I. LUKOVITS AND P. VINKLER

CORRECT CREDIT DISTRIBUTION: A MODEL FOR SHARING CREDIT AMONG COAUTHORS

(Accepted May 3, 1995)

ABSTRACT. On the basis of some assumptions (e.g. Minimum and Maximum Contribution Scores of authors) a simple equation is introduced for calculating individual contribution scores of co-authors of multi-authored papers. The calculated Corrected Contribution Scores are in good agreement with the data obtained empirically, earlier. It is suggested that individual percentage contributions should be declared by the co-authors in the by-line of papers.

INTRODUCTION

Multiple authorship is a common phenomenon in modern natural sciences (Garvey, 1979). The number of papers published in cooperation is steadily increasing (Schubert and Braun, 1990). This is one of the consequences of the mechanism of scientific research in the age of Big Science.

One of the fundamental questions concerning scientific publishing, namely *who* should be listed as an author of a paper is not easy to answer (Culliton, 1988). There are some ethical and publishing guidelines issued by scientific societies or editors of journals in order to regulate the various habits and interests of authors or potential authors (Maddox, 1990). Nevertheless, there exist no generally accepted rules.

Authorship has become an indicator widely applied for the evaluation of research performance. Evaluation of persons, teams, universities or countries by scientometric methods is based primarily on scientometric indicators. The majority of indicators applies the *number of publications* in various contexts. In the evaluations, however, *sharing of the credit of publications* among the cooperating partners (persons, teams or countries, or different disciplines etc.) is problematic. Possible solutions to the problem are discussed in

Social Indicators Research **36:** 91–98, 1995.

a previous paper (Vinkler, 1993) and in references Ajiferuke *et al.*, 1988 and Braun *et al.*, 1992.

There are two simple solutions for the distribution of the credit of papers: *first author counting and linear fractional authorship.* According to the first author counting method the total credit (or score) of the paper is attributed to the first author. This procedure extremely prefers the role of first authors and can be applied with success for large publication pools. The linear fractional authorship model takes the total credit of a paper as unity which is distributed equally according to the number of authors (1/N).

The present paper aims at introducing a new model for sharing credits of papers among researchers (research teams or countries etc.). The model presented here is named Correct Credit Distribution (CCD) model.

CORRECT CREDIT DISTRIBUTION MODEL

Basic Assumptions of the CCD Model

(1) The *quantum of published information* in natural sciences is the scientific paper. (Under the expression 'paper' full paper, review, short communication and letter are meant.) Consequently, total credit of a paper is assumed to be unity.

(2) All persons who made larger *research contribution* to the paper than the '*authorship threshold*' are co-authors of the paper (Vinkler, 1987).

(3) Persons *without* any research *contribution* are not co-authors.

(4) The *rank* of the co-authors *corresponds* to their *contribution* to the paper.

(5) The Minimum Contribution Score (ICS) for the *first authors* of papers with any number of co-authors is equal to the share calculated by the linear fractional authorship model (i.e. 0.50; 0.33; 0.25; 0.20; 0.16 for two, three, four, five and six-authored papers, respectively). The ICS value for all other co-authors is 0.1 (see Table I).

(6) The Maximum Contribution Score (ACS) of any co-author, except for first authors, is equal to the share calculated by the linear fractional authorship model. For first authors ACS is assumed to be unity (see Table I).

(7) Contribution shares of authors between ICS and ACS values are equally probable, i.e. a uniform distribution exists. Consequently, the individual contribution scores of authors can be approximated by the arithmetical means of ICS and ACS values.

Remarks to the Basic Assumptions

In natural sciences about 70 per cent of new information is published in scientific papers (Garvey, 1979). Consequently, it is reasonable to accept papers as main information carriers.

Main activities for producing a paper in natural sciences may be listed as follows (Vinkler, 1993):

- experimental work,
- analysis and evaluation of data,
- supervising the research work,
- writing the text of the paper,
- literature search,
- recommendation of the topic.

Each activity or all of these can be performed by a single person or a whole team consisting of several persons.

The measure of the *authorship threshold* which is needed for being co-author of a chemistry paper in the score system suggested (Vinkler, 1993) was assumed to be two per cent of the total of activities mentioned. However, results of the questionnaire method applied (Vinkler, 1993) revealed that a research contribution of about *10 per cent* is the precondition of being listed as a co-author for papers having 2–5 co-authors. Nevertheless, for 'connectional' reasons (i.e. with little or no research contribution) this limit can decrease to zero. As connectional motivations primarily 'honorary' or/and 'compulsory' authorships are meant in this respect (Croll, 1984).

TABLE I

Minimum (ICS), Maximum (ACS), Uncorrected Mean (UMS) and Corrected Contribution Scores (CCS) for two, three, four, five and six authored papers

Rank of authors	ICS	ACS	UMS	CCS		CCS′
				Theoretical[1]	Empirical[2]	Theoretical[3]
1	0.50	1.00	0.750	0.714	0.71	0.750
2	0.10	0.50	0.300	0.286	0.29	0.250
1	0.33	1.00	0.655	0.563	0.61	0.615
2	0.10	0.50	0.300	0.254	0.26	0.231
3	0.10	0.33	0.215	0.183	0.13	0.154
1	0.25	1.00	0.625	0.475	0.54	0.536
2	0.10	0.50	0.300	0.228	0.31	0.214
3	0.10	0.33	0.215	0.165	0.09	0.143
4	0.10	0.25	0.175	0.133	0.06	0.107
1	0.20	1.00	0.600	0.416	0.34	0.483
2	0.10	0.50	0.300	0.208	0.14	0.201
3	0.10	0.33	0.215	0.150	0.11	0.134
4	0.10	0.25	0.175	0.121	0.17	0.101
5	0.10	0.20	0.150	0.104	0.24	0.081
1	0.16	1.00	0.583	0.374		0.446
2	0.10	0.50	0.300	0.193		0.191
3	0.10	0.33	0.215	0.139		0.127
4	0.10	0.25	0.175	0.112		0.096
5	0.10	0.20	0.150	0.096		0.076
6	0.10	0.16	0.130	0.086		0.064

Remarks

[1] Theoretical CCS values were calculated by the Correct Credit Distribution Model (CCD) presented in this paper applying *0.10* as Minimum Contribution Scores (ICS) for co-authors ($k \geq 2$).

[2] Empirical CCS values are from P. Vinkler, 1993 under the name Total Contribution Factor (TCF).

[3] Theoretical CCS′ values were calculated by the Correct Credit Distribution Model (CCD) applying *zero* as Minimum Contribution Scores (ICS) for co-authors ($k \geq 2$).

The shares of authors (N) in the linear fractional authorship model can be obtained by dividing unity by the number of the authors

(1/N). The resulting shares are supposed to be the possible Minimum Contribution Scores (ICS) for first authors. For the co-authors of papers with ten authors or less the lowest threshold is assumed here to be 10 per cent. For more authors (up to 20) the ICS value can be reduced to 5 and from 20 to 100 authors to 1 per cent. Naturally, the percentage values of the contribution threshold (ICS) can be chosen arbitrarily in the CCD model.

The greatest contribution share (ACS) for first authors is assumed to be unity (100 per cent), whereas that for second, third, fourth, fifth, sixth etc. authors 0.50, 0.33, 0.25, 0.16, etc., respectively regardless of the total number of authors.

CALCULATION OF MEAN UNCORRECTED AND CORRECTED CONTRIBUTION SCORES

Minimum, Maximum and Uncorrected Mean Contribution Scores (ICS, ACS, UMS) are given in Table I for first authors and co-authors of two, three, four, five and six authored papers.

From assumptions 5, 6 and 7 it follows that Uncorrected Mean Contribution Score (UMS) values for first authors (UMS(1)) and co-authors (UMS(k)) can be calculated as arithmetic means of the respective Minimum (ICS) and Maximum Mean Contribution Score (ACS) data by Eqs. 1 and 2, respectively.

$$(1) \qquad \mathrm{UMS}(1) = \frac{\mathrm{ICS}(1)+\mathrm{ACS}(1)}{2} = \frac{N+1}{2N}$$

$$(2) \qquad \mathrm{UMS}(k) = \frac{\mathrm{ICS}(k)+\mathrm{ACS}(k)}{2} = \frac{1}{2k} + \frac{1}{2T}$$

where N is the number of authors, k is the rank of authors ($k \geq 2$) and $T = 100/\mathrm{H}$, where H is the percentage value of the contribution threshold. In this paper $H = 10$ per cent is applied, from which $T = 10$ follows. Note that if we do not consider lower limits of the contribution (ICS), i.e. $H = 0$, $T = \infty$ and $1/T = 0$, $\mathrm{UMS}(k) = 1/2k$ follows from Eq. 2.

The sum of UMS(1) and UMS(k) ($k = 2, \ldots, N$) values is greater than unity (Table I). Consequently, UMS data should be normalized. The normalization factor (F), which is equal to the sum of contribution of all authors, can be calculated by Eq. 3.

$$(3) \qquad \mathrm{F} = \frac{N+1}{2N} + \sum_{k=2}^{N}\left(\frac{1}{2k} + \frac{1}{2T}\right) = \frac{1}{2}\left[\frac{1}{N} + \frac{N-1}{T} + \sum_{k=1}^{N}\frac{1}{k}\right]$$

The normalized contribution scores (Corrected Contribution Score; CCS) for first authors (CCS(1)) and for the k-th co-authors (CCS(k)) can be calculated by Eqs. 4 and 5, respectively.

$$\text{(4)} \qquad \mathrm{CCS}(1) = \frac{N+1}{2NF}$$

$$\text{(5)} \qquad \mathrm{CCS}(k) = \frac{k+T}{2kFT}$$

The CCS values calculated for two, three, four, five and six-authored papers are given in Table I.

Let us present an example here, in which the number of authors is $N = 3$ and authorship threshold is $T = 10$. The normalization factor (F) can be calculated on the basis of Eq. 3 by Eq. 6.

$$\text{(6)} \qquad \mathrm{F} = \tfrac{1}{2}\left[\tfrac{1}{3} + \tfrac{3-1}{10} + \tfrac{1}{1} + \tfrac{1}{2} + \tfrac{1}{3}\right] = 1.183$$

The corrected contribution of the first author (CCS(1)) using Eq. 4, is given by Eq. 7.

$$\text{(7)} \qquad \mathrm{CCS}(1) = \frac{3+1}{2 \cdot 3 \cdot 1.183} = 0.563$$

The normalized contribution scores for the second and the third authors (Eq. 5; CCS(2), (3)), are given by Eqs. 8 and 9, respectively.

$$\text{(8)} \qquad \mathrm{CCS}(2) = \frac{2+10}{2 \cdot 2 \cdot 1.183 \cdot 10} = 0.254$$

$$\text{(9)} \qquad \mathrm{CCS}(3) = \frac{3+10}{2 \cdot 3 \cdot 1.183 \cdot 10} = 0.183$$

Table I shows that the data obtained empirically by the questionnaire method (Vinkler, 1993) do not differ greatly from those calculated by the Correct Credit Distribution (CCD) model, except for fourth and fifth authors of four and five-authored papers, respectively. The mentioned difference may be attributed to the increased number of heads among the fourth and fifth authors as compared to the total (55.0 per cent vs. 29.4 per cent). The fifth authors e.g. declared their activity in *supervising* the research work and offering the *concept* or idea for the investigations significantly greater than all other co-authors except first ones. The share of the co-authors by rank was found to be in supervising and offering the concept as follows 3.95; 1.30; 0.10; 0.83; 2.83 and 3.47; 1.35; 0.10; 0.33; 3.33,

respectively. The dynamic range of both factors was set from zero up to five (Vinkler, 1993).

The CCD-model suggested here would be more consequent by applying 0.9, 0.8, 0.7, 0.6, 0.5 etc. instead of unity as the maximum contribution shares for first authors (ACS(1)) of two, three, four, five and six-authored papers, respectively. This would be a consequence of the lowest authorship threshold (i.e. minimum contribution, ICS) for co-authors applied in this paper (0.1). With the mentioned ACS values, however, for seven and higher-number authored papers co-authors would receive greater contributions than first authors (e.g. UMS(1) = 0.221 whereas UMS(2) = 0.231). Therefore, ACS(1) of unity is allowed for first authors. The application of unity as ACS(1) and the normalization process cause e.g. that the contribution (CCS) of fifth and sixth authors of six-authored papers is lower than 0.1.

Table I also gives the data of Correct Contribution Scores (CCS′) calculated by applying $H = 0$ instead of $H = 10$ as contribution threshold (ICS) for co-authors. The CCS′ values differ only slightly from the respective CCS data. The CCS′(1) (theoretical) data (shares of first authors) are greater than CCS(1) (theoretical) ones, all other CCS′ values (shares of co-authors) are lower than the respective CCS scores.

CONCLUSION

The main goal of the present paper was to introduce a model for calculating shares of credit for co-authors and to compare the data calculated with those obtained earlier empirically (Vinkler, 1993).

The percentage shares suggested by the CCD model presented (CCS values in Table I) could serve as standards for co-authors. The shares calculated can give an orientation for considering the rank of co-authors.

Suggestion for co-authors to declare correct contribution shares

On the basis of our findings we suggest that papers submitted for publication should be provided with the information as follows:

- *Measure of contributions to the paper by each co-author in the by-line by declaring percentage ratios of the individual*

contributions. For example: A. Fair, B. Honest, C. Frank (50–30–20).

Authorship, ranking of co-authors and contribution per cents should be the object of an open discussion of the researchers themselves having taken part in the work resulted in publication.

We do hope that the declaration of contribution per cents by co-authors may contribute to an improved ethical environment of scientific research and would make calculations of publication potentials for science policy makers much more correct.

REFERENCES

Ajiferuke, I., Q. Burrel and J. Tague: 1988, 'Collaborative coefficient: A single measure of the degree of collaboration in research', Scientometrics 14(5–6), pp. 421–433.

Braun, T., J. Gomez, A. Mendez and A. Schubert: 1992, 'International co-authorship patterns in physics and its subfields, 1981–1985', Scientometrics 24(2), pp. 181–200.

Croll, R. P.: 1984, 'The noncontributing author: An issue of credit and responsibility', Perspectives in Biology and Medicine 27(3), pp. 401–407.

Culliton, B. J.: 1988, 'Authorship, data ownership examined', Science 242, p. 658.

Garvey, W. D.: 1979, Communication: The Essence of Science (Pergamon Press, Oxford, etc.).

Maddox, J.: 1990, 'Should camp-followers be policemen?', Nature 348, p. 107.

Schubert, A. and T. Braun: 1990, 'International collaboration in the sciences, 1981–1985', Scientometrics 19(1–2), pp. 3–10.

Vinkler, P.: 1987, 'A quasi-quantitative citation model', Scientometrics 12(1–2), pp. 47–72.

Vinkler, P.: 1993, 'Research contribution, authorship and team cooperativeness', Scientometrics 26(1), pp. 213–230.

Central Research Institute for Chemistry,
Hungarian Academy of Sciences,
Budapest,
Hungary