SOCIAL STRATIFICATION OF AUTHORS REVEALED FROM THE COAUTHORSHIP NETWORK

Hildrun KRETSCHEMER

Institute of Theory, History and Organisation of Science,
Academy of Sciences of the GDR,
Prenzlauer Promenade 149-152, 1100 Berlin (GDR)

Abstract

Seven bibliographies from the fields of medicine, physics and social sciences were used. The authors were classified by groups in accordance with the number of publications per author. Studies were made to determine the statistically expected number of coauthorships by proceeding from assuming an independence of coauthorship between authors from the number of their publications.

Hypothesis: The proportion of the sum of coauthorship found between authors with the same number of publications to the sum of the statistically expected one is larger than the proportion of the sum of coauthorships found between authors with a different number of publications to the sum of the statistically expected one. This hypothesis could be verified in all seven bibliographies. Coauthorships between authors do not come into being independently of the number of their publications, i.e. of their social ranks.

INTRODUCTION

"One of the most notable changes in the social organization of scientific research in the twentieth century has been the rapid spread of collaborative research and teamwork... Since World War I, the practice of collaborative research has grown at a markedly increased rate, compared with its slow progress in the nineteenth and early twentieth centuries... World War II witnessed the emergence of the large team, in close association with the development of Big Science... today, joint research has become the model form of scientific inquiry" [1, p.14].

Thus, at present it is necessary to find measurable regularities between scientific productivity and the scope of collaborative research [2].

D. de Solla Price [3] indicated the relationship between the importance of the scientist and the logarithm of the number of papers he has published during his life.

As a rule, scientists who publish a larger number of publications during the course of their lives, publish also more, within a limited period of time, than other scientists. So it is to be assumed that the rank of a scientist is also influenced, in a tendency-like manner, by the number of publications he publishes within the period of time concerned.

According to a general social psychology theory [4], groups with a rank order show less distance between members with the same rank than between different
ranks, i.e. the frequency of contacts between partners with the same rank is highest and decreases with the growing distance between the ranks. It is important to realise that this stratification form is generally valid for any groups with rank orders. This means that the performance of such groups as a whole is increased by their stratification. Stratification appears to be an optimum form of contacts among the group members, whereby not only the more frequent contact between group members of the same rank is of importance but also less contact between the group members of different ranks. Only in this case the group as a whole exists.

In this sense, the rank of an author was determined by the number of his publications in the analysis of coauthorships, and coauthorships of scientists expressed their contacts. The fact was investigated whether the structure of coauthorships was in line with the social psychological theory mentioned above.

METHODS

The subject of study were publications either in one periodical or of one problem area or specialty respectively within a fixed period of time. Both all single-authored, and all multi-authored papers were investigated.

The number of publications per author was ascertained from a bibliography in accordance to the "normal count procedure". Normal count, leading to full authorship, gives full credit to all contributors because each appearance of the author's name in the by-line is counted.

The authors were classified by groups according to the number of publications per author, that is, all authors with the same number of publications are combined in one group. The matrix of relations between coauthors can be established by two different methods:

1. The result is a symmetric matrix under the condition that the place which a coauthor is given in the order of coauthors of an article is not of importance.
2. The result is a non-symmetric matrix under the condition that this order of places is taken into account.

An example is used to demonstrate the elaboration both of a symmetric and of a non-symmetric matrix. The bibliography of articles of the four authors A, B, C and D is given. The "normal count procedure" is used to find the number of publications per author in this bibliography. In the following, this number is given in brackets behind the name of the authors concerned.

The authors were represented as coauthors in a determined order.

Bibliographies (without headlines of articles):

1. A(1), C(3), B(1), D(4)
2. D(4), C(3)
3. D(4)
4. C(3)
5. D(4)

When establishing the matrix of relations between coauthors, those authors are to be called "collaborating authors" who are to be the viewpoints for considering relations to other authors. The authors to whom these relations go are called "collaborators". A collaborating author is to be given the rank
i according to the number of his publications. A collaborator is to be given
the rank j according to the number of his publications. In the example, the
number of publications is to be equal to the rank. The authors A(1) and B(1)
are combined in a group with the rank i = j = 1. Author C(3) is in the group
with the rank i = j = 3, and author D(4) is in the group with the rank
i = j = 4.

Symmetric matrix:

When author A and author B publish an article together, this relation between
them is recorded into the matrix twice - once from the viewpoint of author A
into the direction of author B, and once from the viewpoint of author B into
the direction of author A. This principle is analogously continued when more
than two authors publish a paper together.

Thus the following relations for the first article appear in the first matrix
(symmetric matrix, cf. Table 1):

- from the viewpoint of author A(1), that is, from the collaborating author
  A(1) : one relation to C(3), one to B(1), one to D(4)
- from the viewpoint of the collaborating author C(3) : one relation to A(1),
  one to B(1) and one to D(4)
- from the viewpoint of the collaborating author B(1) : one relation to A(1),
  one to C(3), one to D(4)
- from the viewpoint of the collaborating author D(4) : one relation to A(1),
  one to C(3), one to B(1).

The relations of the second article were added to this matrix:

- from the viewpoint of the collaborating author D(4) : one relation to C(3)
- from the viewpoint of the collaborating author C(3) : one relation to D(4).

Table 1: Matrix of relations through co-authorships,
symmetrized matrix

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

In case of a larger number of multi-authored papers, this principle would be
applied respectively. The number of relations directed from collaborating
authors with the rank i to collaborators with the rank j is called \( C_{ij} \).

Non-symmetric matrix:

When author A and author B publish an article together, this relation between
them is registered into the matrix only once, that is, from the viewpoint of
the author who is above the other author in the order of places of coauthors.

That's why the second matrix (non-symmetric matrix, cf. Table 2) contains the
following relations for the first article:
Table 2: Non-symmetrical matrix

\[ i, j \text{ - rank of authors} \]

<table>
<thead>
<tr>
<th>i</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td></td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- from the viewpoint of the collaborating author A(1): one relation to C(3), one to B(1), one to D(4)
- from the viewpoint of the collaborating author C(3): in contrast to the symmetric matrix, only to collaborators who were represented behind him, that is, one relation to B(1), one to D(4)
- from the viewpoint of the collaborating author B(1): one relation to D(4)

As no further author is registered behind D(4), there does not appear any relation from D(4) to another author in the matrix (related to the 1. article).

Relations of the second article are added again to this non-symmetric matrix:

- from the viewpoint of the collaborating author D(4): one relation to collaborator C(3).

The number of relations which are directed from collaborating authors with the rank \( i \) to collaborators with the rank \( j \) is called \( C_{ij} \).

In our example, we equated the rank of a group of scientists with the number of publications per scientist. This is possible in case of a sufficiently large bibliography. The minimum number of scientists per group should not be less than 10 because strong statistical deviations in too small groups exist. That means the authors with the highest number of publications per author at least are to be combined in one group. It is also justified to give the same rank to these authors with a larger difference in the number of their publications because the rank of a scientist should correspond to the logarithm of the number of his publications rather than their direct number.

For this reason, it is recommendable, also for investigations of smaller bibliographies, to combine all authors with 1 publication in the first group with rank 1, all authors with 2 - 3 publications in the second group with rank 2, authors with 4 - 7 publications in the third group with rank 3, authors with 8 - 15 publications in the fourth group with rank 4 etc.

The above combinations can be used for more clearly demonstrating the general trend which corresponds to the above mentioned social psychological theory. On the other hand, it is possible to examine the following mathematically formulated hypothesis both by means of summaries and without them (summaries and assignment of ranks in the investigated bibliographies see "Data").

Many types of scientometrics data can be presented as transaction matrix. In all cases the matrix consists of a set of items assigned to each row and column with each cell containing the level of transaction between the row and column items. One may model the level of transaction from item \( i \) to item \( j \) as
independent contributions from a row coefficient and a column coefficient.

In this model, the probability of any transaction from item \(i\) is:

\[ p_{it} = \frac{\text{sum of all transactions in row } i}{\text{total sum of transactions}}, \]

while the probability of any transaction received by item \(j\) is:

\[ p_{jt} = \frac{\text{sum of transactions in column } j}{\text{total sum of transactions}}. \]

Assuming independent contributions from the row and column probabilities, the probability that a transaction from \(i\) to \(j\) occurs is:

\[ p_{ij} = (p_{it})(p_{jt}) \]

Multiplying by the total number of transactions within the matrix yields an expected number of transactions from \(i\) to \(j\),

\[ m_{ij} = (p_{it})(p_{jt}) \text{ (total sum of transactions)} \]

Assuming independence of the observed number of authors with rank \(i\) and their collaborators with rank \(j\) deviates from the statistically expected number. By means of Chi-square-tests it is examined whether the deviations of observed number from expected number are significant in the matrix generally.

The matrix has \(I\) rows and \(J\) columns.

\[ x^2 = \sum_{i=1}^{I} \sum_{j=1}^{J} \frac{(C_{ij} - m_{ij})^2}{m_{ij}} \]

with \(df = (I-1)(J-1)\) degrees of freedom. When \(x^2\) has a statistically significant value, coauthorship between authors does not come into being independent of their ranks. On the assumption of independence, that there is a transaction from \(i\) to \(j\) the probability is \(p_{ij}\), and the expectancy value is \(m_{ij}\) as already mentioned. Probability of the fact that there is either transaction from \(i\) to \(j\), or transaction from \(k\) to \(l\), or that there are both transactions, is \((p_{ij} + p_{kl})\) with an expectancy value of \((m_{ij} + m_{kl})\). This principle can be continued for summing up any number of cells of the matrix. For example, the probability of the sum of transactions for which the difference between \(i\) and \(j\) is constant \((i-j = \text{const})\) is equal to the sum of probabilities

\[ \sum_{i-j = \text{const}} \sum_{i,j} p_{ij}, \]

So the expectancy value of this sum of transactions is

\[ \sum_{i-j = \text{const}} m_{ij}, \]

There is a special case: for all cells in the main diagonal, the rank of collaborating authors is equal to the rank of collaborators, that is, \(i-j = 0\). Thus, the sum of expectancy values of all cells of the diagonal is the expectancy value of the result that collaborating authors and collaborators are of the same rank, assuming that authors establish coauthorships independently of their ranks. In line with social psychological theory, it is useful to compare the sum of the observed number of all cells of the main diagonal with the respective sum of the expected number.
The Freeman-Tukey statistic \([5]\) is a method for displaying the deviations of the observed \((c_{ij})\) from the statistically expected \((m_{ij})\) relations for one single cell of the matrix. The FT statistic

\[
FT = \sqrt{c_{ij}^2 + v_{ij}^2} - \sqrt{m_{ij}^2 + 1}
\]

allows the deviations to be evaluated as samples from a normal distribution. FT deviates lower than \(-1.64\) show a significant deficit in the number of observed transactions while deviates in excess of \(1.64\) show a significant excess in the observed number of transactions (significance on the 5 per cent level). The Freeman-Tukey statistic can also be used for comparing the observed number and statistically expected one for a combination of cells, for example, for the combination of the cells of the main diagonal, or, on the other hand, for combining all the remaining cells containing transactions between authors of different ranks.

**DATA**

Seven bibliographies were studied:


   \[
   \sum_{i,j} c_{ij} = 1280
   \]

   Rank 1 was given to the group of authors with 1 publication per author, rank 2 to the group of authors with 2 publications, rank 3 to the group of authors with 3 publications, rank 4 to the group of authors with 4 publications, and rank 5 to the group of authors with 5 or more publications. In addition, the bibliography was sub-divided as follows:

   1.1. Bibliography with publications from the USA

   1.2. Bibliographies with publications from other countries.

2. Bibliography with publications about the chemotherapeutic Endoxan (1982-1985), source of this bibliography: MEDLARS (medical publications) [7]

   \[
   \sum_{i,j} c_{ij} = 12222
   \]

   The assignment of ranks to groups of authors corresponds to the assignment applied with regard to the first bibliography.


   \[
   \sum_{i,j} c_{ij} = 2276
   \]

   Source of bibliography: INSPEC.

Although the total sum of \(c_{ij}\) in the matrix is smaller in this bibliography than in the bibliography of articles about Endoxan, the assignment of ranks to the groups of authors can be extended because there is a larger number of authors with a high number of publications. In the bibliography about Endoxan, there was only one author with more than 7 publications whereas in the above case there are 17 authors with more than 7 publications. Rank 1 was given to authors with 1 publication, rank 2 to authors with 2 publications and so on up to rank 8 to authors with 8 or more publications.

$$\sum_{i} \sum_{j} C_{ij} = 2234$$

The assignment of ranks corresponds to the assignment of ranks in the second bibliography.

The four preceding bibliographies contained publications from different periodicals from all over the world. The following three bibliographies, however, refer to one periodical each.

5. Bibliography of publications in "Czechoslovak Journal of Physics".

The authors were classified by groups according to the number of their publications published from 1981 to 1984. The matrix of $C_{ij}$, however, contains only articles from the year 1984 [10]

$$\sum_{i} \sum_{j} C_{ij} = 816$$

Due to the small number of $C_{ij}$ in the matrix, the groups of authors were combined more strongly than in the preceding bibliographies. Rank 1 was given to the group of authors with 1 publication, rank 2 to authors with 2-3 publications and rank 3 to the group of authors with 4 or more publications.


$$\sum_{i} \sum_{j} C_{ij} = 393$$

The assignment of ranks corresponds to the preceding one.

7. Bibliography of publications in "Deutsche Zeitschrift für Philosophie".

The authors were classified by groups according to the number of their publications from 1953 to 1977. A random selection of articles was the basis of investigating coauthorships [10],

$$\sum_{i} \sum_{j} C_{ij} = 586$$

The assignment of ranks corresponds to that of the other two periodicals.

RESULTS AND DISCUSSION

The first example is the symmetric matrix (cf. Table 3) which was elaborated according to the first bibliography (Quantitative studies of science). $X^2 = 716.6 > X^2_{0.01}$, that is coauthorships do not come into being independently of the ranks of collaborating authors and collaborators. According to the social psychological theory, the smallest distance is between members of the same rank, this distance increases with the distance between ranks. To illustrate this fact, a proportion is established - under the condition that the difference between the ranks of collaborating authors and collaborators is constant: $(i-j = \text{const})$ - between the sum of observed values and the sum of expected values:
Table 3: Symmetrized matrix of the first bibliography

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>63</td>
<td>63</td>
<td>32</td>
<td>17</td>
<td>135</td>
</tr>
<tr>
<td></td>
<td>(475)</td>
<td>(103)</td>
<td>(48)</td>
<td>(32)</td>
<td>(121)</td>
</tr>
<tr>
<td>2</td>
<td>63</td>
<td>78</td>
<td>13</td>
<td>1</td>
<td>13</td>
</tr>
<tr>
<td></td>
<td>(103)</td>
<td>(22)</td>
<td>(10)</td>
<td>(7)</td>
<td>(26)</td>
</tr>
<tr>
<td>3</td>
<td>32</td>
<td>13</td>
<td>14</td>
<td>9</td>
<td>11</td>
</tr>
<tr>
<td></td>
<td>(48)</td>
<td>(10)</td>
<td>(49)</td>
<td>(3)</td>
<td>(12)</td>
</tr>
<tr>
<td>4</td>
<td>17</td>
<td>1</td>
<td>9</td>
<td>8</td>
<td>17</td>
</tr>
<tr>
<td></td>
<td>(31.8)</td>
<td>(6.8)</td>
<td>(3.2)</td>
<td>(2.1)</td>
<td>(8.0)</td>
</tr>
<tr>
<td>5</td>
<td>35</td>
<td>13</td>
<td>11</td>
<td>17</td>
<td>122</td>
</tr>
<tr>
<td></td>
<td>(121)</td>
<td>(26)</td>
<td>(12)</td>
<td>(8)</td>
<td>(31)</td>
</tr>
</tbody>
</table>

$\sum_{i,j} c_{ij} = 1280; \chi^2 = 716.6$

(the groups with $i,j = 3$ and with $i,j = 4$ are combined for calculation of $\chi^2$)

This proportion is presented in dependence on the increasing difference between $i$ and $j$ (cf. Fig.1). To extend the example, a sub-division into bibliography 1.1 and bibliography 1.2 is being used. The results are given in Fig.2. The results of the second bibliography and the results of the third bibliography are given in Fig.3.

The structure of coauthorship reflects the above mentioned social psychological theory.

The conclusion is drawn that the following main hypothesis should be investigated:

The sum of observed relations by coauthorships in proportion to the sum of expected ones is larger between authors of the same rank than between authors of different ranks ($i$ or $j$ = rank of authors in one group).

$\frac{\sum_{i,j} c_{ij}}{\sum_{i,j} m_{ij}} > \frac{\sum_{i\neq j} c_{ij}}{\sum_{i\neq j} m_{ij}}$

Table 5 shows the results for the symmetric matrices of all bibliographies.
$\sum_{i-j}^{C_{ij}}$, with $i-j = const$

Fig. 1: Stratification of authors in the bibliography No.1
$$\frac{\sum C_{ij}}{\sum m_{ij}}, \text{ with } i-j = \text{ const.}$$

Fig. 2: Stratification of authors in the bibliography No. 1.1 and in the bibliography No. 1.2

- full line: bibliography of all countries without the USA
- broken line: bibliography of the USA
\[ \frac{\sum_{i-j} C_{ij}}{\sum_{i-j} m_{ij}} \], with \( i-j = \text{const.} \)

Fig. 3: Stratification of authors in the bibliographies No. 2 and No. 3

- full line: bibliography No. 2
- broken line: bibliography No. 3
According to the Chi-square-test, coauthorships between authors do not come into being independently of their ranks. Error probability is smaller than 0.1 per cent. This applies to all bibliographies. In addition, the observed number of coauthorships between authors of the same rank is significantly higher than the expected one ($\chi^2_{ij} > 1.96$, significance on the 2.5 percent level).

This means

$$\sum C_{jj} \frac{\chi^2}{\sum m_{jj}} > 1.$$ 

For all bibliographies, the observed number of coauthorships between authors with different ranks is significantly smaller than expected. ($\chi^2_{ij} < -1.96,$
significance on the 2.5 percent level). This means

\[
\sum_{i \neq j} \frac{C_{ij}}{m_{ij}} < 1.
\]

Thus the main hypothesis can be verified for all bibliographies. It should be investigated whether the main hypothesis can also be verified when, in case of sufficiently comprehensive bibliographies, the rank of a group of scientists is equalled to the number of publications per scientist in any case, that is, no combination of groups of authors is made. This was investigated for the first, second and third bibliographies (cf. Table 6).

Table 6: Results of the symmetrized matrices

<table>
<thead>
<tr>
<th>Bibliography</th>
<th>(\sum_{i,j} C_{ij})</th>
<th>(\sum_{i,j} m_{ij})</th>
<th>(FT_{jj})</th>
<th>(\sum_{i \neq j} C_{ij})</th>
<th>(\sum_{i \neq j} m_{ij})</th>
<th>(FT_{ij})</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>744</td>
<td>513.3</td>
<td>9.25</td>
<td>536</td>
<td>766.7</td>
<td>-9.02</td>
</tr>
<tr>
<td>2</td>
<td>8454</td>
<td>6666.1</td>
<td>20.6</td>
<td>3788</td>
<td>5556</td>
<td>-26.28</td>
</tr>
<tr>
<td>3</td>
<td>584</td>
<td>282.2</td>
<td>14.7</td>
<td>1692</td>
<td>1993.6</td>
<td>-7.03</td>
</tr>
</tbody>
</table>

The results hardly differ from the results in Table 5. The reason is that the number of combined authors with a high number of publications per author in proportion to the number of all the other authors is such a small one that even statistical deviations exercise only a small effect. For this reason, the modified main hypothesis reads as follows:

The sum of observed relations by coauthorships in proportion to the sum of expected ones is higher between authors with the same number of publications per author than between authors with a different number of publications per author. (When \(i\) or \(j\) are equal to the number of publications per author of a group, the formula of the main hypothesis is valid too).

However, the social psychological theory shows not only that the frequency of contacts between partners of the same rank is the highest - which corresponds to the main hypothesis - but also that the frequency of contacts decreases with a growing difference between ranks. This tendency is shown in Fig.1, Fig.2 and Fig.3.
Thus, the following hypothesis reads:

The proportion of the sum of observed relations by coauthorships to the sum of expected ones show the following order: The proportion for authors with the same rank \((i-j = 0)\) is higher than the proportion for authors with the difference between the ranks \(i-j = 1\), which is higher than the proportion for authors with the larger difference between the ranks \(i-j > 1\) (cf. Table 7)

\[
\frac{\sum C_{ij}}{\sum m_{ij}} > \frac{\sum C_{i,j=1}}{\sum m_{ij}} > \frac{\sum C_{i,j>1}}{\sum m_{ij}}
\]

<table>
<thead>
<tr>
<th>Bibliography</th>
<th>(\frac{\sum C_{jj}}{\sum m_{jj}})</th>
<th>(\frac{\sum C_{i,j=1}}{\sum m_{ij}})</th>
<th>(\frac{\sum C_{i,j&gt;1}}{\sum m_{ij}})</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.59</td>
<td>0.82</td>
<td>0.44</td>
</tr>
<tr>
<td>1.1</td>
<td>1.46</td>
<td>1.032</td>
<td>0.465</td>
</tr>
<tr>
<td>1.2</td>
<td>1.66</td>
<td>0.73</td>
<td>0.43</td>
</tr>
<tr>
<td>2</td>
<td>1.28</td>
<td>0.72</td>
<td>0.59</td>
</tr>
<tr>
<td>3</td>
<td>2.03</td>
<td>0.91</td>
<td>0.7</td>
</tr>
<tr>
<td>4</td>
<td>1.45</td>
<td>1.07</td>
<td>0.68</td>
</tr>
<tr>
<td>5</td>
<td>1.2</td>
<td>0.81</td>
<td>(\nabla) 0.86</td>
</tr>
<tr>
<td>6</td>
<td>1.2</td>
<td>0.92</td>
<td>0.67</td>
</tr>
<tr>
<td>7</td>
<td>1.34</td>
<td>0.94</td>
<td>0.65</td>
</tr>
</tbody>
</table>

The results in Table 7 reflect this tendency generally. One exception is the bibliography of "Czechoslovak Journal of Physics". It is possible that there is a modification in case of extending the scope of this bibliography. On the other hand, results of an absolute conformity cannot be expected when presenting tendencies. It is possible that every expansion of the main hypothesis brings about larger deviations. In this context, the kind of combining groups of authors with a different number of publications per author is of growing importance and should be investigated in further studies. As indicated in "Data", it might be important to take the logarithm of the number of publications into account.

For the purpose of demonstration, the results of the non-symmetric matrices of the first, second and third bibliographies are shown in Fig. 4. Fig. 4 allows the assumption that generally, there is no fundamental difference between investigating symmetric and non-symmetric matrices. It only shows whether the behavior of coauthors exist in different bibliographies. In the articles of the first bibliography, coauthors with a higher number of publications have frequently their place before the authors with a smaller number of publications. In the third bibliography, there is the reverse order.
The stratification of authors was discussed in this paper in dependence on the number of publications per author, which is reflected in the structure of coauthorships. In preceding publications written by J. Cole and S. Cole [11], Oromaner [12], Snizek [13] and Green [14], the stratification of scientists was demonstrated in another way, that is, in dependence on the number of citations. The above mentioned authors found out that those authors of highly cited papers mainly tend to cite papers of other highly cited authors.

Independently of this fact, there exists a high correlation between the number of citations which a scientist receives, and the estimation of the quality of his work as well as a high correlation between the number of these citations and the number of his publications [15]. That is the reason why it is useful to compare results achieved by means of citations with results that could be found by means of coauthorships.

In contrast to the study by J. and S. Cole [11], the present paper
- involves not only selected (highly cited) papers but all available ones,
- is concentrated not only on elite but likewise on all available authors.

As a conclusion it was possible to integrate the whole social psychological theory (cf. Introduction) into our interpretation which says that stratification is an optimum form of contacts between the members of a group with rank orders which are of importance for the existence of the group as a whole, that is, not only for the elite.

The bibliographies which were used for investigating the structures of coauthorships were selected under varying conditions (cf. Data) under the condition of the analysis of the structure of coauthorships.

- in journals,
- in problem areas or specialties respectively independently of the place where the publications were published,
- in social sciences,
- in exact and life sciences,
- in bibliographies with varying interval of the years of publication,
- in bibliographies of a different scope,
- with different assignment of rank to the number of publication per author, that is, with different combinations of groups of authors.

The investigations should be continued systematically in order to answer the question under which special conditions the assumption

\[ \frac{\sum C_{ij}}{\sum M_{ij}} > \frac{\sum C_{ij}}{\sum M_{ij}} \]

is valid in bibliographies.

ACKNOWLEDGEMENTS

First I want to thank my mother that she found, when collecting basic data for the structure of citations, that properties characteristic of the structure of citations [16] also apply to the structure of coauthorships, that is, that authors with the same number of publications wrote articles together in coauthorship more frequently than authors with a different number of publications. Thus, she initiated the statistical structure analysis as presented in this paper.

REFERENCES


