems of communicative relationships and communication structure of science. Dr. Sharabchiev considers communications an essential part of the social organization of science, of which there are three main aspects : informational, social, and cognitive. Being a social system, communications include the whole bulk of existing scientific information; scientists as a social category of information producers and users (problems of priorities, prestige, stratification, etc.); and the channels for information transmission and use (libraries, informational services, editorial boards of the scientific and abstract journals, etc.). Use of sociometric approach (citation and cluster analysis, content analysis, co-authorship studies, etc.) in the studies of communicational process helps to reveal the mechanisms of scientific communications, *hot* topics and trends in science, interrelationships of the nucleus and research front of a science, etc.

The impressive bulk of scientometric literature (268 references) reviewed in the book cover the following issues : social relationships and interactions in science, concept and methodology of sociometrics (or scientometrics, informetrics) of scientific communications, structure and characteristics of the channels of scientific communications (scientific communications as a social and informational process). There are many kinds of communications, and each performs its own specific socially oriented communicational function. According to Dr. Sharabchiev, these include advance information (bibliographic descriptions, catalogues, etc.), interdisciplinary and international communication (abstract journals, international multidisciplinary periodicals), operative informing (abstracts of scientific conferences), scientific generalization of facts and knowledge (scientific reviews and monographs), educational functions (text books).

Sociometric studies of the patterns of communicative behaviour of biomedical scientists and practicing medical professionals in Belorussia revealed a substantial difference between these two groups. Because of the lack of the developed network of specific *innovation* communications, a great deal of the most important scientific results in medicine remain within the sphere of scientific activity and cannot reach the specialists of practical medicine who use an entirely different channels of scientific communications and a different information base. A hierarchy in the communicational structure of medicine in Belorussia was shown to base on the existing structure of the health care system that is different from the hierarchy of priorities of medical science.

Development of science is shown to be accompanied by the growth of the interdisciplinary interaction of scientists and by the increase in the number of co-authors, but on the other hand, the works that had the greatest impact were performed by one or two authors.

The chapter called *Sociometric assessment of communicative functions in science* deals with the following issues : diffusion and concentration of information in the channels of scientific communications, *aging* of information in the system of scientific communications, and communicative barriers between the scientists and scientific communities (including geosocial barriers in the structure of center-periphery relations and language barriers). Analysis of current trends in the publishing activity of scientists and the use of techniques based on extrapolation of the results of the studies of publication dynamics enabled Dr. Sharabchiev to suggest that the problem of informational explosion exists not on the level of new ideas which number is limited and does not

exceed 0.1% of the total number of annually published works, but on the level of total volume of publications, 1/3 of which can be regarded as the *informational noise.*

The international scientific community can be characterized by a certain pattern of center-periphery relationships, the distribution of social roles between the countries is such that some of the countries act as the centers of intellectual influence, while others act as communication or innovation centers. The countries that do not participate in the exchange of information are practically isolated from the international scientific community. Communicative barriers are shown to be socially determined and new information technologies not only fail to eliminate the old bar-riers, but also create the new ones.

Dr. Sharabchiev suggests that the importance of a periodical as a part of the system of scientific communications can be assessed using the following indicators : number of publications devoted to the matter in the complete set of journals for one year, number of references to the articles published in this journal, and the number of library orders for this journal.

Dr. Sharabchiev believes that according to stratification of scientists into the highly productive elite and low productive majority, there exist two massifs of scientific documents and two communication systems. One of these is a massif of highly cited documents that cite other highly cited documents as their predecessors and consequently lead to appearance of other highly cited documents. Another is a much-bulkier massif of uncited documents that cite uncited and low-cited documents and do not generate new knowledge. The system of scientific communications that recurrently generates new knowledge includes *nuclear* journals related to the topic and prestigious scientific periodicals of a general character, monographs, and reviews.

The aspect of communications between the nucleus and research front of science is another important topic covered by Dr. Sharabchiev in his book. Several interesting generalizations had been done in this area. Thus it was shown that the level of self-citation lowers after the formal acknowledgement of the scientist's achievements by his colleagues. The elite group of scientists behaves as a rather closed system widely intercommunicating within its own, developed media. Social communications of medical scientists of the CIS countries with foreign scientists is based on the use of different information bases, and the network of scientific communications employed by foreign scientists only partly overlaps with that of Russian or Belorussian scientists, which also hinders the development of medical science not only in Belorussia and other CIS countries but also in the foreign countries. It is also concluded that the problem of overproduction of scientific information should be regarded as a normal result of the progress in science and technology since the most important works (as well as the most outstanding scientists) may function normally only in the context of the works and scientific community of a lower level, the existence of which is a necessary phenomenon of scientific activity. The nucleus and the research front of science are interconnected by the network of social contacts that provide the transformation of research front into science nucleus introducing changes in the existing paradigm in the course of numerous expertises.

The ideas and conclusions stated in this book are supported by the extensive data summarized in 43 tables.

**Maria M Klavdieva**
Institute of the History of Science and Technology
Russian Academy of Sciences
Staropansky per., 1/5,
Moscow, 103012, Russia

# JISSI Data Sheets

**Title of the journal** : Jissi : the international journal of scientometrics and informetrics

**ISSN** : 0971 - 6696

**CODEN :**

**Established** : March,1995

**Periodicity** : Quarterly (issued March, June, September, December

**Subjects covered** : Scientometrics, Informetrics, Bibliometrics, Technometrics, Librametrics. Library statistics, Citation & co-citation analysis, Co-word analysis, Co-reference studies, Science indication and forecasting, Psychometrics and sociometrics of invention, discovery, authorship and informational and communicational activities, Econometrics of information, Historiometrics of knowledge

**Broad subjects** : Science of science; History of science, technology, social science, arts and humanities; Sociology of knowledge and information; Library and information science; Information management; Economics of information; Psychology and socio-psychology of information and knowledge; Applied statistics; applied mathematics.

**Place of publication** : New Delhi, India

**Publishers** : Brzark Information System (P) Ltd

**Publishers, address** : 112, Humayun Pur, Safdarjung Enclave, New Delhi - 110029, India, Telephone : +91–11–618-2366, Telefax : +91–11–617-7148

**Editor** : Subir K Sen (Honorary Executive Editor)

**Editorial address** : 85 (Old) Devi Nivas Road, P. O. Motijhil, Calcutta 700 074, India Telefax : +91–33–551-2180

**Editorial board** : David Andrich (Australia), Subbiah Arunachalam (India), Donald deB. Beaver (USA), Quentin Burrel (UK), Jean Pierre Courtial (France), Mari Davis (Australia), Leo Egghe (Belgium). Hajime Eto (Japan), Belver C Griffith (USA), S D Haitun (Russia), Ashok Jain (India), Michael E D Koenig (USA), Hildrun Kretschmer (Germany), Stephen Lawani (Nigeria), Loet leydesdorff (Netherlands), Cesar A Macias-Chapunal (Mexico), Sinisa Maricic (Croatia), William E McGarth (USA), Jack Meadows (UK), Bluma Peritz (Israel), I K Ravichandra Rao (India), Ronald Rousseau (Belgium), B K Sen (Malaysia), Shan Shi (China), Jean Tague-Sutcliffe (Canada), Dimiter T Tomov (Bulgaria).

**Collaboration** : In association with the International Society for Scientometrics and Informetrics (ISSI)

**Attributes** : JISSI is an international journal representing a challenging new field of scientometrics and informetrics (or rather sciento-infor-metrics). It provides a forum for dissemination and exchange of ideas and scholarship in this field.

**The subject field :** Sciento-infor-metrics is a fast growing multidisciplinary as well as interdisciplinary fiedl of research and study. The field may very well represent quantitative ( and qualitative ) assessments of man's cultural progress ( sociometry and historiometry) and advent into information age. University departments and academic institutions dealing with history and sociology of science, science of science, library and information science etc are imparting regular and irregular teaching courses and running doctoral programmes in different areas of this field.

**Target readership :** Research workers, teachers, students, practising information scientist, librarians, information managers, science and technology policy makers, historians of science, social scientists, statisticians, etc.

**Coverage :** JISSI is a refereed journal devoted to the quantitative aspects ( including mathematical, statistical, computational analysis and modelling ) of - science of science & technology ( including technometrics ) - studies on social sciences, arts and humanities ( particularly based on bibliographic data) - all types of information processes e.g., generation, growth, diffusion, dynamics, use etc of information systems and services (including libraries, archives, databases, information centres or any other type of document collections) which may go by the name of librametrics any study that may fall under the purview of bibliometrics

**Features :** Research papers, reviews, short communications, letter to the editor, book reviews, reprint of classical and very important but apparently obscure papers relevant to the field, profiles of important personalities ( these should be serious essays dealing with the work and just not life sketches), glossries of scientometric and informetric terms, editorials and guest editorials (solicited and unsolicited), issues and unsolved problems, ISSI Newsletters ( as and when provided for by the ISSI board), selected and refereed papers from a relevant Conference, Seminar, Symposium etc.

**Size :** A4 (27.5cm X 20.5cm approx)

**Print area :** 24cm X 16cm

**Advertisements :** JISSI publishes selected ads

**Price (Single copy) :** US $ 55 = 00/Rs 750.00

**Subscription Information**

| **Annual subscription :** | INSTITUTIONS | INDIVIDUALS | ISSI MEMBERS |
|---|---|---|---|
| **Developed countires :** | US $ 200=00 | US $ 100=00 | US $ 75=00 |
| **Developing countries :** | US $ 125=00 | US $ 50=00 | US $ 30=00 |
| **Indian subcontinent :** | Rs. 2500=00 | Rs. 1000=00 | Rs. 750=00 |

** Spcial offer for advance subscription for three years : Double the amount of Annual Subscription