Research Evaluation, 18(3), September 2009, pages 215–220
DOI: 10.3152/095820209X466919; http://www.ingentaconnect.com/content/beech/rev

Multi-authored publications: their influence in the distribution of the financing costs in world licenses

Juan Gorraiz and Martin Wieland

Due to the rising costs of periodicals and the necessity of assuring access to the research results of the scientific community, new publishing models are required. Most of them are based on the distribution of financing costs among all participating authors, affiliations or countries according to the number of their scientific publications. Thus the distribution of costs may depend very much on the manner in which the problem of multi-authorship is handled. Our study focuses on the dependence of the distribution of financing costs on multi-authorship handling for a case study: the hypothetical Austrian participation in the world project SCOAP³. Our results show that the distributions of costs differ widely. We also survey whether there is any correlation between the number of full-text downloads operated by several universities with the number of their publications in these journals. The correlation results vary depending on the type of journal: broader or specific journals.

THE IMPORTANCE of open access, offering free access to the research results of the scientific community, is increasing with the steadily rising costs of periodicals, obliging libraries and institutions to cancel a considerable amount of their subscriptions and reducing the access of their members to important scientific literature. Although most of the published articles in theses periodicals are publicly funded, their actual 'producers or creators' pay exorbitant prices to access and use them. This steadily increasing awareness is one of the reasons of the success and multiplication of repositories, also called the 'Green Road or Route to Open Access' (OA). However, repositories cannot provide the quality control guaranteed by peer-reviewing processes as offered by high-quality journals, while the publishers of these 'excellent' journals justify their high prices with the costs of peer reviewing and editorial processing, and with access control and subscription administration.

Juan Gorraiz and Martin Wieland are in the Bibliometrics Department, Library and Archive Services, University of Vienna, Boltzmanngasse 5, A-1010 Vienna, Austria; Email: juan.gorraiz@univie.ac.at.

The second possibility or route to open access (also called the 'Golden Road') is to create open access journals or platforms warranting free information availability of peer-reviewed articles on the public Internet. The costs arising by organising and maintaining these OA initiatives are, however, higher than presumed, reaching approximately €2,000 per publication.[1] So their viability is very strongly questioned and new alternatives and impulses are demanded. Concepts for the financing of the OA publications have especially focused on the author-pays business model (Hagenhoff *et al*, 2008).

At the same time, in order to fulfil their main function, warranting access to scientific information, libraries and institutions have reacted by building consortia in the sense of the collaborative acquisition of access rights to electronic journals. Currently some different organisational models are applied in addition to the consortium approach (consortia on a subject or geographical basis) or the pay-per-view model (end-users are allowed to register via publisher, licensing library or institution and to access the digital resource on a pay-per-use basis). Going beyond both models are the ideas of national licenses[2] and 'world licenses' in the broadest possible sense.

One of the most interesting and advanced initiatives in 'world licenses' is SCOAP³ (Sponsoring

Consortium for Open Access Publishing in Particle Physics) presented by CERN, a plan for the conversion of six peer-reviewed high energy physics (HEP) journals to OA, eliminating the costs of access control and subscription administration, through a consortium of international agencies, laboratories and libraries (SCOAP³ Working Party, 2007). The main goals are a further push to open access initiatives while retaining all the advantages of the peer-review system and, at the same time, to bring the spiralling cost of journal subscriptions under control and to build a sustainable model for OA publishing.

All in all, due to the steadily rising costs of periodicals and the necessity of assuring access to the research results of the scientific community, new publishing models are required. Most of the new alternatives are based on the replacement of subscription fees by the author-pays business model (as in BioMed Central [BMC]) or by paying publishers for the organisation of peer-review services and for guaranteed OA (as in SCOAP³).

In the first model (BMC), whoever submits the manuscript is responsible for making or arranging the payment. Usually it will be the first author (corresponding author) or for instance his institution. Authors or institutions then have the opportunity to find their own fair agreement. In the second model (SCOAP³), the financing costs will be distributed among all participating countries according to the number of its scientific publications, taking into account the relative fraction of authorship of articles published in these journals during the last two years. So, in both models, the distribution of the financing costs may depend very much on the manner in which the problem of multi-authorship is handled.

On the other hand, the number of multi-authored papers is increasing permanently. In his study, 'The future of single-authored papers' performed for four journals in the fields of astronomy, physics, chemistry and biology, Abt (2007) shows that the frequency of single-authored papers has been decreasing exponentially during the last 30 years (1975–2005), varying in 2005 between 5% (chemistry) and 20% (physics). Abt concludes that in coming years single-authored papers will decrease into an exponential that will never reach zero. These data are in agreement with those compiled by Farber (2005) at Israeli universities. She reports pproportions of single-authored articles oscillating between 27% (mathematics) and 10% (biomed). Significantly higher are the percentages in the social sciences (22–36%) and in the arts and humanities (higher than 80%).

ScienceWatch tracked multi-authored papers published in Thomson ISI-indexed journals from 1990 through 2003 (*ScienceWatch*, 2004). After the never entirely explained dip in 1991, the number of multi-authored papers ascended sharply. The average number of authors per Thomson ISI-indexed paper climbed from 2.6 in 1990 to 3.6 in 2003. Papers with 50 or more authors reached a zenith in 1997, while those with 100+ authors are permanently increasing.

In 1999 John Taylor introduced the term *e-Science (*or *eScience*) to describe a large funding initiative of the United Kingdom's Office of Science and Technology. The term is now commonly used to describe computationally intensive science that is carried out in highly distributed network environments. Currently there are large focuses on social simulations, particle physics, earth sciences and bioinformatics.

For example, HEP is a discipline with a well-developed e-Science infrastructure due to its need for adequate computing facilities for the analysis of results and storage of data originating from the CERN large Hadron Collider as well as its widely distributed collaborative nature. The results of such distributed collaboration, that involves large teams managed and developed by several research laboratories, universities and even governments, can be compiled only in publications with more than 300 authors (if not more than a thousand), and contribute to the increase in large multi-authored publications.

Study purpose and methodology

For all countries participating in world projects or licenses, as for example SCOAP³, the following questions now arise:

1. How to make their own costs distribution among the involved institutions?
2. Is the use of the fractional count method as proposed by SCOAP³ the more adequate one?
3. Are the efforts retrieving the required data justified?
4. How significant is the deviation by using other, simpler methods?
5. Which other alternatives are possible?
6. Is there some concordance between the distributions basing on the number of scientific publications and those basing on their usage (downloads)?

All these questions have been studied for the hypothetical case of an Austrian participation in SCOAP³. Our study focused on the dependence of the distribution of the financing costs on the way to handle the problem of multi-authorship. Furthermore we are discussing the most usual methods for allocating credits to authors and giving some recommendations for their use in the distribution of the financing costs.

Our analysis was conducted for the same six peer-reviewed HEP-journals considered in SCOAP³:

1. *Physical Review D* (American Physical Society)
2. *Physics Letters B* (Elsevier)
3. *Nuclear Physics B* (Elsevier)
4. *Journal of High Energy Physics* (SISSA/IOP)
5. *European Physical Journal C* (Springer)
6. *Physical Review Letters* (American Physical Society)

and extended to another 'broadband' journal with HEP instrumentation articles:

7. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* (Elsevier)

This data pool is very adequate for our analysis because multi-authored works and publications of international collaborations and working groups are very common in HEP, a discipline with a well-developed e-Science infrastructure. The journal selection is also very well balanced because the most important publishers in the sector are represented.

We limited our analysis to all these publications with at least one Austrian affiliation. The data (368 articles) were retrieved in Web of Science (WoS) for the years 2005 and 2006, where the seven journals are cover-to-cover indexed, but authorship counts (and the necessary affiliation disambiguation) could be performed only by autopsy, because in WoS affiliations and authorships are not correctly correlated. One author with two or more affiliations was directly assigned only to his first affiliation. So 4.9% of publications with Austrian affiliations could not be considered following the rules used in SCOAP[3]. All the Austrian affiliations were manually disambiguated and allocated to their corresponding university or institution.

Subsequently we calculated the costs distribution for all the involved Austrian institutions, allocating credits to their authors in three different ways (Egghe, 1990):

1. By using adjusted or fractional count: every co-author is assigned a fraction of the authorship.
2. By using normal counts: giving full credit to all co-authors.
3. By using straight counts: giving the first author all the credit.

Other methods, such as the proportional counting proposed by Van Hooydonk (1997), refining the fractional counting by considering also the rank, were not used because of their difficult calculation and the general argument that only the first and last ranks are of importance by citing.

Because of the data confidentiality we are replacing the real names of institutions with standardised denominations such as University 1, University 2, etc.

Results

Figures 1 and 2 show the distributions of co-authors and affiliations. The number of single-authored publications remains under the expected 5%, only 2.2%. Co-operations between two and six authors account for more than 50% of the publications.

The mean number of authors is between 5 and 6. Interesting is the low number of publications with 11 to 100 authors and the large increase of publications with more than 100 authors, which could represent a threshold for e-Science activity. Publications with 101 to 200 authors are almost a quarter of the total.

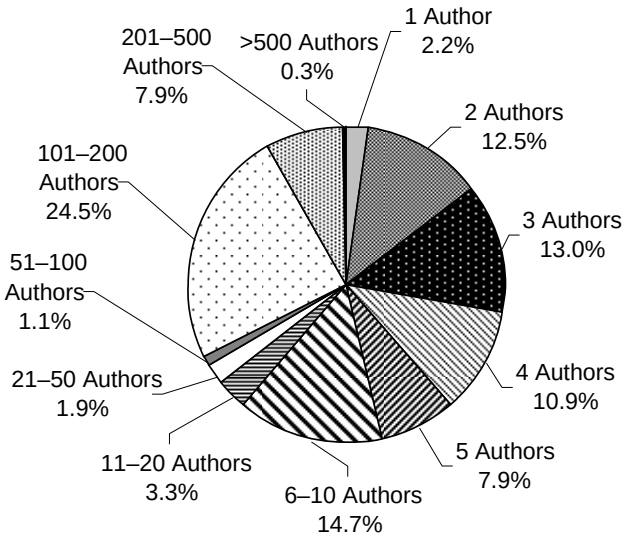

Figure 1. Co-authors distribution for the sample studied

The distribution of affiliations is very similar (see Figure 2). Co-operations of four affiliations account for more than 50% of the publications. The large number of publications with 101 to 200 authors must correspond with 41 to 50 affiliations.

The results for the varying costs distributions are shown in Figures 3, 4, 5 and 6.

Discussions and conclusions

First, the costs distributions are completely different and could be easily manipulated in order to defend the interests of individual institutions. In Austria, academies are the institutions with larger multi-authored publications and therefore are profiting the

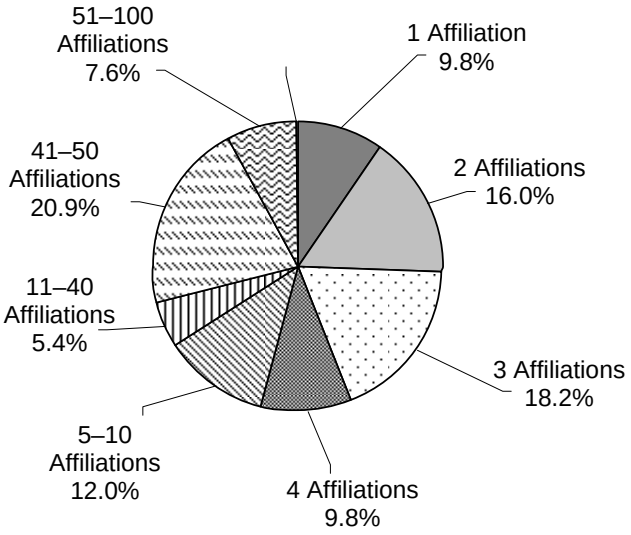

Figure 2. Distribution of the affiliations for the sample studied

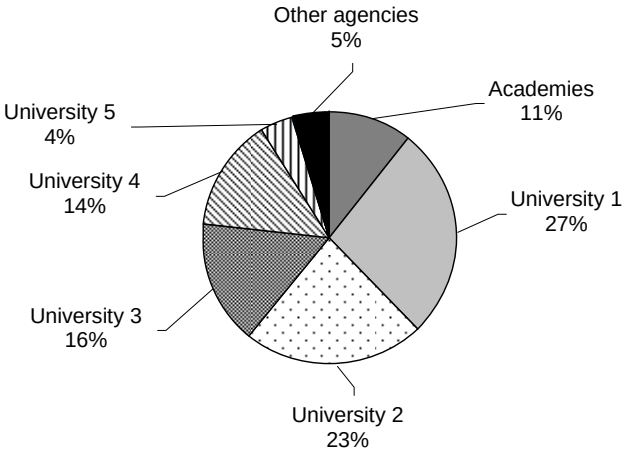

Figure 3. Costs distribution among Austrian institutions by using fractional counts

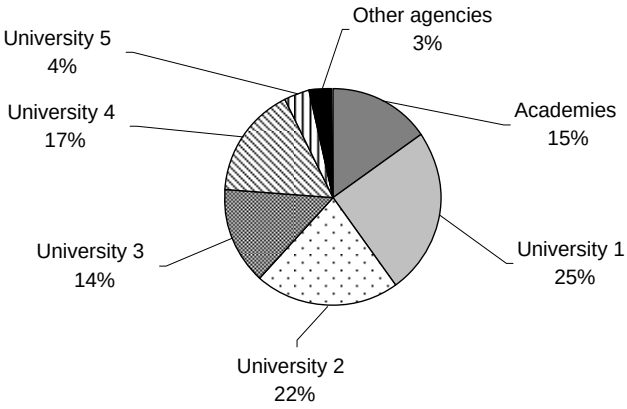

Figure 5. Costs distribution among Austrian institutions by using straight counts

most by using the adjusted or fractional count method for the distribution of the costs (11% instead of 54% by using normal counts) in detriment of the universities.

Second, the method by using adjusted or fractional count seems to be the more equitable one. Co-authors of very large multi-authored publications, however, meet disadvantages. So we found several authors with approximately 50 to 100 publications not having even one third of the credit! One possible solution to this problem could be the introduction of a threshold for the fractional count, of maximum 10.

Another disadvantage of this time-consuming method is the difficulty in retrieving the correct data by very large international co-operations or licensing projects. There is no citation database calculating automatically the fractional count for every author. For example, in WoS, it would not be possible because of the missing correlation between authors and affiliations mentioned above. Scopus offers this possibility but the fractional count can be calculated only manually. The database arXiv.org is surely the best alternative for this special area (HEP). Citebase Search, a product of Tim Brody from the University of Southampton, offers even the opportunity to retrieve the number of citations and downloads of every article, and their fractional counts. However, this database is suitable only for some areas of physics, mathematics, computer science, statistics and quantitative biology.

Third, the approach by using normal counts inflates the publication scores of those researchers producing many multi-authored papers and produces dramatic changes in the costs distributions (compare 'Academies' in Figures 3 and 4), particularly for institutions engaged in large international co-operations. This method also inflates the sum of the number of publications of all authors becoming larger than the number of papers under study (1,036:368). Curiously, in this study, the approach by using fractional counts deflates the number of papers at approximately the same factor of one third (368:126). Finally, and as indicated by Van Hooydonk (1998), it should be mentioned that most of the personal or institutional citation scores are obtained by standard (also inflationary) normal counting and not by fractional counts, which is inconsistent.

Fourth, the approach by using straight counts is the simplest one, reducing the work required to collect the data, but is also the most disputed. In this case study for Austria, it leads to very similar results by using fractional counts in the costs distribution (compare Figures 3 and 5) as well as in the deflation of the number of papers under study (368:125). While the first author place does not have the same significance in all disciplines, it is however very

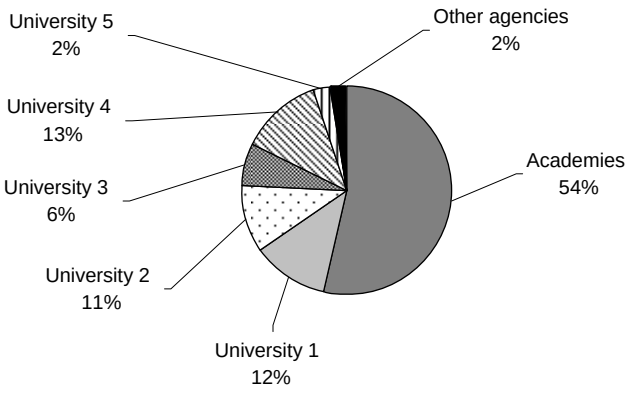

Figure 4. Costs distribution among Austrian institutions by using normal counts

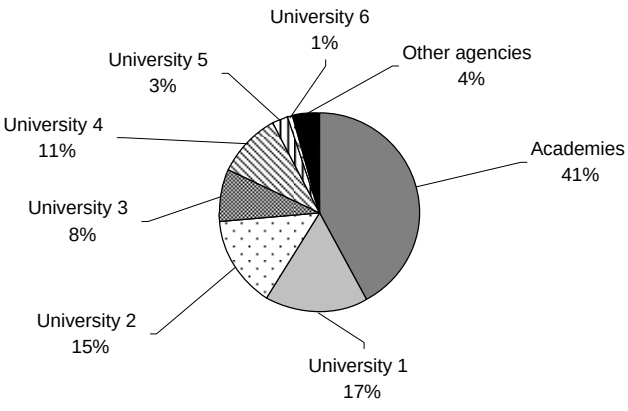

Figure 6. Costs distribution among Austrian institutions by using affiliation normal counts

often used by publications with a very large number of co-authors. This approach could be used for very quick approaches and also to help to correct the deflation of publication scores of researchers with a very large number of co-authors by using the fractional count.

Fifth, probably the most equitable method would be a combination of these three approaches depending on the publications characteristics of each discipline, but this effort would not compensate workload and time investment.

Why not directly use affiliations and not author counts?

Hence we have tried an easier way, calculating the costs distribution by directly using affiliation (and not author) normal counts. The results are shown in Figure 6 and they obviously reduce the inflation of the number of papers under study produced by normal counts in a considerable way (429 instead of 1073).

The pros of this approach are:

- It involves easier calculation, not requiring autopsy;
  and
- It seems to give a middle balance between allocating credits to their authors by using fractional and normal counts (compare Figures 3, 4 and 6).

The cons are that the proportionality to authorship is lost, and that multiple affiliations are not considered and can produce a significant shift of the results.

In a further study we analyzed the correlation of these approaches with the passive usage (number of downloads) operated by some of the involved institutions.

Correlations with downloads

For some of these journals, depending on data availability, we compared the number of full-text downloads operated by several universities with the number of their publications in these journals, allocating their credits in different ways: fractional count and normal count considering authorship and affiliation.

Due to their confidentiality and marketing strategies, the collection of the required data was not easy and we had some restrictions. Unfortunately it was not possible to get data from the academies. Finally, we succeed in collecting data for five journals from five universities. The results are shown in Tables 1, 2, 3, 4 and 5.

The correlation results are very different depending on the type of journals. For journals with a large number of downloads (broader journals as journal A) there is a very good correlation between number of operated downloads and number of publications. A better correlation is reached by considering authorships instead of affiliations. No considerable

Table 1. Downloads versus number of publications considering authorship and affiliation, fractional and normal counting for Journal A

Journal A	Univ. 1	Univ. 2	Univ. 3	Univ. 4	Univ. 5
Downloads	24%	22%	11%	10%	33%
Authorship, FC	29%	22%	9%	10%	30%
Authorship, NC	26%	23%	9%	8%	34%
Affiliations, FC	28%	25%	9%	10%	28%
Affiliations, NC	28%	27%	8%	7%	30%

Note: FC = fractional counting; NC = normal counting

Table 2. Downloads versus number of publications considering authorship and affiliation, fractional and normal counting for Journal B

Journal B	Univ. 1	Univ. 2	Univ. 3	Univ. 4	Univ. 5
Downloads	26%	26%	12%	26%	10%
Authorship, FC	15%	26%	0%	59%	0%
Authorship, NC	11%	24%	0%	65%	0%
Affiliations, FC	12%	37%	0%	51%	0%
Affiliations, NC	13%	38%	0%	49%	0%

Note: FC = fractional counting; NC = normal counting

Table 3. Downloads versus number of publications considering authorship and affiliation, fractional and normal counting for Journal C

Journal C	Univ. 1	Univ. 2	Univ. 5
Downloads	18%	38%	44%
Authorship, FC	52%	26%	22%
Authorship, NC	9%	12%	79%
Affiliations, FC	57%	24%	19%
Affiliations, NC	23%	23%	54%

Note*: FC = fractional counting; NC = normal counting*

Table 4. Downloads versus number of publications considering authorship and affiliation, fractional and normal counting for Journal D

Journal D	Univ. 1	Univ. 2	Univ. 4	Univ. 5
Downloads	87%	7%	3%	10%
Authorship, FC	39%	50%	11%	0,5%
Authorship, NC	33%	23%	4%	39%
Affiliations, FC	41%	48%	11%	0,1%
Affiliations, NC	43%	25%	13%	19%

Note: FC = fractional counting; NC = normal counting

Table 5. Downloads versus number of publications considering authorship and affiliation, fractional and normal counting for Journal E

Journal E	Univ. 1	Univ. 2
Downloads	42%	58%
Authorship, FC	100%	0%
Authorship, NC	100%	0%
Affiliations, FC	100%	0%
Affiliations, NC	100%	0%

Note: FC = fractional counting; NC = normal counting

differences between fractional counts and normal counts were found, but normal counting shows the strongest correlation.

For journals with a reduced number of downloads (specialized journals such as journal B, C and D), the results give a very weak or no correlation between number of operated downloads and number of publications. There was no better correlation considering affiliations instead of authorships and no better results using fractional counts.

However, universities without publications operated downloads. A socket from 10% was found (around 100 downloads, see Table 2). Interestingly, for Journal E, University 2 without its own publications operated more downloads than University 1 with 9 publications (normal count) or 3.5 (fractional count).

General conclusions

The following general conclusions can be drawn from our study:

1. The costs distributions are strongly varying depending on the way to handle the problem of multi-authorship.
2. The use of fractional counting as proposed in SCOAP³ is not so evident: the controversy 'fractional or normal counts' is open.
3. The effort involved in calculating costs distributions by using fractional counting is considerable and problematic for very large licensing projects: there are no citation databases considering fractional counts.
4. The effort of considering authorship and using fractional counting is not supported by correlations considering number of publications versus downloads except for journals with very large numbers of downloads.
5. Using affiliations instead of authorships reduces the effort but it is not evidently preferable.
6. Database producers are not paying adequate attention to this fact and there are only a few that are considering or allocating credits to the authors by using adjusted or fractional counts. Bibliometric data sources will barely be able to abstain from different options to handle the problem of multi-authorship (inclusion of fractional counts).
7. New publishing models are basing on the distribution of the financing costs among all participating authors, affiliations or countries according to the number of their scientific publications. So, the way to handle the problem of multi-authorship will be crucial and, until now, there is no commonly accepted and satisfactory solution to this problem. Most of the personal or institutional citation scores are obtained by standard normal counting, which is inconsistent with the use of fractional counts when calculating costs distributions.

Acknowledgements

The authors would like to thank Wolfgang Lucha from the Austrian Academy of Sciences, and Brigitte Kromp and Wolfgang Mayer from the Vienna University Library for their support, input and comments.

Notes

1. See, for example, 'Comparison of BioMed Central's Article Processing Charges with those of other publishers' under the link <http://www.biomedcentral.com/info/authors/apccomparison>.
2. See, for example, <http://www.nationallizenzen.de/>.

References

Abt, H A 2007. The future of single-authored papers. *Scientometrics*, **73**, 353–358.

Egghe, L and R Rousseau 1990. *Introduction to Informetrics, Quantitative Methods in Library, Documentation and Information Science*. Elsevier, Amsterdam.

Egghe, L, R Rousseau and G van Hooydonk 2000. Methods for accrediting publications to authors or countries: consequences for evaluation studies. *Journal of the American Society for Information Science*, **51**, 145–157.

Farber, M 2005. Single-authored publications in the sciences at Israeli universities. *Journal of Information Science*, **31**(1), 62–66.

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

Hagenhoff, S, M Blumenstiel and B Ortelbach 2008. An empirical analysis of the amount of publication fees. *Serials Review*, **34**(4), 257–266.

ScienceWatch 2004. Crowd control multiauthor papers appear to level off in recent years. *ScienceWatch*, **15**(4), July/August <http://www.sciencewatch.com/july-aug2004/sw_july-aug2004_page1.htm>.

SCOAP³ Working Party 2007. Towards Open Access Publishing in High Energy Physics. Report of the SCOAP³ Working Party, CERN, Geneva, available at <http://open-access.web.cern.ch/Open-Access/Scoap3WPReport.pdf>.

Van Hooydonk, G 1998. Fractional counting of multi-authored publications: consequences for the impact of authors. *Journal of the American Society for Information Science*, **48**, 944–945.