Field-Normalized Citation Impact Indicators and the Choice of an Appropriate Counting Method

Ludo Waltman and Nees Jan van Eck

{waltmanlr, ecknjpvan}@cwts.leidenuniv.nl
Centre for Science and Technology Studies, Leiden University, Leiden (The Netherlands)

Abstract
Bibliometric studies often rely on field-normalized citation impact indicators in order to make comparisons between scientific fields. We discuss the connection between field normalization and the choice of a counting method for handling publications with multiple co-authors. Our focus is on the choice between full counting and fractional counting. Based on an extensive theoretical and empirical analysis, we argue that properly field-normalized results cannot be obtained when full counting is used. Fractional counting does provide results that are properly field normalized. We therefore recommend the use of fractional counting in bibliometric studies that require field normalization, especially in studies at the level of countries and research organizations.

Conference Topic
Citation and co-citation analysis; Indicators

Introduction
In discussions on bibliometric indicators, two topics that receive a considerable amount of attention are field normalization and counting methods. Field normalization is about the problem of correcting for differences in citation practices between scientific fields. The challenge is to develop citation-based indicators that allow for valid between-field comparisons. Counting methods are about the way in which co-authored publications are handled. For instance, if a publication is co-authored by two countries, should the publication be counted as a full publication for each country or should it be counted as half a publication for each country?

The topics of field normalization and counting methods are usually discussed separately from each other. However, we argue that there is a close connection between the two topics. Our argument is that proper field normalization is possible only if a suitable counting method is used. In particular, we claim that properly field-normalized results cannot be obtained when one uses the popular full counting method, in which co-authored publications are fully assigned to each co-author. The fractional counting method, which assigns co-authored publications fractionally to each co-author, does provide properly field-normalized results. The problem of full counting basically is that co-authored publications are counted multiple times, once for each co-author, which creates a bias in favor of fields in which there is a lot of co-authorship and in which co-authorship correlates with additional citations. This is the essence of the argument that we present in this paper. Our argument builds on an earlier paper (Waltman et al., 2012), but in the present paper we elaborate the argument in more detail and we also present an extensive empirical analysis.

This paper is a shortened version of a more extensive working paper (Waltman & Van Eck, 2015). The working paper includes additional empirical analyses comparing different counting methods at the level of institutions and countries. Furthermore, the working paper considers different variants of fractional counting and also studies first author and corresponding author counting methods.

Counting methods

Our focus is on the comparison between full counting and fractional counting. In the case of full counting, a publication is fully assigned to each co-author. For instance, a publication co-authored by four countries counts as a full publication for each of the four countries. In the fractional counting case, a publication is fractionally assigned to each co-author. The weight with which a publication is assigned to a co-author indicates the share of the publication allocated to that co-author. The sum of the weights of all co-authors of a publication equals one. An example of fractional counting is the situation in which a publication co-authored by four countries is assigned to each country with a weight of 1 / 4 = 0.25.

There is a quite extensive literature on counting methods. Because of space limitations, we mention only a few selected studies. A systematic terminology for counting methods is proposed by Gauffriau, Larsen, Maye, Roulin-Perriard, and Von Ins (2007). They refer to full counting as whole counting and to fractional counting as normalized counting. Gauffriau, Larsen, Maye, Roulin-Perriard, and Von Ins (2008) present a comparison of counting methods at the country level. They also provide an overview of earlier literature on counting methods. Another country-level comparison is reported by Aksnes, Schneider, and Gunnarsson (2012). At the institution level, Waltman et al. (2012) present a comparison between full and fractional counting. Interesting work on counting methods can also be found in various papers by Ruiz-Castillo and colleagues, who propose the idea of a so-called multiplicative counting method (e.g. Albarrán, Crespo, Ortuño, & Ruiz-Castillo, 2010).

Relation between counting methods and field normalization

Our aim in this section is to demonstrate the close connection between counting methods and field normalization. In particular, we aim to make clear that full counting is fundamentally inconsistent with the idea of field normalization. We argue that full counting yields results that suffer from a bias in favor of fields in which there is a lot of co-authorship and in which co-authorship correlates with additional citations. This bias is caused by the fact that co-authored publications are counted multiple times in the case of full counting, once for each co-author.

We present our argument by providing two simple examples. Both examples take countries as the unit of analysis and focus on the mean normalized citation score (MNCS) indicator (Waltman, Van Eck, Van Leeuwen, Visser, & Van Raan, 2011). However, the underlying ideas of the two examples are more general, and similar examples can be given with authors or organizations as the unit of analysis and with other field-normalized indicators.

Table 1. Example involving a single field.

	Authors	No. of cit.	Norm. cit. score
Publication 1	Country A	3	0.6
Publication 2	Country A	6	1.2
Publication 3	Country B	1	0.2
Publication 4	Country A; Country B	10	2.0

Example involving a single field

We consider a world in which there are just four publications. These publications have been produced by two countries, labeled as country A and country B. Table 1 shows for each publication the countries by which the publication is authored and the number of citations the publication has received. The table also shows the normalized citation score of each publication. For simplicity, it is assumed that all four publications are in the same field. The normalized citation score of a publication is therefore obtained simply by dividing the number of citations of the publication by the average number of citations of all four publications. The

average number of citations of the four publications equals (3 + 6 + 1 + 10) / 4 = 5, and therefore the normalized citation score of for instance publication 1 equals 3 / 5 = 0.6. Of course, the average of the normalized citation scores of the four publications equals one.
We now calculate both for country A and for country B the MNCS. Using full counting, we obtain

$$\mathrm{MNCS}_\mathrm{A} = \frac{0.6 + 1.2 + 2.0}{3} = 1.27 \text{ and } \mathrm{MNCS}_\mathrm{B} = \frac{0.2 + 2.0}{2} = 1.10.$$

On the other hand, using fractional counting, we get

$$\mathrm{MNCS}_\mathrm{A} = \frac{1.0 \times 0.6 + 1.0 \times 1.2 + 0.5 \times 2.0}{1.0 + 1.0 + 0.5} = 1.12 \text{ and } \mathrm{MNCS}_\mathrm{B} = \frac{1.0 \times 0.2 + 0.5 \times 2.0}{1.0 + 0.5} = 0.80,$$

where publication 4 has been assigned with a weight of 0.5 to country A and with a weight of 0.5 to country B.
The important thing to observe in this example is that in the case of full counting country A and country B both have an MNCS above one. One of the main ideas of field-normalized indicators such as the MNCS indicator is that the value of one can be interpreted as the world average. Under this interpretation, country A and country B both perform above the world average. Since there are no other countries in our example, the conclusion would be that all countries in the world perform above the world average. There are no countries with a below-average performance. In our opinion, the conclusion that everyone is above average does not make much sense. Moreover, this conclusion is fundamentally different from the conclusion that is reached in the case of fractional counting. Using fractional counting, country A has a performance above the world average while the performance of country B is below the world average.
Looking a bit more in detail at our example, we observe that in the fractional counting case we have

$$\frac{2.5 \times \mathrm{MNCS}_\mathrm{A} + 1.5 \times \mathrm{MNCS}_\mathrm{B}}{2.5 + 1.5} = \frac{2.5 \times 1.12 + 1.5 \times 0.80}{2.5 + 1.5} = 1.$$

Hence, the weighted average of the MNCS of country A and the MNCS of country B, with weights given by each country's fractional number of publications, equals exactly one. This is a general property of fractional counting. The weighted average of the MNCSs of all countries in the world will always be equal to exactly one.
In the full counting case, the weighted average of the MNCS of country A and the MNCS of country B equals

$$\frac{3 \times \mathrm{MNCS}_\mathrm{A} + 2 \times \mathrm{MNCS}_\mathrm{B}}{3 + 2} = \frac{3 \times 1.27 + 2 \times 1.10}{3 + 2} = 1.20,$$

where the weight of each country is given by the number of publications of the country obtained using full counting. So in the full counting case the world average at the country level does not equal one but instead equals 1.20. Taking 1.20 as the world average, we conclude that country A, with an MNCS of 1.27, has an above-average performance while

country B, with an MNCS of 1.10, performs below average. This is in agreement with the conclusion reached using fractional counting.

So in our example there is a difference of 1.20 – 1 = 0.20 between the world average obtained using full counting and the world average obtained using fractional counting. We refer to this difference as the full counting bonus. In principle, the full counting bonus can be either positive or negative, but we will see that in practice the bonus is usually positive. The full counting bonus is caused by the fact that publications co-authored by multiple countries are counted multiple times in the case of full counting, and therefore the citation impact of multi-country publications relative to single-country publications determines whether the full counting bonus is positive or negative. The bonus will be positive if publications co-authored by multiple countries receive more citations than publications authored by a single country. Conversely, a negative bonus will be obtained if multi-country publications are cited less frequently than single-country publications. As can be seen in Table 1, in our example the only publication co-authored by multiple countries is publication 4, and this is also the most highly cited publication. In the full counting case, publication 4 is fully assigned both to country A and to country B. Hence, the most highly cited publication in our example is counted two times, once for country A and once for country B. This double counting of publication 4 explains why both countries have an MNCS above one and why the full counting bonus is positive.

Example involving multiple fields

In the example discussed above, all publications are in the same field. We now consider an example that involves more than one field. This example is presented in Table 2. There are six publications, three in field X and three in field Y, and there are four countries. Countries A and B are active only in field X, while countries C and D are active only in field Y. The three publications in field X have all received the same number of citations, and therefore these publications all have a normalized citation score of one. This is not the case in field Y, in which publication 6, co-authored by countries C and D, has received more citations than publications 4 and 5, which are single-country publications. Of course, the average normalized citation score of the publications in field Y equals one, just like in field X.

Table 2. Example involving multiple fields.

	Field	Authors	No. of cit.	Norm. cit. score
Publication 1	Field X	Country A	10	1.0
Publication 2	Field X	Country B	10	1.0
Publication 3	Field X	Country A; Country B	10	1.0
Publication 4	Field Y	Country C	4	0.8
Publication 5	Field Y	Country D	4	0.8
Publication 6	Field Y	Country C; Country D	7	1.4

Using fractional counting, the four countries all have an MNCS of exactly one. For countries A and B this is immediately clear. In the case of countries C and D, the MNCS is calculated as $(1.0 \times 0.8 + 0.5 \times 1.4) / (1.0 + 0.5) = 1$. So fractional counting tells us that all four countries perform at the world average. This is indeed the outcome that we would expect to obtain. The publications of countries A and B have all been cited equally frequently as the average of their field, so countries A and B obviously perform at the world average. In the case of countries C and D, we observe that these countries have exactly the same performance and that they are the only countries active in field Y. Based on these two observations, it is natural to conclude that the performance of countries C and D is at the world average.

We now consider the full counting case. Using full counting, countries A and B have an MNCS of one, while countries C and D have an MNCS of $(0.8 + 1.4) / 2 = 1.10$. The full

counting results seem to suggest that countries C and D have a better performance than countries A and B. However, a more careful analysis shows that this is not a correct interpretation of the results. To see this, we calculate both for field X and for field Y the average of the MNCSs of the countries active in the field. The average MNCS of the countries active in field X equals one, while the average MNCS of the countries active in field Y equals 1.10. Hence, both countries A and B active in field X and countries C and D active in field Y perform at the world average of their field. Like in the fractional counting case, we conclude that all four countries have an average performance. Countries C and D have a higher MNCS than countries A and B only because they are active in a field with a higher full counting bonus. Field Y has a full counting bonus of $1.10 - 1 = 0.10$, while the full counting bonus in field X equals zero.

Conclusions based on the examples

Based on the above examples, two important conclusions can be drawn. The first conclusion is that there is a need to carefully distinguish between two field normalization concepts. We refer to these concepts as weak field normalization and strong field normalization. Weak field normalization requires the average of the normalized citation scores of all publications in a field to be equal to one. Strong field normalization is more demanding. It requires the weighted average of the MNCSs of all countries active in a field to be equal to one, where the weight of a country is given by its number of publications in the field.

As shown in the above examples, full counting yields results that are in agreement with the idea of weak field normalization, but these results may violate the idea of strong field normalization. For instance, in the first example discussed above, the average normalized citation score of the four publications equals one (weak field normalization), but the average MNCS of the two countries does not equal one (no strong field normalization). Fractional counting results, on the other hand, satisfy not only the idea of weak field normalization but also the idea of strong field normalization. Using fractional counting, the weighted average of the MNCSs of all countries active in a field will always be equal to one.

When citation-based indicators are calculated using full counting, there is a risk of misinterpretation. People may confuse the concepts of weak and strong field normalization, and they may fail to understand that the idea of strong field normalization does not apply in the case of full counting. In the second example presented above, they may for instance draw the incorrect conclusion that countries C and D perform above the world average. In the fractional counting case, people will not draw such an incorrect conclusion, because fractional counting results are in agreement with the idea of strong field normalization.

We now turn to the second conclusion that follows from our examples. The fact that full counting yields results that are incompatible with the idea of strong field normalization may in itself be regarded as just a minor issue. Instead of having a world average of one, the average of all countries in the world may for instance be equal to 1.10 or 1.20. Although a world average of one might be somewhat more convenient, the exact value of the world average may in the end seem to be of limited importance.

However, our second conclusion is that deviations of the world average from one actually do have serious consequences, at least when making comparisons between fields. This is what is shown in the second example given above. Using full counting, the average MNCS of the countries active in field X equals one, while the average MNCS of the countries active in field Y equals 1.10. So in field X the world average equals one, while in field Y we have a world average of 1.10. Direct comparisons of the MNCSs of the countries active in field X and the countries active in field Y therefore do not yield valid conclusions. Based on their MNCSs, the countries active in field Y seem to perform better than the countries active in field X, but

taking into account the fact that field Y has a higher world average than field X, it actually should be concluded that all countries perform at the same level.

Essentially, the second conclusion that we draw based on our examples is that full counting is fundamentally inconsistent with the idea of field normalization. Citation-based indicators calculated using full counting yield results that do not allow for valid comparisons between fields, and this is the case even when field-normalized indicators, such as the MNCS indicator, are used. When full counting is used in the calculation of field-normalized indicators, countries that focus their activity on fields with a high full counting bonus have an advantage over countries that are active mainly in fields with a low full counting bonus. Fractional counting does not suffer from this problem. Fractional counting results are compatible with the idea of strong field normalization, and these results therefore do allow for proper between-field comparisons.

Empirical analysis of the full counting bonus

In the previous section, we have introduced the idea of the full counting bonus and we have illustrated this idea using theoretical examples. In this section, we present a large-scale empirical analysis of the full counting bonus. This analysis for instance makes clear which fields benefit most from the full counting bonus, and the analysis shows the differences between fields caused by the bonus.

Calculation of the full counting bonus

We first explain in more detail the way in which we calculate the full counting bonus. For simplicity, we assume that our interest is in the full counting bonus at the level of countries. However, the full counting bonus can be calculated in a similar way at the level of for instance authors or organizations.

Suppose we have a set of n publications. This could be for instance the set of all publications in a specific field and in a specific year. For each publication i, we have a citation score c_i. The citation score of a publication can be defined in different ways. It may be simply the number of times a publication has been cited, but it may also be something more advanced, for instance a field-normalized citation score. We also know for each publication the countries by which the publication has been co-authored. We use m_i to denote the number of countries that have co-authored publication i.

In order to obtain the full counting bonus, we first calculate for each country the average citation score of its publications. We perform this calculation both using full counting and using fractional counting. Next, we calculate a weighted average of the average citation scores of all countries. In the case of full counting, we use the number of publications of a country obtained using full counting as the weight of the country. In the case of fractional counting, we use a country's number of publications obtained using fractional counting as the country's weight. Finally, we calculate the full counting bonus as the difference between the weighted average in the full counting case and the weighted average in the fractional counting case.

The above approach to calculating the full counting bonus is somewhat complicated. However, a mathematically equivalent but much simpler approach is available. In this approach, the full counting bonus is calculated as

$$\mathrm{FCB} = \frac{\sum_{i=1}^{n} m_i c_i}{\sum_{i=1}^{n} m_i} - \frac{\sum_{i=1}^{n} c_i}{n},$$

where the first term equals the above-mentioned weighted average in the full counting case while the second term equals the weighted average in the fractional counting case. In the first term, the citation score c_i of publication i co-authored by m_i countries is counted m_i times. This is because in the full counting case publication i is fully assigned to each of the m_i countries. In the second term, the citation score c_i of publication i is counted only once, regardless of the number of countries m_i by which publication i has been co-authored. This is because in the fractional counting case the total weight with which publication i is assigned to the m_i countries equals one.

In our empirical analysis, we consider two definitions of the citation score of a publication. Both definitions include a normalization for field. In the first definition, the citation score of a publication is obtained by dividing the number of citations of the publication by the average number of citations of all publications in the same field and in the same year. Averaging the citation scores of multiple publications then gives us the MNCS indicator. This indicator was also used in the theoretical examples presented in the previous section. In the second definition of the citation score of a publication, we determine whether a publication belongs to the top 10% most frequently cited publications of its field and publication year. A publication belonging to the top 10% has a citation score of one, while a publication belonging to the bottom 90% has a citation score of zero. When this second definition is used, averaging the citation scores of multiple publications yields the $PP_{top\ 10\%}$ indicator, where $PP_{top\ 10\%}$ stands for the proportion of top 10% publications (Waltman et al., 2012; Waltman & Schreiber, 2013). When the full counting bonus is calculated for the set of all publications in a specific field and in a specific year, the second term in the above equation for the full counting bonus will be equal to one in the case of our first definition of the citation score of a publication. This term will be equal to 0.1 (or 10%) in the case of our second definition.

Empirical results

We perform our analysis using the Web of Science (WoS) database. The analysis is based on publications in the period 2009–2010. Only publications of the WoS document types 'article' and 'review' are taken into account. A four-year citation window is used, including the year in which a publication appeared. For the purpose of the calculation of the field-normalized citation scores of publications, fields are defined by the WoS journal subject categories.

We consider three units of analysis: Authors, organizations, and countries. To determine the number of organizations and the number of countries by which a publication has been co-authored, we take into account both the regular addresses of the publication and the reprint address. The number of organizations and the number of countries of a publication is obtained by counting the number of distinct organization names and the number of distinct country names mentioned in the addresses of the publication.

The full counting bonus depends on two factors. On the one hand, it depends on the variation among publications in the number of authors, organizations, or countries. For instance, if all publications have the same number of authors, there can be no full counting bonus at the level of authors. On the other hand, the full counting bonus also depends on the relation between the number of authors, organizations, or countries of a publication and the citation score of the publication. There can for instance be no author-level full counting bonus if publications with different numbers of authors on average all have the same citation score.

Figure 1 presents the distribution of publications based on their number of authors, organizations, and countries. Not surprisingly, the figure shows that the variation among publications in the number of authors is largest while the variation among publications in the number of countries is smallest. Figure 2 presents the relation between the number of authors, organizations, and countries of a publication and the average citation score given by the MNCS indicator. In general, an increasing relation can be observed between the number of

authors, organizations, and countries of a publication and the average citation score. The relation is strongest for countries and weakest for authors. In fact, when the number of authors is between two and five, there is hardly any dependence of the average citation score of a publication on the number of authors. Publications with three or four authors on average even have a slightly lower citation score than publications with two authors. Results for the $PP_{top\ 10\%}$ are not shown, but are similar to the results for the MNCS indicator.

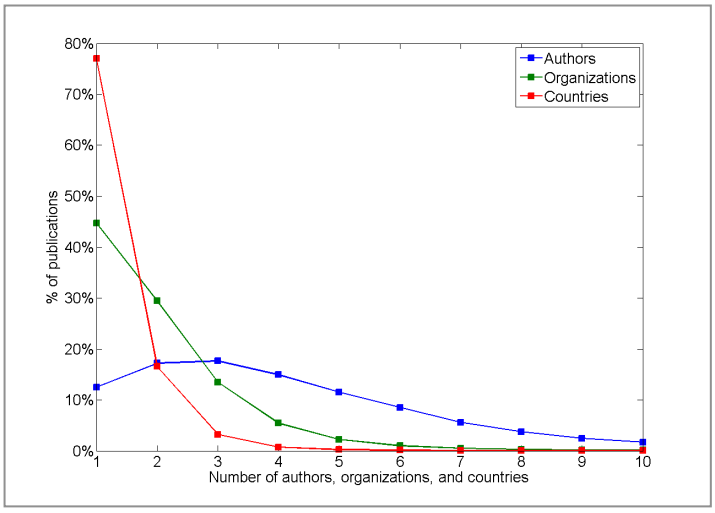

Figure 1. Distribution of publications based on their number of authors, organizations, and countries.

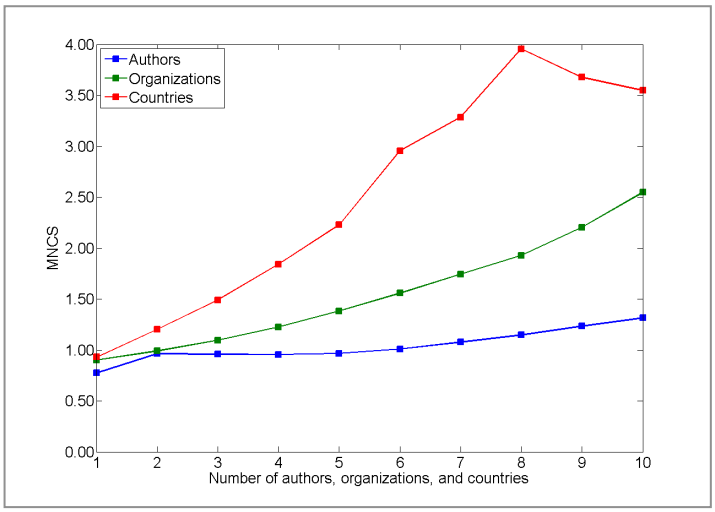

Figure 2. Relation between the number of authors, organizations, and countries of a publication and the MNCS indicator.

Figures 1 and 2 make clear that publications often have multiple co-authors and that the citation impact of a publication tends to increase with the number of co-authors. Co-authored publications are counted multiple times in the case of full counting, and our expectation based on Figures 1 and 2 therefore is to observe full counting bonuses that are positive and of significant size. This is indeed what is reported in Tables 3 and 4. The tables show the full counting bonus at the level of authors, organizations, and countries for five broad fields of science and also for all fields of science taken together. Table 3 relates to the MNCS indicator, while Table 4 relates to the $PP_{top\ 10\%}$ indicator. In order to facilitate comparison between the results obtained for the two indicators, the full counting bonus is presented as a percentage of the average value of the indicator. For instance, in the case of the MNCS

indicator, we obtain a full counting bonus of 0.248 at the level of authors for all fields of science. The average value of the MNCS indicator equals one, and therefore the full counting bonus is reported as 0.248 / 1 = 24.8% in Table 3. Likewise, the $PP_{top\ 10\%}$ indicator has an average value of 0.1 (or 10%), and therefore a full counting bonus of 0.0304 (or 3.04%) is reported as 0.0304 / 0.1 = 30.4% in Table 4.

Table 3. Full counting bonus for the MNCS indicator at the level of authors, organizations, and countries, including a breakdown into five broad fields of science.

	Authors	Organizations	Countries
All fields	24.8%	21.1%	12.6%
Biomedical and health sciences	20.9%	26.8%	16.7%
Life and earth sciences	14.7%	16.2%	12.7%
Mathematics and computer science	8.2%	8.0%	6.9%
Natural sciences and engineering	35.2%	19.3%	10.8%
Social sciences and humanities	14.7%	11.2%	5.6%

Table 4. Full counting bonus for the PPtop 10% indicator at the level of authors, organizations, and countries, including a breakdown into five broad fields of science.

	Authors	Organizations	Countries
All fields	30.4%	26.5%	17.1%
Biomedical and health sciences	24.9%	34.5%	22.6%
Life and earth sciences	22.8%	24.3%	19.7%
Mathematics and computer science	11.3%	11.3%	9.7%
Natural sciences and engineering	43.3%	20.6%	13.0%
Social sciences and humanities	21.3%	17.2%	8.3%

Based on the results for the MNCS indicator presented in Table 3, a number of conclusions can be drawn. At all three analysis levels (i.e., authors, organizations, and countries), there turns out to be a full counting bonus that is positive and of significant size. In general, the bonus is highest at the level of authors and lowest at the level of countries. We have seen in Figure 2 that the number of countries of a publication has a much stronger effect on a publication's citation score than the number of authors, but apparently this is offset by the fact that publications with a large number of countries occur much less frequently than publications with a large number of authors, as shown in Figure 1. The full counting bonus at the level of organizations is generally in between the country-level and author-level bonuses, although there are two main fields (i.e., 'Biomedical and health sciences' and 'Life and earth sciences') in which the organization-level bonus is higher than the author-level one.

The results reported in Table 3 also indicate that at the levels of authors and organizations the full counting bonus is lowest in the 'Mathematics and computer science' main field. At the country level, 'Social sciences and humanities' is the main field with the lowest bonus. The 'Natural sciences and engineering' main field has the highest bonus at the level of authors, while the highest bonus at the organization and country level can be found in the 'Biomedical and health sciences' main field.

The results for the $PP_{top\ 10\%}$ indicator reported in Table 4 are quite similar to the MNCS results presented in Table 3. However, full counting bonuses turn out to be consistently higher for the $PP_{top\ 10\%}$ indicator than for the MNCS indicator.

More detailed results at the level of 250 WoS journal subject categories can be found in an Excel file that is available at www.ludowaltman.nl/counting_methods/. The Excel file also indicates how the five main fields listed in Tables 3 and 4 are defined in terms of the WoS journal subject categories. There turn out to be rather large differences between subject categories in the full counting bonus. For instance, the subject categories with the highest MNCS full counting bonus at the level of organizations and countries are 'Medicine, general

& internal' and 'Physics, nuclear'. The subject categories have bonuses of, respectively, 148% and 176% at the organization level and 89% and 70% at the country level. Other subject categories have bonuses that are close to zero or even negative. Examples of such subject categories include 'Chemistry, organic' and 'Ergonomics'.
It is important to be aware of the consequences of the large differences between subject categories in the full counting bonus. Consider a university that has a full counting MNCS of 2.50 in the 'Medicine, general & internal' subject category and a full counting MNCS of 1.00 in the 'Chemistry, organic' subject category. What should we conclude based on these values? The obvious conclusion may seem to be that in terms of citation impact our university is performing much better in the 'Medicine, general & internal' subject category than in the 'Chemistry, organic' subject category. However, this conclusion does not take into account the effect of the full counting bonus. As mentioned above, the 'Medicine, general & internal' subject category has an organization-level full counting bonus of almost 150%, while the full counting bonus for the 'Chemistry, organic' subject category is close to zero. Taking into account the effect of the full counting bonus, we need to conclude that in both subject categories our university performs around the average level of all organizations worldwide.

Commonly used arguments in favor of full counting

In practice, most bibliometric analyses use full counting instead of fractional counting. Below we list three arguments that are often given to argue against the use of fractional counting and to justify the use of full counting. We also provide a response to each argument.

Argument 1: The different co-authors of a publication usually have not contributed equally. By giving equal weight to each co-author, fractional counting fails to properly represent the contributions made by the different co-authors. Hence, giving equal weight to each co-author is arbitrary and lacks a sound justification.

It is true that there can be large differences between co-authors in the contribution they have made to a publication. At the level of an individual publication, fractional counting may therefore significantly misrepresent the contributions made by individual co-authors. However, at the level of a large set of publications, for instance all publications of an organization or a country, we believe that it is reasonable to assume that the error will be within an acceptable margin. This is because errors at the level of individual publications are likely to cancel out. The contribution of an organization or a country to certain publications may be overestimated, but most probably there will then be other publications for which the contribution of this organization or this country is underestimated.
Furthermore, the argument that giving equal weight to each co-author of a publication is arbitrary may equally well be used as an argument against full counting. Like fractional counting, full counting gives the same weight to each co-author of a publication.

Argument 2: Fractional counting provides an incentive against collaboration, which is often considered undesirable.

We believe that citation impact and collaboration represent different dimensions of scientific performance and that in general these dimensions can best be measured separately from each other. Citation-based indicators should be assessed based on the degree to which they measure citation impact in an accurate way. In this respect, we believe that for many purposes fractional counting performs better than full counting. If in addition to citation impact one also considers collaboration to be a relevant dimension of scientific performance, then additional indicators should be used to measure this dimension. If one desires to do so, these indicators can then be used to provide an incentive to collaboration. By assessing citation-

based indicators based on the effect they may have on collaboration, one fails to make a proper distinction between the citation impact dimension of scientific performance and the collaboration dimension.

Argument 3: Fractional counting is more difficult to understand and less intuitive than full counting.

To a certain degree, we agree with this argument. Fractional counting yields non-integer publication and citation counts. These non-integer counts are more difficult to understand and require more explanation than the integer publication and citation counts provided by full counting. Fractional counting may also be less intuitive than full counting. For instance, consider a researcher who has produced some of his publications on his own while he has produced other publications with one or two co-authors. The researcher may feel that his co-authored publications are of similar importance to his oeuvre as his single-author publications. However, fractional counting gives less weight to the co-authored publications of the researcher than to his single-author publications. This is not in agreement with the feelings the researcher has about the importance of the different publications in his oeuvre, and therefore from the point of view of the researcher fractional counting can be regarded as less intuitive than full counting.

On the other hand, from a different point of view, it can also be argued that fractional counting is actually more intuitive than full counting. Earlier in this paper, we have given two examples showing that field-normalized citation impact indicators calculated using full counting can easily be misinterpreted. Field-normalized indicators calculated using fractional counting are much more easy to interpret in a correct way. As we have explained, this is because indicators based on fractional counting yield results that are compatible with the idea of strong field normalization. Unlike full counting indicators, fractional counting indicators therefore allow comparisons between fields to be performed in an easy and intuitive way. So from this point of view indicators based on fractional counting can be considered more intuitive than their full counting counterparts.

Conclusions

In this paper, we have presented a new perspective on the choice between different counting methods, leading to an important new argument in favor of fractional counting. Building on our earlier work (Waltman et al., 2012), this argument is based on the observation that the problem of choosing an appropriate counting method is closely connected to the problem of field normalization of citation-based indicators.

We have argued that from a field normalization point of view fractional counting is preferable over full counting. As we have shown, properly field-normalized results cannot be obtained using full counting, and field-normalized indicators calculated using full counting can easily be misinterpreted. Fractional counting does provide properly field-normalized results, and these results can be interpreted in a much more straightforward way than results obtained using full counting. Essentially, the problem of full counting is that co-authored publications are counted multiple times, once for each co-author, which creates an unfair advantage to fields with a lot of co-authorship and with a strong correlation between co-authorship and citations. For instance, the average full counting MNCS of all organizations or all countries active in these fields is significantly higher than one. On the other hand, fields in which co-authorship is less common or in which co-authorship does not correlate with citations are disadvantaged. Full counting yields results that are biased against organizations and countries whose activity is focused on these fields. Fractional counting does not suffer from this problem. In the case of fractional counting, each publication is counted only once, regardless

of its number of co-authors, and this ensures that comparisons between fields can be made in an unbiased way.

What are the practical implications of the analysis presented in this paper? In our view, this depends on the level of aggregation at which a bibliometric study is performed. In the case of a study at a high aggregation level, such as the level of countries or organizations (e.g., university rankings), we consider it absolutely essential to use fractional counting instead of full counting. At this level, there is a serious risk of misinterpretation of full counting results. Moreover, we believe that arguments in favor of full counting, such as the ones discussed in the previous section, are of limited relevance at a high aggregation level.

The situation is more difficult at a low level of aggregation, for instance at the level of researchers or research groups. At this level, we believe that reasonable arguments can be given in favor of both full and fractional counting. Especially the third argument discussed in the previous section plays an important role at this level. As pointed out in this argument, full counting is in agreement with the intuitive idea that all publications of a researcher or a research group should be considered of equal importance.

However, there is a more fundamental reason why the argument presented in this paper in favor of fractional counting is less relevant at a low level of aggregation. The argument depends on the connection between counting methods and field normalization, but the entire idea of field normalization may be seen as problematic at a low aggregation level. Field-normalized indicators have a limited accuracy (e.g., Van Eck, Waltman, Van Raan, Klautz, & Peul, 2013), and it is questionable whether these indicators are sufficiently accurate for applications at a low aggregation level. If the accuracy of field-normalized indicators at a low aggregation level is considered insufficient, the argument presented in this paper in favor of fractional counting has no relevance at this level.

In this paper, we have not shown how results obtained using full and fractional counting differ in practice. We refer to our working paper (Waltman & Van Eck, 2015) for an extensive comparison of full and fractional counting in bibliometric studies at the level of institutions and countries. The working paper also considers different variants of fractional counting, and it studies first author and corresponding author counting methods.

References

Aksnes, D.W., Schneider, J.W., & Gunnarsson, M. (2012). Ranking national research systems by citation indicators. A comparative analysis using whole and fractionalised counting methods. *Journal of Informetrics*, *6*(1), 36-43.

Albarrán, P., Crespo, J.A., Ortuño, I., & Ruiz-Castillo, J. (2010). A comparison of the scientific performance of the U.S. and the European Union at the turn of the 21st century. *Scientometrics*, *85*(1), 329-344.

Gauffriau, M., Larsen, P.O., Maye, I., Roulin-Perriard, A., & Von Ins, M. (2007). Publication, cooperation and productivity measures in scientific research. *Scientometrics*, *73*(2), 175-214.

Gauffriau, M., Larsen, P.O., Maye, I., Roulin-Perriard, A., & Von Ins, M. (2008). Comparisons of results of publication counting using different methods. *Scientometrics*, *77*(1), 147-176.

Van Eck, N.J., Waltman, L., Van Raan, A.F.J., Klautz, R.J.M., & Peul, W.C. (2013). Citation analysis may severely underestimate the impact of clinical research as compared to basic research. *PLoS ONE*, *8*(4), e62395.

Waltman, L., Calero-Medina, C., Kosten, J., Noyons, E.C.M., Tijssen, R.J.W., Van Eck, N.J., & Wouters, P. (2012). The Leiden Ranking 2011/2012: Data collection, indicators, and interpretation. *Journal of the American Society for Information Science and Technology*, *63*(12), 2419-2432.

Waltman, L., & Schreiber, M. (2013). On the calculation of percentile-based bibliometric indicators. *Journal of the American Society for Information Science and Technology*, *64*(2), 372-379.

Waltman, L., & Van Eck, N.J. (2015). Field-normalized citation impact indicators and the choice of an appropriate counting method. arXiv:1501.04431.

Waltman, L., Van Eck, N.J., Van Leeuwen, T.N., Visser, M.S., & Van Raan, A.F.J. (2011). Towards a new crown indicator: Some theoretical considerations. *Journal of Informetrics*, *5*(1), 37-47.