Methods for accrediting publications to authors or countries: consequences for evaluation studies

Peer-reviewed author version

EGGHE, Leo; ROUSSEAU, Ronald & Van Hooydonk, G. (2000) Methods for accrediting publications to authors or countries: consequences for evaluation studies. In: Journal of the American Society for Information Science, 51(2). p. 145-157.

DOI: 10.1002/(SICI)1097-4571(2000)51:2<145::AID-ASI6>3.0.CO;2-9
Handle: http://hdl.handle.net/1942/792

Methods for accrediting publications to authors or countries: consequences for evaluation studies

Leo Egghe*, Ronald Rousseau* and Guido Van Hooydonk**

*LUC, Universitaire Campus, B-3590 Diepenbeek, Belgium

and UIA, IBW, Universiteitsplein 1, B-2610 Wilrijk, Belgium

legghe@luc.ac.be ronald.rousseau@kh.khbo.be

** University of Ghent, Department of Library Sciences,

Rozier 9, B-9000 Ghent, Belgium

guido.vanhooydonk@rug.ac.be

Abstract

One aim of science evaluation studies is to determine quantitatively the contribution of different players (authors, departments, countries) to the whole system. This information is then used to study the evolution of the system, for instance to gauge the influence of special national or international programs. Taking articles as our basic data, we want to determine the exact relative contribution of each co-author or each country. These numbers are then brought together to obtain country scores, or department scores, etc. It turns out, as we will show in this article, that different scoring methods can yield totally different rankings. In addition to this, a relative increase according to one method can go hand in hand with a relative decrease according to another counting method. Indeed, we present examples in which country (or author) c has a smaller relative score in the total counting system than in the fractional counting one, yet this smaller score has a higher importance than the larger one (fractional counting). Similar anomalies were constructed for total versus proportional counts and for total versus straight counts. Consequently, a ranking between countries, universities, research groups or authors, based on one particular accrediting method does not contain an absolute truth about their relative importance. Different counting methods should be used and compared. Differences are illustrated with a real-life example. Finally, it is shown that some of these anomalies can be avoided by using geometric instead of arithmetic averages.

I. Introduction

One aim of science evaluation studies is to determine quantitatively the contribution of different players (authors, departments, countries) to the whole system. This information is then used to study the evolution of the system, for instance to gauge the influence of special national or international (e.g. the European Union) programs. Taking articles as our basic data, we want to determine the exact relative contribution of each co-author or each country. These numbers are then brought together to obtain country scores, or department scores, etc. Making evaluations would be easier if any method of scoring would yield the same rank. It turns out, however, as we will show in this article, that different scoring methods can yield totally different rankings. In addition to this, a relative increase according to one method can go hand in hand with a relative decrease according to another counting method. This is one aspect of the 'multiple-author-problem' as reviewed by Harsanyi (1993).

What is the idea behind the use of one particular method of counting? Probably, a certain method is used because the evaluator thinks that this specific method reflects best the contribution of each collaborator to the article. Note that we assume that co-authors have not explicitly stated their exact contribution (as is the practice nowadays), but that the order in which the co-authors are mentioned gives an indication about their relative impact on the whole study. This also implies that pure alphabetical ordering (as used for this article) is not considered here. Indeed, alphabetical ordering is ambiguous as it can mean two things: it is either a custom, implying nothing about the relative contribution of each co-author, or it is an indication that all contributions are equal.

Anyway, a counting procedure can be seen as an estimation method to determine the real, but unknown, relative contribution of each co-author.

Using different accrediting methods is no longer a purely academic or mathematical problem: academic careers get more and more dependent on bibliometric evaluations of the quality of research (de Bruin et al., 1993; Reed, 1995; Rousseau, 1998; Toutkoushian, 1994; Van Hooydonk, 1998).

We will now recall the most used methods of counting (Lindsey, 1980; Egghe & Rousseau, 1990).

(1) First-author counting (Cole & Cole, 1973)

Only the first of the A authors of a paper receives a credit equal to one. The other authors do not receive any credit. This method is also known as straight counting. It should be seen as an easy-to-use sampling method. One hopes that, e.g., an author who has written n papers with on average 2 co-authors, is first author in one third of the cases. This method is popular mainly because the Science Citation Index used to present citation data (not publication data) by first author only (this has been improved in the Web of Science).

(2) Total author counting

Here, each of the A authors receives one credit. This counting method is also called normal, or standard counting.

(3) Fractional counting (Price, 1981)

Here, each of the A authors receives a score equal to 1/A. This counting method is sometimes called adjusted counting. Fractional counting has been studied e.g. in (Burrell and Rousseau, 1995; Egghe, 1996; Van Hooydonk, 1997).

To these well-known counting methods we will add a few lesser-used or new ways of counting relative contributions.

(4) Proportional counting (Van Hooydonk, 1997)

If an author has rank R in the author list of an article with A collaborators (R = 1, ..., A), then she/he receives a score of

$$\frac{2}{A}\left(1-\frac{R}{A+1}\right) \qquad (1)$$

Note that equation (1) is determined in such a way that the sum of all scores is equal to 1. Before this normalization the score of the collaborator at rank R was A+1-R.

(5) Pure geometric count (a new suggestion)

If an author has rank R in an article with A co-authors (R = 1,...,A) then she/he receives a credit of:

$$\frac{2^{A-R}}{2^{A}-1} \qquad (2)$$

This counting method too is normalized so that the sum of all relative contributions is 1. Before normalization the score for the author at rank R was 2^{A-R}.

(6) Noblesse oblige, cf. (Zuckerman, 1968)

Here the most important author closes the list. She/he receives a credit of 0.5, while the other A-1 authors receive a credit of 1/(2(A-1)) each (this is but one suggestion, among many more that are possible here).

(7) A variation of total counting could be termed 'absolute country counting' (our terminology). Using this method, a country receives at most one credit (depending on whether or not one of the authors works in this country). In (Nederhof & Moed, 1993) this method was called the on-line fractionation approach.

For counting methods (2), (3) and (7) rankings play no role. For the other ones it does, fully (4),(5) or partially (1),(6).

In this article we will study several problems. First, we want to know how the relative score of an author or country is affected by the counting procedure used. Note that when making comparisons only relative scores are important. Absolute scores are incomparable as the total sum of weights differs between methods that use normalization (i.e. the total weight is equal to the number of articles in the system) and methods that do not. This means that we are interested in the total score of each author (country, etc.) within a certain counting method, divided by the total sum of scores given in this counting method. Counting methods (5) and (6) are just given as suggestions. In this article they will not be studied further. Also, absolute country counting (7) will not be studied here (although it was used e.g. in (Nederhof & Moed, 1993; Ojasoo, 1996). In Ojasoo's article it was stated that this method led to an increase of 9% with respect to a method that uses normalization (each article weights one unit). Anderson et al. (1988) also discussed absolute country counting, where an increase of 7.3% has been reported. The difference between these two values can easily be explained by an increased worldwide collaboration tendency over the last decades.

Section II illustrates the problems with different counting methods on a real-life problem. In Section III we will develop formulae for relative scores. We will show theoretically (presenting exact formulae) and by examples, that it is possible to have a larger score for author (or country) a than for author (or country) b for one counting method, while the opposite is true for another one. As a consequence of this, changing the counting procedure reverses the relative importance, expressed by the obtained ranks.

In Section IV we, moreover, illustrate the following paradox. It is possible to have an increase (between methods) of the relative score of an author while at the same time this author occurs lower in the ranked list (an increase of the rank of this author). Stated otherwise, this author becomes more important according to her/his relative score and less important according to her/his rank, and this using the same two counting methods! Examples of this behavior are presented.

In Section V we search for possible solutions for these anomalies. It is first noted that the results obtained with the total, fractional and proportional counting method are close to each other in case we are dealing with systems in which collaboration (number of co-authors, denoted as A) is high. Furthermore, if one replaces arithmetic averages by geometric ones, anomalies disappear. This has already been remarked in (Egghe & Rousseau, 1996b). Indeed, by using geometric averages fractional counting is the same as total counting, in the sense that relative scores become equal. Hence, there cannot be any change of rank between these methods.

II. A real-life example

In this section we illustrate the problems associated with different counting methods. The data refer to citations of a sample of multi-authored papers, published by a biotechnology research group at the University of Ghent. The senior researcher, M. Van Montagu is a well-reputed scientist in his field, who generally applies the 'noblesse oblige' principle. We note that e.g. A.F.J. Van Raan (CWTS, Leiden, The Netherlands) practices the same principle in the field of scientometrics-informetrics. This analysis answers some of the questions raised by Oppenheimer (1998), concerning citation counts for senior scientists and illustrates the differences between accrediting methods.

In this sample of 36 multi-authored articles by Van Montagu's group, 4 were published in 1993 and 32 in 1994. Citation counts were collected in October 1997. The impact is defined as C/P (citations per publication). We have always used the same counting method for both citations and publications.

Insert Figs 1 and 2 about here

Fig. 1 represents the overall view with respect to impact and counting procedures. Data on the senior author are given first. Sometimes the term 'senior author' is used for 'first author'. We do not agree with this terminology and do not follow it. Although the trends in the three impact values are similar, the deviations are larger for proportional counting than for fractional counting. Reversals in impact are observed for almost every author. More detail becomes visible in Fig. 2, where the impact obtained by fractional and proportional counting has been scaled by the impact as calculated by total counting (the de facto standard method used for scientometric evaluations).

For the three cases (d, g and h) where the deviations from total count are large, the total count impact is nearly the same and almost equals that of the senior scientist (about 11), cf. Fig. 1. Authors d and h in Figs. 1 and 2 rank invariantly first and second on highly cited papers. Author g ranks second on less cited papers and only fifth on a highly cited paper. Author j, the one with the lowest over-all impact (see Fig. 1), sees his proportional impact increase (relatively), which is due to the fact that he is the last author in the three papers.

We conclude that counting methods may have important practical consequences for career opportunities. Proportional counting adds additional detail into the processing of citation data. We notice, for instance, that first authors on highly cited papers invariantly get a higher proportional impact than their co-authors. A conditio sine qua non is that the rank of an author on a multi-authored paper is a reflection of his/her real contribution, the general validity of which remains a problem. The impact of the senior author seems almost always independent of the used counting method, at least in this sample. This result could have been anticipated intuitively, and answers, in part, some of Oppenheimer's remarks on this question (Oppenheimer, 1998).

III. Formulae for counting procedures

In this section we will first introduce notation and terminology. Then we derive formulae for the relative contribution of countries, depending on the used counting method. The term 'country' must be interpreted in a general way: it can be a real country or a university, a department or even an author. Finally we give an interpretation of some of these results.

III.1 Notation

We will consider a system of N articles, the total number of authors of article i is denoted as $a_i > 0$ (excluding anonymous articles). Each author of an article is assumed to work in exactly one country. Let the total number of entries be denoted as E, see further (6).

If country c has an occurrence on rank r in article i, $d_{irc} = 1$, otherwise d_{irc} is 0. Then the following relations hold:

$\sum_{r=1}^{a_i} d_{irc} = a_i(c)$ denotes the number of occurrences of country c in article i (3)

$\sum_{i,r} d_{irc}$: the total number of occurrences of country c (4)

$\sum_{r,c} d_{irc} = a_i$: the number of authors of article i (5)

$\sum_{i,r,c} d_{irc}$ = E : total number of entries in the system (6)

$\sum_{i,c} d_{i1c} = N$: total number of articles (7)

$\sum_{r,c} r\, d_{irc} = \frac{a_i(a_i+1)}{2}$: sum of all ranks of article i (8)

$\sum_{r=1}^{a_i} r\, d_{irc} = R(i,c)$: sum of all ranks occupied by country c in article i (9)

We note that the sum of all ranks occupied by any country (c) in any article (i) is always smaller than the sum of all ranks in this article. Mathematically, this is:

$$1 \leq R(i,c) \leq \frac{a_i(a_i+1)}{2} \qquad (10)$$

If c is interpreted as an author $a_i(c) = 0$ or 1. In this way the important 'author case' follows from the 'country case' as studied here.

With the three-dimensional matrix $D = (d_{irc})_{irc}$ we will associate new three-dimensional matrices, depending on the counting method we want to study. These matrices are called the characteristic matrices of the counting method. Use of this notion allows us to define other (weighting) variables such as W, W(c) and Q(c) (see further for their definition) in the same way.

For straight or first-author counting we will use the matrix $S(D) = (s_{irc})_{irc}$, with

$$s_{irc} = \delta_{1,r}\, d_{irc} \qquad (11)$$

where δ denotes the Kronecker delta, defined as: $\delta_{k,l} = 1$ if $k = l$ and $\delta_{k,l} = 0$ if $k \neq l$.

For total counting we use T(D) = D, with

$$t_{irc} = d_{irc} \qquad (12)$$

For fractional counting we need F(D) = $(f_{irc})_{irc}$, with

$$f_{irc} = \frac{d_{irc}}{a_i} \qquad (13)$$

Finally, for proportional counting, we use P(D) = $(p_{irc})_{irc}$, with

$$p_{irc} = \frac{2\,d_{irc}}{a_i}\ \frac{a_i + 1 - r}{a_i + 1} \qquad (14)$$

III.2 Total weight of a scoring system

The total weight of a scoring system is denoted as W, and is defined as the sum of all elements of its characteristic matrix. For the four scoring systems under study the total weights are as follows.

Straight (or first-author) counting:

$$W_S = \sum_{irc} s_{irc} = \sum_{ic} d_{i1c} = N \qquad (15)$$

where we have used equations (11) and (7).

Total counting:

$$W_T = \sum_{irc} t_{irc} = \sum_{irc} d_{irc} = \sum_{i} a_i = E \qquad (16)$$

where we have used (12), (5) and (6).

Fractional counting:

$$W_F = \sum_{irc} f_{irc} = \sum_{irc} \frac{d_{irc}}{a_i} = \sum_{i=1}^{N} \frac{a_i}{a_i} = N \qquad (17)$$

where we have used (13) and (5).

Proportional counting:

$$W_P=\sum_{irc}p_{irc}=\sum_{irc}\frac{2d_{irc}(a_i+1-r)}{a_i(a_i+1)}=2\sum_i\frac{\sum_{rc}d_{irc}}{a_i}-2\sum_i\frac{\sum_{rc}rd_{irc}}{a_i(a_i+1)}=2N-2\sum_{i=1}^{N}\frac{1}{2}=N \quad (18)$$

as it, of course, should be. Here we have used (14), (5) and (8).

III.3 Total score of country c

The total score of country c in a certain scoring system is denoted as W(c), and is defined as the sum of all elements of row c of its characteristic matrix. This gives the following formulae.

Straight (or first-author) counting:

$$W_S(c)=\sum_{ir}s_{irc}=\sum_{i=1}^{N}d_{i1c} \qquad (19)$$

Total counting:

$$W_T(c)=\sum_{ir}t_{irc}=\sum_{ir}d_{irc}=\sum_{i=1}^{N}a_i(c) \qquad (20)$$

where we have used (3).

Fractional counting:

$$W_F(c)=\sum_{ir}f_{irc}=\sum_{ir}\frac{d_{irc}}{a_i}=\sum_{i=1}^{N}\frac{a_i(c)}{a_i} \qquad (21)$$

where, again, we have used (3).

Proportional counting:

$$W_P(c)=\sum_{ir} p_{irc}=\sum_{ir}\frac{2d_{irc}(a_i+1-r)}{a_i(a_i+1)}=2\sum_{ir}\frac{d_{irc}}{a_i}-2\sum_{ir}\frac{rd_{irc}}{a_i(a_i+1)}=\sum_{i=1}^{N}\frac{2}{a_i}\left(a_i(c)-\frac{R(i,c)}{a_i+1}\right) \quad (22)$$

where we have used equations (3) and (9).

III.4 Relative contribution of a country

Finally, we determine, for the four counting methods, the relative contribution of country c. Relative contribution is denoted as Q(c), and is obtained as:

$$0 \leq Q(c)=\frac{W(c)}{W} \leq 1 \qquad (23)$$

For the four counting methods this yields:

$$Q_S(c) = \frac{\sum_{i=1}^{N} d_{i1c}}{N} \qquad (24)$$

This is the relative number of occurrences of country c on rank one.

For total counts we obtain:

$$Q_T(c)=\frac{\sum_{i=1}^{N} a_i(c)}{\sum_{i=1}^{N} a_i} = \frac{\sum_{i=1}^{N} a_i(c)}{E} \qquad (25)$$

For fractional counts this becomes:

$$Q_F(c)=\frac{1}{N}\sum_{i=1}^{N}\frac{a_i(c)}{a_i} \qquad (26)$$

and finally, for proportional counting:

$$Q_P(c) = \frac{1}{N}\sum_{i=1}^{N}\frac{2}{a_i}(a_i(c) - \frac{R(i,c)}{a_i+1}) = 2\,Q_F(c) - \frac{1}{N}\sum_{i=1}^{N}\frac{2\,R(i,c)}{a_i(a_i+1)} \qquad (27)$$

The second term in (27) can be interpreted as the sum of all relative shares of country c in all ranks of article i.

III.5 Interpretation and discussion of these formulae

III.5.1 Q_T versus Q_F

The difference between $Q_T(c)$ versus $Q_F(c)$ has been studied e.g. in (Egghe and Rousseau, 1996a,b), where the following result was proved:

Theorem. $Q_T(c) >$ (resp. $<$) $Q_F(c)$ if and only if the slope of the regression line of

$\frac{a_i(c)}{a_i}$ over a_i ($i = 1,\ldots, N$) is strictly positive (resp negative).

This result makes it clear that there must exist countries c and c' for which

$$Q_T(c) > Q_F(c) \text{ and } Q_T(c') < Q_F(c')$$

within one group of countries of which the publication output is investigated. In fact, it is extremely simple to construct such an example.

Example 1

Let c and c' denote two countries with a co-authorship distribution as indicated in the following table:

	c	c'
article 1	x	
article 2	x	x

Here $Q_T(c) = 2/3$, $Q_T(c') = 1/3$, while $Q_F(c) = 3/4$ and $Q_F(c') = 1/4$. Hence

$$Q_T(c) < Q_F(c) \quad \text{while } Q_T(c') > Q_F(c').$$

These differences in behavior can be considered as 'normal' or 'expected' since we deal here with different methods of measuring relative scores. In the next section, however, we will construct examples, showing undesirable (or at least unexpected) properties of the comparison of both counting methods. Since these examples are non-trivial we have reserved a special section for it (see Section IV).

III.5.2 Q_F versus Q_P

Considering (27) we may expect to find countries c and c' for which $Q_P(c) < Q_F(c)$ and $Q_P(c') > Q_F(c')$. Indeed, take a country c for which $R(i,c) = a_i > 1$ for all $i = 1,\ldots,N$ (this country occurs exactly once in each article, namely at the last rank). Then, by (27):

$$Q_P(c)=2Q_F(c)-\frac{1}{N}\sum_{i=1}^{N}\frac{2}{a_i+1} < 2Q_F(c)-\frac{1}{N}\sum_{i=1}^{N}\frac{1}{a_i} \qquad (28)$$

As $a_i(c) = 1$, we see from (28) that $Q_P(c) < 2Q_F(c) - Q_F(c) = Q_F(c)$.

Now take a country c' for which, for all $i = 1,\ldots,N$, $R(i,c') = 1$ (hence, this country always occupies the first place and has no author at other places). Then:

$$Q_P(c')=2Q_F(c')-\frac{1}{N}\sum_{i=1}^{N}\frac{2}{a_i(a_i+1)}>2Q_F(c')-\frac{1}{N}\sum_{i=1}^{N}\frac{1}{a_i} \qquad (29)$$

Hence, (29) implies that (since again $a_i(c') = 1$), $Q_P(c') > Q_F(c')$.

Note

If, in a certain article, every country occurs exactly once, i.e. $a_i(c) = 1$ (with i fixed and for a_i countries c) then the average relative proportional score of all these countries is:

$$\overline{Q_P} = \frac{1}{a_i}\sum_{j=1}^{a_i}\frac{2}{a_i}\left(1-\frac{j}{a_i+1}\right) = \frac{1}{a_i}\frac{2}{a_i}\left(a_i - \frac{a_i(a_i+1)}{2(a_i+1)}\right) = \frac{1}{a_i} = Q_F(c) \qquad (30)$$

for all c.

This value can even been reached by a concrete country. Indeed, take a_i odd and consider the country c occupying rank $(a_i+1)/2$. Then:

$$Q_P(c) = \frac{2}{a_i}\left(1-\frac{a_i+1}{2(a_i+1)}\right) = \frac{1}{a_i} = Q_F(c) \qquad (31)$$

These elementary ideas lead to the following theorem.

Theorem

For any system of collaboration between countries (institutions, authors,...), denoted as system I, we can construct another one, denoted as system II, such that the proportional counting result in system II equals the fractional counting result in system I, as well as in system II. Furthermore, also total counting and straight counting results are the same in both systems. In symbols:

$$Q_F^{I}(c) = Q_F^{II}(c) = Q_P^{II}(c), \qquad (32)$$

$$Q_T^{I}(c) = Q_T^{II}(c) \qquad (33)$$

$$\text{and } Q_S^{I}(c) = Q_S^{II}(c) \qquad (34)$$

Proof.

Take any system of N articles. Consider the i-th article. In this article country c appears $a_i(c)$ times. The ranks of the authors (each working in a certain, not necessarily different, country) are $1,2,\ldots,a_i$. Now mirror this situation so that we obtain an article with $2a_i$ authors, where, if country c has an author at rank r in system I, it also has an author at rank $2a_i - r + 1$. Fig.3 illustrates this construction.

Insert Fig. 3 about here

Do this for every article. For each country c the number of occurrences in the i-th article is now $2a_i(c)$. We will denote by $Q_X^{I}(c)$ and $Q_X^{II}(c)$, the relative scores – counted according to X = S,T,F or P - of country c in the original system (system I) and in the system where all articles have the double number of authors (system II, obtained by the mirroring operation).

We will now prove that for every country: $Q_F^{I}(c) = Q_F^{II}(c) = Q_P^{II}(c)$.

For every c we have, by definition and construction:

$$Q_P^{II}(c) = \frac{1}{N}\sum_{i=1}^{N}\frac{2}{2a_i}\left(2a_i(c) - \frac{R^{II}(i,c)}{2a_i+1}\right),$$

where $R^{II}(i,c)$ denotes the sum of the ranks occupied by c in article i in the system II. By construction, this is always equal to $(2a_i + 1)a_i(c)$. Hence,

$$Q_P^{II}(c) = \frac{1}{N}\sum_{i=1}^{N}\frac{2}{2a_i}\left(2a_i(c) - \frac{(2a_i+1)a_i(c)}{2a_i+1}\right)$$

$$= \frac{1}{N}\sum_{i=1}^{N}\frac{a_i(c)}{a_i} = Q_F^{I}(c) = \frac{1}{N}\sum_{i=1}^{N}\frac{2a_i(c)}{2a_i} = Q_F^{II}(c)$$

Finally, we also note that, clearly

$$Q_T^{II}(c) = \frac{\sum_{i=1}^{N} 2a_i(c)}{\sum_{i=1}^{N} 2a_i} = Q_T^{I}(c)$$

and, because of the number of times a country occurs first has not changed,

$$Q_S^{II}(c) = Q_S^{I}(c)$$

This proves the theorem.

This result will enable us to construct anomalous examples for Q_T versus Q_P, whenever we find anomalous examples for Q_T versus Q_F. This will be done in the next section.

We finally note that it is easy to obtain anomalies involving first-author count. Some examples are presented in the appendix. In the extreme case, an author who always applies alphabetical name ordering, or an important author who generally let junior collaborators be first author will not even show up in straight order rankings. A case in point is Van Raan's absence on the White-McCain co-citation map (White & McCain, 1998).

IV Non-trivial examples of anomalies between scores of total and fractional scores, and between scores of total and proportional scores

In this section we will explore whether it is possible to have systems such that

$$Q_T(c) > Q_T(c') \quad (35)$$

and at the same time

$$Q_F(c) < Q_F(c') \quad (36)$$

This implies that e.g. publication rankings between countries depend on the used counting method. We will, moreover, extend inequalities (35) and (36) such that:

$$Q_T(c') < Q_T(c) < Q_F(c) < Q_F(c') \quad (37)$$

Inequality (37) indicates that country (or author) c has a smaller relative score in the total counting system than in the fractional counting one, yet this smaller score has a higher importance (as compared to c', using total counting) than the larger one (fractional counting). We will develop examples and counterexamples in a systematic way. Further, denoting the rank of a country by rk, one might expect from (37) that

$$rk_T(c') > rk_T(c) > rk_F(c) > rk_F(c') \quad (38)$$

but even this is not true in general.

Because of the theorem shown in the previous section, the 'anomaly' also holds for total versus proportional counting. The validity of, say publication, rankings between countries, universities, research groups or authors, is not only restricted because of the used reference set (often only articles published in journals covered by ISI), but also because of the used author-counting method. So, such rankings do not contain an absolute truth, but are a mere indication to be corroborated by other methods. It seems that many people are not aware of the implications of (37).

We will now begin the construction of examples satisfying inequality (37). Sometimes we will need relatively complex systems (a high number of articles), but it will suffice to consider only three authors (or, equivalently countries such that $a_i(c) = 1$, for each i = 1,...,N).

Let us consider the general situation of three authors a, b and c in a system with x+y+z single-author articles, namely x with a as single author, y with b as single author and z with c as single author; x'+y'+z' articles with two authors, namely x' with b and c as authors, y' with a and c as authors and z' with a and b as authors; and finally α articles with a, b and c as authors. So, for this system N = x+y+z+x'+y'+z'+α. We will try to determine the seven unknowns x,y,z,x',y',z' and α such that (37) becomes valid. For this system we have:

$$Q_T(a)=\frac{x+z'+y'+\alpha}{x+y+z+2x'+2y'+2z'+3\alpha} \qquad (39)$$

$$Q_T(b)=\frac{y+z'+x'+\alpha}{x+y+z+2x'+2y'+2z'+3\alpha} \qquad (40)$$

and

$$Q_T(c) = \frac{z + x' + y' + \alpha}{x + y + z + 2x' + 2y' + 2z' + 3\alpha} \qquad (41)$$

For the fractional counts we obtain:

$$Q_F(a) = \frac{x + z'/2 + y'/2 + \alpha/3}{x + y + z + x' + y' + z' + \alpha} \qquad (42)$$

$$Q_F(b) = \frac{y + z'/2 + x'/2 + \alpha/3}{x + y + z + x' + y' + z' + \alpha} \qquad (43)$$

and,

$$Q_F(c) = \frac{z + x'/2 + y'/2 + \alpha/3}{x + y + z + x' + y' + z' + \alpha} \qquad (44)$$

Condition (37) can be expressed as:

$$x + y' > y + x' \qquad (45)$$

$$x + \frac{y'}{2} < y + \frac{x'}{2} \qquad (46)$$

$$\frac{(y + z - x)(z' + y')}{2} + \frac{\alpha(2y + 2z - 4x + x' - z'/2 - y'/2)}{3} < xx' \qquad (47)$$

This can be solved for positive integer values, as we will show in the following examples.

Example 2

We take x = z = x' = 0, in order to construct a 'minimal' example. Then, with some elementary algebra we see that y=2, y'=3, α = z' = 13 is one possible solution. This leads to the following table of collaborations (rankings are not indicated, and play no role here).

article	a	b	c
1		x	
2		x	
3	x	x	
4	x	x	
...	x	x	
15	x	x	
16	x		x
17	x		x
18	x		x
19	x	x	x
...	x	x	x
31	x	x	x

This results in the following ranked values for Q_T and Q_F.

Q_T	Q_F
$Q_T(a) = 0.3973$	$Q_F(b) = 0.4140$
$Q_T(b) = 0.3836$	**$Q_F(a) = 0.3978$**
$Q_T(c) = 0.2192$	$Q_F(c) = 0.1882$

Note that $Q_T(b) < Q_T(a) < Q_F(a) < Q_F(b)$, as required. Moreover a ranks first using total counts and is only second using fractional counts.

Example 3. We will now construct an example in which a, b and c never collaborate all together, i.e. $\alpha = 0$. A solution is given as $x = x' = 5$, $y = 7$, $y' = 8$, $z = 1$, $z' = 2$. This gives the following collaboration table.

article	a	b	c
1	x		
2	x		
3	x		
4	x		
5	x		
6		x	
7		x	
8		x	
9		x	
10		x	
11		x	
12		x	
13			x
14	x	x	
15	x	x	
16		x	x
17		x	x
18		x	x
19		x	x
20		x	x
21	x		x
22	x		x
23	x		x
24	x		x
25	x		x
26	x		x
27	x		x

This leads to the following Q-values:

Q_T	Q_F
$Q_T(a) = \mathbf{0.3488}$	$Q_F(b) = 0.3750$
$Q_T(b) = 0.3256$	$Q_F(a) = \mathbf{0.3571}$
$Q_T(c) = 0.3256$	$Q_F(c) = 0.2679$

Note again that $Q_T(b) < Q_T(a) < Q_F(a) < Q_F(b)$. Also, a ranks first according to total counts, and only second according to fractional counts.

Note 1. By the theorem in the previous section these examples yield also examples such that $Q_T(b) < Q_T(a) < Q_P(a) < Q_P(b)$. This shows that these 'paradoxical' examples also exist between the total counting system and the proportional system.

Note 2. If there are, altogether, only two authors, such anomalies are not present. It is indeed trivial that, if there are only two authors that, if $Q_T(a) > Q_T(b)$, then $Q_F(a) < Q_F(b)$ can only occur if $Q_F(a) < Q_T(a)$ since $Q_T(a) + Q_T(b) = Q_F(a) + Q_F(b) = 1$. Of course, the case of two countries represented in more than two co-authorship situations can again lead to such anomalies. This is left as an exercise.

V.Solutions to these anomalies

It is conceivable that scientists doing evaluations do not want to get involved in the problems and anomalies sketched in the previous sections. So, how can one exclude or at least diminish their influence?

We will show that Q_T, Q_F and Q_P are close to each other if the number of collaborators is high. This gives a partial solution to the problems encountered in the previous section, although we are not convinced that one should rely on it for a complete solution. In the last subsection we will show, based on earlier findings (Egghe and Rousseau, 1996b) that a complete solution exists if one replaces the arithmetic average by the geometric mean.

V.1. Partial solutions

The fact that $Q_T(c)$ and $Q_F(c)$ can be close to each other is shown in the unpublished note of Kranakis and Kranakis (1988) (the result can also be found in (Egghe and Rousseau, 1990)). We will repeat the easy result for the sake of completeness.

Theorem (Kranakis and Kranakis, 1988)

For any country c,

$$|Q_T(c) - Q_F(c)| \leq \left(\frac{1}{m} - \frac{1}{M}\right)\frac{\sum_{i=1}^{N} a_i(c)}{N} \qquad (48)$$

where m = min{$a_1,\ldots,a_N$}, M = max{$a_1,\ldots,a_N$}.

Proof. Since, for every c, equation (26) applies

$$Q_F(c) = \frac{1}{N}\sum_{i=1}^{N}\frac{a_i(c)}{a_i} \qquad (26)$$

we have:

$$\frac{1}{MN}\sum_{i=1}^{N} a_i(c) \leq Q_F(c) \leq \frac{1}{mN}\sum_{i=1}^{N} a_i(c), \qquad (49)$$

Since, further for very c,

$$Q_T(c) = \frac{\sum_{i=1}^{N} a_i(c)}{\sum_{i=1}^{N} a_i}, \qquad (25)$$

(49) also holds for $Q_F(c)$ replaced by $Q_T(c)$. This proves inequality (48).

Inequality (48) shows that if m = M, i.e. if the number of co-authors is the same in all papers, then $Q_F(c) = Q_T(c)$. Further, if m and M are approximately equal, or if m (and hence M) is very large, then also $Q_F(c) \approx Q_T(c)$. As to $Q_F(c)$ versus $Q_P(c)$ we have the following result.

Theorem

For any c, such that $a_i(c) = 1$, for every i = 1,…,N, the following inequality holds:

$$\left|Q_F(c) - Q_P(c)\right| \leq \frac{1}{N}\sum_{i=1}^{N}\frac{a_i - 1}{a_i(a_i + 1)} \qquad (50)$$

Proof. If $a_i(c) = 1$, for every i = 1, ...,N, we have

$$Q_F(c) = \frac{1}{N}\sum_{i=1}^{N}\frac{1}{a_i}$$

$$Q_P(c) = \frac{1}{N}\sum_{i=1}^{N}\frac{2}{a_i}\left(1 - \frac{R(i,c)}{a_i + 1}\right),$$

where $R(i,c) \in \{1,\ldots,a_i\}$, since $a_i(c) = 1$, for every i = 1, ...,N. So,

$$Q_F(c) - Q_P(c) = \frac{1}{N}\left(\sum_{i=1}^{N}\frac{1}{a_i} - \sum_{i=1}^{N}\frac{2R(i,c)}{a_i(a_i + 1)}\right)$$

can be bounded as follows:

$$\frac{1}{N}\sum_{i=1}^{N}\frac{1 - a_i}{a_i(a_i + 1)} \leq Q_F(c) - Q_P(c) \leq \frac{1}{N}\sum_{i=1}^{N}\frac{a_i - 1}{a_i(a_i + 1)}$$

Consequently, since $a_i \geq 1$:

$$\left|Q_F(c) - Q_P(c)\right| \leq \frac{1}{N}\sum_{i=1}^{N}\frac{a_i - 1}{a_i(a_i + 1)}$$

This proves this theorem.

Corollary 1

Under the assumptions of the above theorem we also have:

$$\left|Q_F(c) - Q_P(c)\right| \leq average \quad of \ \{\frac{1}{a_1 + 1},\ldots,\frac{1}{a_N + 1}\} \qquad (51)$$

Corollary 2

Let m = min$\{a_1,\ldots,a_N\}$ as before. Then

$$|Q_F(c) - Q_P(c)| \leq \frac{1}{m+1} \qquad (52)$$

Hence, $Q_F \approx Q_P$ if m is high. Together with the earlier found result, we hence have:

$$\text{if m is high: } Q_T \approx Q_F \approx Q_P \qquad (53)$$

V. 2. A complete solution to the encountered anomalies

In this section we will show that, using geometric instead of arithmetic averages, eliminates the occurrence of these anomalies. Recall that the use of geometric means was already suggested in (Egghe and Rousseau, 1996b). Let us first define these geometric versions (denoted with a g as superscript).

By (27)

$$Q_T(c) = \frac{\textit{arithmetic average of } \{a_1(c),\ldots,a_N(c)\}}{\textit{arithmetic average of } \{a_1,\ldots,a_N\}}$$

hence, we define:

$$Q_T^g(c) = \frac{(a_1(c)a_2(c)..a_N(c))^{1/N}}{(a_1a_2..a_N)^{1/N}} \qquad (54)$$

By (28), Q_F is:

$$Q_F(c) = \textit{arithmetic average of } \{\frac{a_1(c)}{a_1},\ldots,\frac{a_N(c)}{a_N}\}$$

hence, we define:

$$Q_F^g(c)=\left(\frac{a_1(c)}{a_1}\cdot\frac{a_2(c)}{a_2}\cdots\frac{a_N(c)}{a_N}\right)^{1/N} \qquad (55)$$

It is clear that

$$Q_T^g(c) = Q_F^g(c) \qquad (56)$$

for any system and any c. Furthermore, by (56) all rankings based on the total counting method are the same as those based on the fractional counting method. Consequently, all ambiguities are gone!

We leave it as an open problem to study the behavior of the proportional counting system when arithmetic averages are replaced by geometric ones. In other words: what is the relation between $Q_P^g(c)$ *and* $Q_F^g(c)$?

Note. In the theoretical sections of this article we have restricted ourselves to publication scores. In case we want to deal with citation scores (e.g. impacts) the situation becomes even more complicated as publication as well as citations scores can be counted in different ways. This leads even easier to all kind of anomalies (Van Hooydonk, 1997), as shown in Figs. 1 and 2.

We conclude this section by referring to the data sample presented in Section II.Fig. 4 represents the data for the 14 co-operating authors in this group, having as one common element the co-authorship with the same senior scientist (author 1 in Fig.4). The number of authors on these 36 papers varies from three to nine. A log-scale has been used to obtain a more explicit picture of the observed trends in, and reversals caused by, the different counting procedures. We see that the fractional counting method roughly follows the total count results. The proportional method leads to reversals in the number of accredited publications, which, expressed as percentages can be rather large. It appears that in this sample the output of the senior author, who is an author on all 36 publications, but is never first author, is almost invariant to the counting method. This is in line with the trends observed in Figs. 1 and 2. Since Fig.4 gives an idea of the consequences of applying different accrediting methods in the real world, it is an interesting illustration of the theoretical matters dealt with in the previous sections.

Insert Fig.4 about here

A similar illustration can be found in (Rousseau and Rousseau, 1998, p. 80) for efficiency scores of countries, based on the DEA (data envelopment analysis) method.

VI. Conclusion

We have presented a review of different accrediting or scoring methods for countries, institutes, research groups or authors, based on their publication or citation output. It is shown that different methods give not only different relative numerical results, but that unexpected (paradoxical) results may occur. Indeed, we have presented examples in which country (or author) c has a smaller relative score in the total counting system than in the fractional counting one, yet this smaller score has a higher importance than the larger one (fractional counting). Similar anomalies were constructed for total versus proportional counts and for total versus straight counts. Consequently, a ranking between countries, universities, research groups or authors, based on one particular accrediting method does not contain an absolute truth about the relative importance of the players in the system. Different counting methods should be used and compared. Differences are illustrated with a real-life example. Finally, it is shown that some of these anomalies can be avoided by using geometric instead of arithmetic averages.

References

Anderson, J. , Collins, P.M.D., Irvine,J., Isard, P.A., Martin, B.R., Narin, F, and Stevens K. (1988). On-line approaches to measuring national scientific output: a cautionary tale. *Science and Public Policy*, 15(3),153-161.

Burrell, Q. and Rousseau,R. (1995). Fractional counts for authorship attribution: a numerical study. *Journal of the American Society for Information Science*, 46, 97-102.

Cole,J.R. and Cole,S. (1973). *Social stratification in science.* Chicago: The University of Chicago Press.

De Bruin, R., Kint, A., Luwel, M. and Moed, H.F. (1993). A study of research evaluation and planning: the University of Ghent. *Research Evaluation*, 3, 25-41.

Egghe, L. (1996). Source-item production laws for the case that items have multiple sources with fractional counting of credits. *Journal of the American Society for Information Science*, 47, 730-748.

Egghe,L. and Rousseau, R. (1990). Introduction to informetrics. Quantitative methods in library, documentation and information science. Amsterdam: Elsevier.

Egghe,L. and Rousseau, R. (1996a). Average and global impact of a set of journals. *Scientometrics,* 36, 97-107.

Egghe, L. and Rousseau, R. (1996b). Averaging and globalizing quotients of informetric and scientometric data. *Journal of Information Science*, 22, 165-170.

Harsanyi, M.A. (1993). Multiple authors, multiple problems – bibliometrics and the study of scholarly collaboration: a literature review. *Library and Information Science Research* 15, 325-354.

Kranakis, E. and Kranakis E. (1988). Comparing two weighting methods in citation analysis. Unpublished paper. Amsterdam.

Lindsey, D. (1980). Production and citation measures in the sociology of science: the problem of multiple authorship. *Social Studies of Science*, 10, 145-162.

Nederhof, A.J. and Moed, H.F. (1993). Modeling multinational publication: development of an on-line fractionation approach to measure national scientific output. *Scientometrics*, 27, 39-52.

Ojasoo, Tiiu (1996). National scientific endeavour from 1981 to 1992. *Science Tribune*. August 1996. www.tribunes.com/tribune/art96/ojas1.htm

Oppenheimer,C. (1998). Fractional counting of multiauthored publications. *Journal of the American Society for Information Science,* 49(5), 482.

Price, D. de Solla (1981). Letter to the editor. *Science*, 212, 987.

Reed, K.L. (1995). Citation analysis of faculty publication: beyond Science Citation Index and Social Science Citation Index. *Bulletin of the Medical Librar. Association*, 83(4), 503-508.

Rousseau, R. (1998). A scientometric study of the scientific publications of LUC. Period 1984-1996. Report: LUC.

Rousseau, S. and Rousseau, R. (1998).The scientific wealth of European nations: taking effectiveness into account. *Scientometrics,* 42(1), 75-87.

Toutkoushian, R.K. (1994). Using citations to measure sex discrimination in faculty salaries. *The Review of Higher Education*, 18(1), 61-82.

Van Hooydonk, G. (1997). Fractional counting of multi-authored publications: consequences for the impact of authors. *Journal of the American Society for Information Science,* 48, 944-945.

Van Hooydonk, G. (1998). Standardizing relative impacts: estimating the quality of research from citation counts. *Journal of the American Society for Information Science,* 49(10), 932-941.

White,H. and McCain,K. (1998). Visualizing a discipline: an author co-citation analysis of information science, 1972-1995. *Journal of the American Society for Information Science,* 49(4), 327-355.

Zuckerman, H. (1968). Patterns of name-ordering among authors of scientific papers: a study of social symbolism and its ambiguity. *American Journal of Sociology*, 74, 276-291.

Appendix

Anomalies involving straight (or first-author) and total (or normal) counts. As in the main text, the symbol c_i denotes a country.

Example 4

article	first author	second	third	fourth
1	c_1	c_3	c_3	c_3
2	c_2	c_1	c_1	c_3
3	c_2	c_3	c_3	c_3

This yields the following scores

	Q_S	Q_T
c_1	1/3	3/12
c_2	2/3	2/12
c_3	0	7/12

Note that country 3 does not score according to the straight counting method, but is first according to total counts. Moreover:

$$Q_T(c_2) < Q_T(c_1) < Q_s(c_1) < Q_s(c_2) \qquad (57)$$

Equation (57) is an analogue of equation (37). Yet, country c_1 always ranks second. We will next give an example where also ranks change.

Example 5

article	first author	second	third	fourth	fifth
1	c_1	c_3	c_3	c_3	c_3
2	c_2	c_1	c_1	c_3	c_3
3	c_2	c_3	c_3	c_1	c_3
4	c_4	c_3	c_3	c_3	c_3
5	c_4	c_3	c_3	c_3	-
6	c_4	c_3	c_3	c_3	-

This yields the following scores

	Q_S	Q_T
c_1	1/6	4/28
c_2	2/6	2/28
c_3	0	19/28
c_4	3/6	3/28

Note that country 3 does not score according to the straight counting method, but is (again) first according to total counts. Moreover, also here

$$Q_T(c_2) < Q_T(c_1) < Q_s(c_1) < Q_s(c_2) \qquad (57)$$

Now, denoting the rank of a country by rk, one might expect from (57) that

$$rk_T(c_2) > rk_T(c_1) > rk_s(c_1) > rk_s(c_2)$$

but even this is not true: indeed $rk_T(c_2) = 4 > rk_T(c_1) = 2$, and $rk_s(c_1) = 3 > rk_s(c_2) = 2$, yet $rk_T(c_1) = 2 \not\geq rk_s(c_1) = 3$.

It is not difficult to write down a result similar to the theorem relating fractional and proportional counting. Indeed we have the following (trivial) result.

Proposition on straight and normal counting

For any system of collaboration between countries (institutions, authors,...), denoted as system I, we can construct another one, denoted as system II, such that the straight counting result in system II equals the total counting result in system I, as well as in system II. Moreover, in system II the four counting methods studied in this article coincide. In symbols:

$$Q_T^{I}(c) = Q_T^{II}(c) = Q_S^{II}(c) = Q_F^{II}(c) = Q_P^{II}(c).$$

Proof. It suffices to define system II as the system consisting of only single-author articles, namely each author of system I.

An example

System I consists of three articles with the following (ranked) authors

a a' a"	a a" a'	a' a"

System II consists of eight articles with the following authors

a	a'	a"
a	a"	a'
a'	a"	

Fig. 1 Impact C/P according to 3 counting methods for 14 authors, cooperating with the same senior author

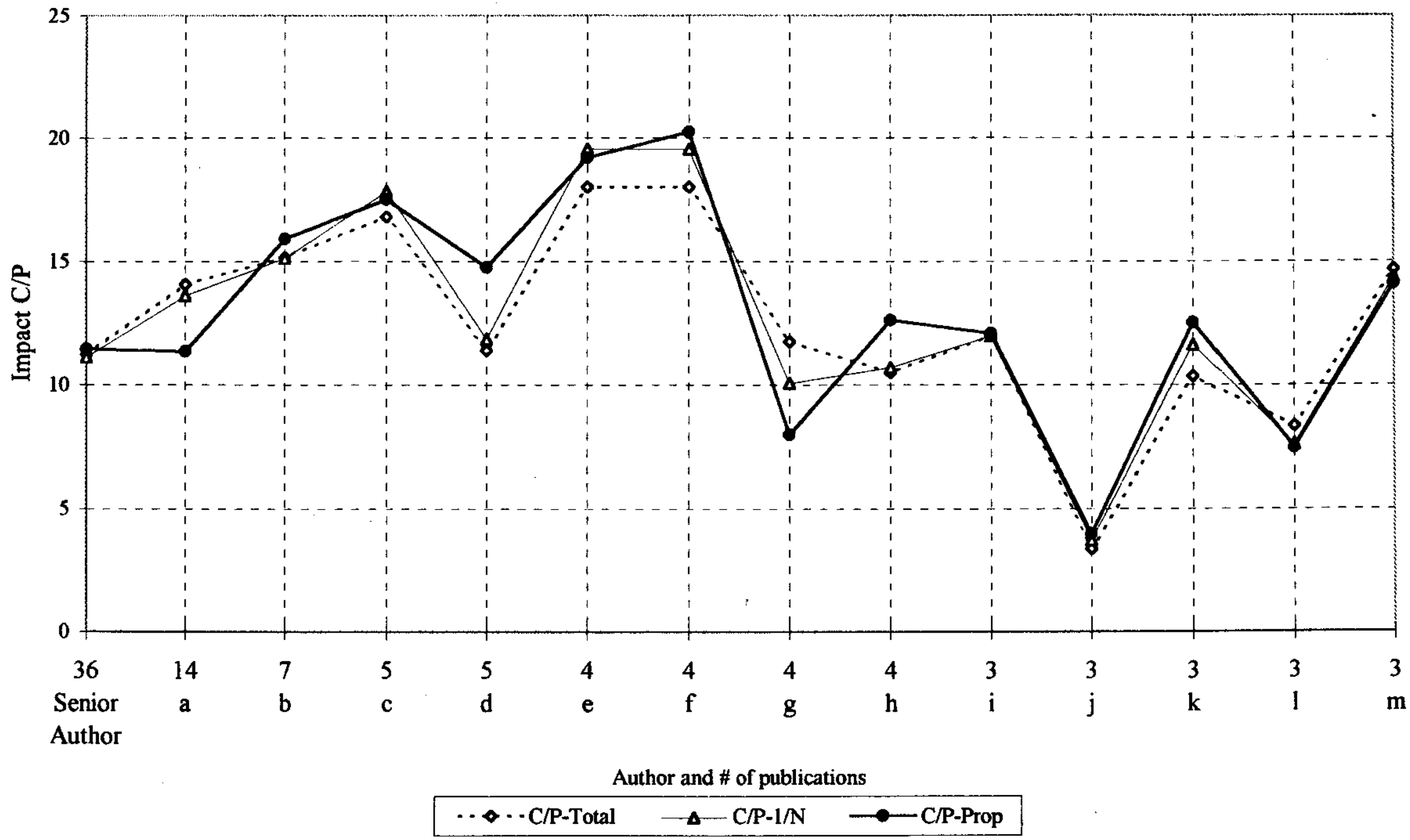

Fig. 2 Impact ratio (scaled by total count impact)
Ratio 1 for fractional counting; Ratio 2 for proportional counting

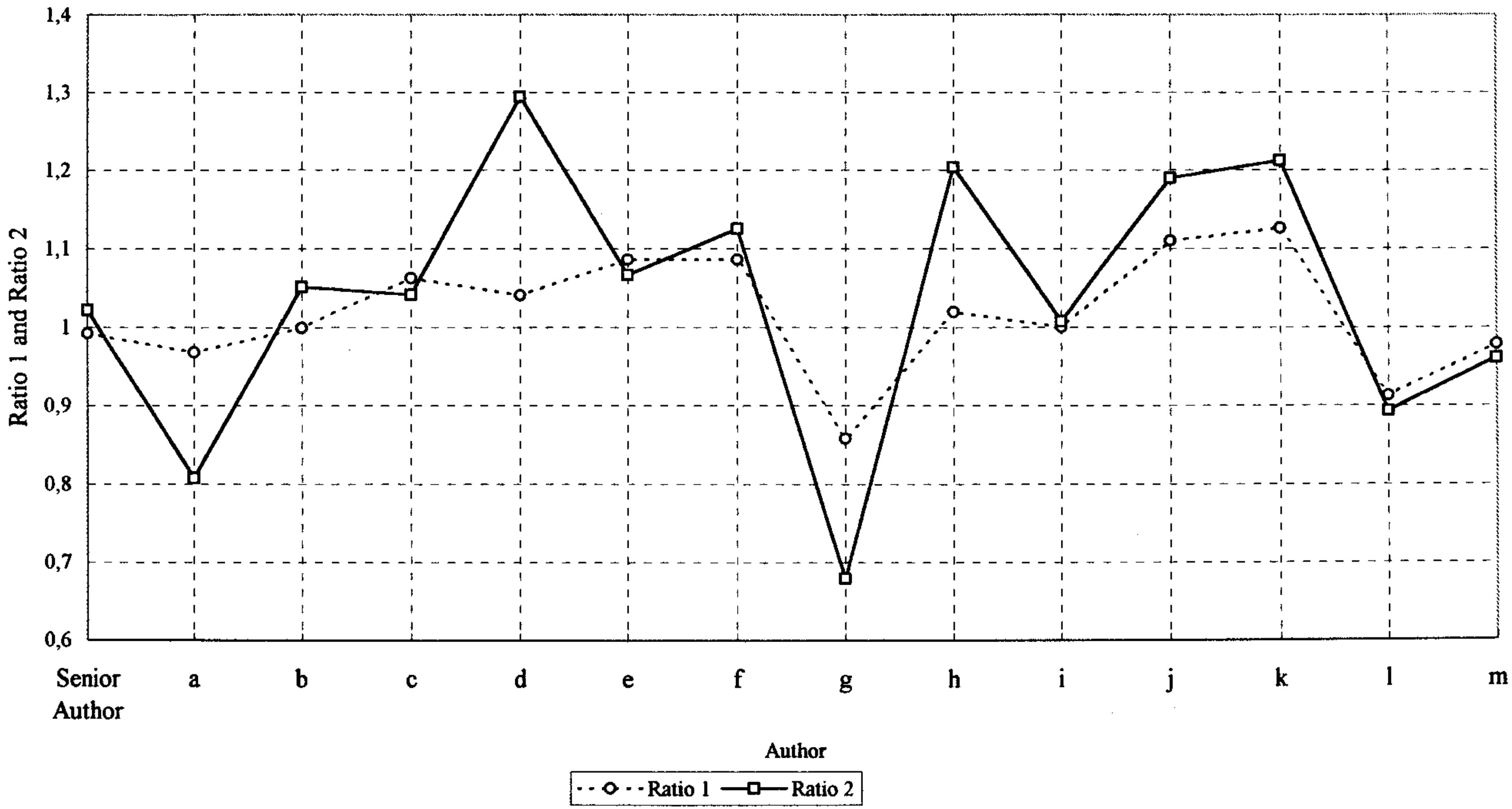

Fig. 3 Article correlation between systems I and II

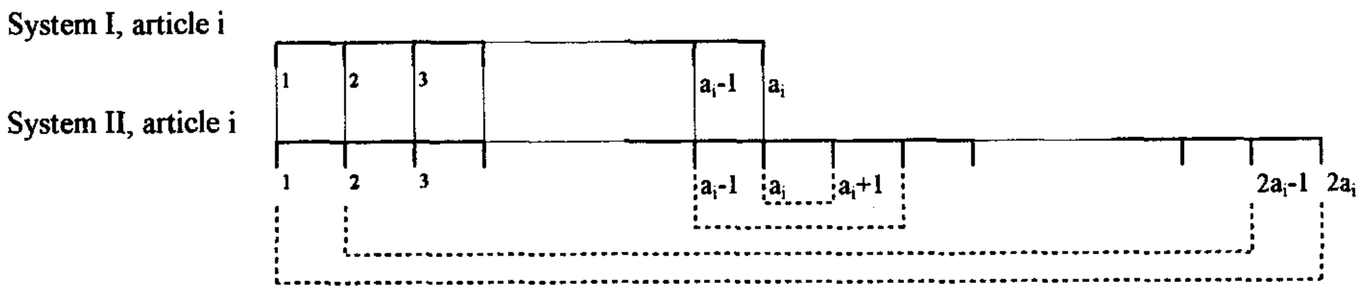

Fig. 4 Publication counts according to three methods (Tot, Fract, Prop) for 14 cooperating authors (senior author first, log scale, same authors as in Fig. 1 and 2)