



RESEARCH ARTICLE

# A validation of coauthorship credit models with empirical data from the contributions of PhD candidates

Paul Donner 

German Centre for Higher Education Research and Science Studies, Department 2, Research System and Science Dynamics, Schützenstraße 6a, 10117 Berlin, Germany

**Keywords:** author contribution, coauthorship, coauthorship credit, fractional counting, harmonic counting, validation study

## ABSTRACT

A perennial problem in bibliometrics is the appropriate distribution of authorship credit for coauthored publications. Several credit allocation methods and formulas have been introduced, but there has been little empirical validation as to which method best reflects the typical contributions of coauthors. This paper presents a validation of credit allocation methods using a new data set of author-provided percentage contribution figures obtained from the coauthored publications in cumulative PhD theses by authors from three countries that contain contribution statements. The comparison of allocation schemes shows that harmonic counting performs best and arithmetic and geometric counting also perform well, while fractional counting and first author counting perform relatively poorly.

## 1. INTRODUCTION

The social creation of knowledge and its dissemination as publications is of critical importance to the modern science system. The publications under a researchers' name attest to their accomplishments and form the basis of the reward system in science by having a major influence on recognition by colleagues and career advancement. The fair sharing of authorship credit of coauthored publications, especially in the context of research evaluation, is therefore a significant topic of research in itself (Egghe, Rousseau, & van Hooydonk, 2000; Gauffriau, Larsen, et al., 2008).

Coauthorship credit is a theoretical concept that refers to the idea that one may conceive of a publication as being associated with a mathematical object with unity value (1.0), on which mathematical operations can be performed. By this transformation the publication's authors can be credited with relative shares of the whole unit. These shares can be referred to as their *respective credits* or *partial publication equivalents*. A paper's credit is an abstraction, useful for analytical and especially evaluative purposes, but it does not exist as such in reality. A real paper can of course not be arbitrarily divided between coauthors. To give an example, a two-author paper may be accompanied by a statement indicating that both authors contributed equally to the work. Here, each author made a 50% contribution and claims 50% of the paper's credit. Because only a few publications state contributions explicitly and numerically, it is crucial in bibliometric practice to use a credit allocation method that is on average in close agreement with typical contributions.

an open access  journal



Citation: Donner, P. (2020). A validation of coauthorship credit models with empirical data from the contributions of PhD candidates. *Quantitative Science Studies*, 1(2), 551–564. [https://doi.org/10.1162/qss\\_a\\_00048](https://doi.org/10.1162/qss_a_00048)

DOI:  
[https://doi.org/10.1162/qss\\_a\\_00048](https://doi.org/10.1162/qss_a_00048)

Received: 15 January 2020  
Accepted: 04 April 2020

Corresponding Author:  
Paul Donner  
[donner@dzhw.eu](mailto:donner@dzhw.eu)

Handling Editor:  
Staša Milojević

Copyright: © 2020 Paul Donner.  
Published under a Creative Commons  
Attribution 4.0 International (CC BY 4.0)  
license.



In this connection the number of authors and the position in the author order can be valuable clues to individual contributions, because in many fields, by convention, author ordering is based on relative contribution. In general, the first author made the greatest contribution, with successive authors having made successively smaller or equal contributions.

For any specific publication, the combination of author order and author count (and corresponding authorship as a further clue) can never be an accurate indicator of authors' contributions, as these vary between papers for the same combinations of byline position and number of coauthors. However, author credit models that show high agreement with empirical data can be very valuable for improving credit calculation on the level of aggregated publication sets, where such case-to-case variation can be expected to cancel out with large enough numbers of observations. It has been demonstrated a number of times that the choice of authorship credit allocation method (also called *counting method*) has a substantial influence on the results of bibliometric studies, not only on the level of authors, but even on the levels of institutions and countries (Gauffriau et al., 2008; Gauffriau & Olesen Larsen, 2005; Huang, Lin, & Chen, 2011; Lin, Huang, & Chen, 2013). Furthermore, the values of the credit allocation method are often also constituent parts of advanced citation indicators. They are used as weights for the citation indicators of specific publications for the units (i.e., authors, organizations, countries) that contributed. The question of the most appropriate coauthor credit allocation method therefore deserves comprehensive investigation.

In this paper we only consider methods which split the total publication credit in some way across the coauthors and add up to 1.0, as only these methods avoid artificially inflating publication counts. Furthermore, flexible credit allocation schemes modified by parameters are not considered, as these would need tuning based on extensive empirical data appropriate to an anticipated application (i.e., a discipline), which could lead to overfitting in the general case. An overview of various author credit assignment methods is given in Waltman (2016). The "whole counting" scheme, in which every coauthor gets a full publication credit regardless of the number of coauthors, leads to distortions of paper counts, a phenomenon also referred to as *authorship inflation*. Lindsey (1980) argued against the use of whole counting and first author counting, which was then prevalent in the social studies of science literature. While whole counting causes multiplication of authorship credit, first author counting also leads to distortions, as it is not a viable sampling strategy to consider authors' first-authored papers as representative of their entire work because the order of authors in coauthored papers is not random. He proposed to divide the unit authorship credit by the number of authors and called this "adjusted counts." De Solla Price (1981) also suggested—"in the absence of evidence to the contrary"—to divide one whole unit of publication credit equally by the number of authors to counteract authorship inflation. Their method, now mostly referred to as "fractional counting," is nowadays commonly used (Waltman, 2016) but apart from Lindsey (1980) there are no studies that contain empirical data on how prevalent the different methods are in bibliometric studies.

To briefly review the methods investigated here, the fractional counting method assumes the contribution of each author of multi-authored papers to be equal and assigns each author  $\frac{1}{N}$  of the total credit (which is 1.0), where  $N$  is the number of authors. As other methods also assign fractions of a whole publication to authors, it might be better to refer to this method as *equal fractional counting*. In the first author counting method, the first author receives all of the credit and all further coauthors nothing, under the assumption that the first author is the most important contributor (cf. Cole & Cole, 1974), who called it *straight count*). One can consider these two methods as the extremes of total equality and total inequality. Geometric counting (Egghe et al., 2000), arithmetic counting (also called *proportional* or

*positionwise counting*) (Kalyane & Vidyasagar Rao, 1995; van Hooydonk, 1997) and harmonic counting (Hagen, 2008; Hodge & Greenberg, 1981), on the other hand, assign different credit shares based on the number of authors  $N$  and on the authors' positions  $i$  of a publication, using the formulas given below.

$$C_{\text{geometric}} = \frac{2^{N-i}}{2^N - 1}$$

$$C_{\text{arithmetic}} = \frac{N + 1 - i}{1 + 2 + \dots + N}$$

$$C_{\text{harmonic}} = \frac{\frac{1}{i}}{1 + \frac{1}{2} + \dots + \frac{1}{N}}$$

Details can be found in the respective cited sources. In many of these publications, the respective methods have merely been proposed, with no attempts made to validate them.

We mention in passing that there are plenty further suggestions of even more elaborate credit allocation schemes, based more or less on the whim of their proponents rather than any empirical data. By all logic the empirical investigation of which authors contributed how much to publications should be conducted first. Only after that is it reasonable that models be proposed that approximate the reality as much as possible.

Hagen (2010), being the exception to that observation, validated (perceived) coauthor credit shares from prior studies in chemistry, psychology, and medicine. It was found that harmonic counting fits the data better than arithmetic, geometric, or fractional counting. It should be noted that the mean values of credit values were used, such that there is only one value for each possible combination of author position and author count. In contrast, for the present study we are able to make use of observations on the level of individual credit statements of authors of particular publications. The disadvantage of the former is that some of the aggregate data points could be computed from more primary data points than others, thus distorting the influence of these data points, which could be avoided by appropriate weighting. Furthermore, information on variability within the observed values is lost.

A further conceptual difference is that in the present study we do not utilize *perceptions* of typical author credit of *readers* of multiauthored papers but public *statements* by *authors* themselves. In Hagen's (2010) study, the data for psychology, obtained from Maciejovsky, Budescu, and Ariely (2009) and medicine, obtained from Wren et al. (2007), was not based on authors' judgments about their own contributions, but on the perceptions of researchers facing typical publications with specific numbers of authors and corresponding authorship statements. In order for this type of data to be applicable for the purpose of validating author credit assignment methods, it would first need to be corroborated by comparison with data of judgments by authors themselves. It goes without saying that for the determination of authorship credit, the authors' statements of contribution are more relevant than the impressions (or guesses) of readers.

For the chemistry data, Hagen (2010) uses figures from Vinkler (2000, Table 4, p. 608). However, these data are not the empirical data but a simplified derived scheme devised by Vinkler for use in his institute's internal evaluation practice based on the research reported in Vinkler (1993). In this latter paper, Vinkler was concerned about the unreflected use of first author counting prevalent in scientometric research at the time. He surveyed authors from his

chemistry research institute on their activities for specific publications on seven types of activities, such as experimental work, analysis and evaluation of experimental data, and writing the text. For each publication and activity, participants judged how much they contributed on a scale of five levels (100%, >50%, ~50%, <50%, ~10%). But it seems for later calculations that these were transformed to the integers 1 to 5 in descending order (Vinkler, 1993, note to Table 5). This means that the ratio across the range of values of originally 10:1 (100%:~10%) has been compressed to 5:1. For each of the activities, the author derived importance weighting factors by averaging the responses of researchers as to how important those activities are in producing chemistry papers, which were given in percentages. These weights range from 1.0 to 3.0. The data that come closest to empirical credit distribution are what Vinkler calls *Percentage Total Contribution Factors* and are derived from the empirical activity data and weights, which were summed and averaged across authorship positions and author counts. The exact calculation is difficult to reconstruct from the original publication. The important part is that the equivalent to credit shares are the rows labeled “TCF%” in Table 5 of Vinkler (1993). Summarizing these, we present the relevant figures in Table 1. We have also added the numbers of observations for each row as extracted from the text of Vinkler (1993, p. 217).

If one compares the figures in Table 1 with those in Table 4 of Vinkler (2000), which are the same as in Hagen (2010), one can see that Hagen did not use Vinkler’s empirical data. Rather, he used a weighting scheme employed by the evaluation committee of Vinkler’s institute at the time. Kim and Kim (2015) use both the 1993 and 2000 data, assuming both to be empirical data. But even using the more correct data from Vinkler (1993), we would object that these are not contribution statements from authors regarding the amount of work they contributed to a paper but elaborate attempts at reconstructions based on very roughly quantified statements to various activities.

There is also an issue with the usage of the Wren et al. (2007) data. In this study, perceived contributions in three pre specified categories, initial conception, work performed, and supervision, were collected, but overall contribution was not. Instead, the figures for the three categories are averaged and these are the numbers used by Hagen (2010). This simple averaging implicitly assumes that the three contribution categories are equally important, but this assumption is not substantiated in the paper at all.

### 1.1. Contribution-Based Authorship Order

The ordering of coauthor names on publications by contribution is widespread. It is conventional, for example, in management, according to journal editors (von Glinow & Novelli Jr.,

**Table 1.** Percentage total contribution factors, extracted from Vinkler (1993, Table 5)

Author count (observations)	1st author	2nd author	3rd author	4th author	5th author
2 (6)	71	29			
3 (12)	61	26	13		
4 (13)	54	31	9	6	
5 (11)	34	14	11	17	24

1982), in nursing, according to nurses expected to publish (Butler & Ginn, 1998), in library science, according to authors (Hart, 2000), medicine, according to Cochrane reviews authors (Mowatt, Shirran, et al., 2002), editorial board members (Bhandari, Einhorn, et al., 2003), and promotion committee members (Wren et al., 2007) and educational science, according to authors (Moore & Griffin, 2006).

## **1.2. Authorship Status Signals Not Based on Contribution**

### **1.2.1. Alphabetical authorship**

In disciplines in which alphabetical authorship order is the accepted convention, it would obviously be unreasonable to apply credit allocation methods that apply weights based on author position. However, the problem arises of how to handle publications in these “alphabetized” disciplines that manifestly deviate from the norm of alphabetical order. In those cases it can be reasoned that the authors were either not aware of the norm or they intentionally refrained from submitting to it. In either case, the assumption of equal contributions ought to be rejected. The next difficulty is that in many cases it is simply not possible to determine whether a group of authors followed the alphabetical ordering convention or if their deliberately chosen author order just happened to be alphabetical, even though it was intended to reflect relative contribution. For example, the prevalence of alphabetical order for two authors by chance alone is 50%. But as the number of coauthors increases, the chance of coincidental alphabetical ordering quickly tends to zero. From the perspective of authors who would like to express their relative contributions by authorship order, alphabetical ordering convention can be an impediment. In cases when the contribution order coincides with alphabetical order, the alphabetical norm would inadvertently lead to distorted perceptions of contributions.

Waltman (2012) has conducted a large-scale empirical study on the prevalence of intentionally alphabetical authorship in coauthored publications; that is, alphabetical coauthorship corrected for the probability of incidental alphabetical ordering. The share of intentionally alphabetically ordered coauthored publications has decreased from about 9% in 1981 to about 4% in 2011. On the level of disciplines, as approximated by Web of Science subject categories, there are stark differences in the use of intentional alphabetical authorship order. It is common in “Mathematics,” “Business, finance,” “Economics,” and “Physics, particles & fields.” However, also in these disciplines, alphabetical order is far from universal in the studied period (2007–2011). The percentages of alphabetical order range from 73% to 57%. The rates seem to be declining lately in “Mathematics” and “Economics” while alphabetical authorship has increased over time in “Business, finance.” It is important to keep in mind that the rates may appear higher to researchers from these fields because of incidental alphabetical ordering.

### **1.2.2. Last authorship and corresponding authorship**

There is as of yet no conclusive empirical evidence from prior studies about the specific signal of corresponding authorship and last authorship with respect to contribution to publications across different disciplines. While there are conventions in disciplines by which group leaders are indicated by last authorship, (a) this does not allow the conclusion that all last authors are group leaders and (b) neither last authorship nor corresponding authorship in themselves tell anything about how much these authors have, on average, contributed.

Laudel (2002), p. 11, based on interviews and publication data, reports on authorship position conventions in highly collaborative work in interdisciplinary research at the intersection of biology, physics, and chemistry:

Nearly all coauthors were ordered in the following way. The first author is the scientist who conducted the experimental work; that is, a doctoral student or a postdoctoral fellow. [...] The seven publications with the permuted order are of special interest. In these cases the group leader was first author, while another experimenter was listed last. In two cases the group leader did not change from the experimental role to the conceptual role but assumed both. This seems due to the group's specific content of work (development of research techniques and instruments). [...] In the case of a DOL [division of labor], usually the scientist who conducted the larger part of the experimental work is the first author, followed by the experimenter of the collaborating group. Both group leaders are last authors on the coauthorship list.

In this case, the conventions of listing research group leaders last and of authorship order by decreasing overall contribution to the work were used. The mentioned exceptions to the last position rule suggest that the contribution-based ordering convention overrides the group-leaders-last convention for the studied community of researchers.

Wren et al. (2007) report on a survey of the perceived authorship credit and influence of last author position and corresponding authorship in medicine. They elicited perceptions of North American promotion committee members on the perceived credit percentages of hypothetical three- and five-author papers in three specified contribution categories: initial conception, work performed, and supervision. The results indicate that last authors are perceived as having made important contributions to initial conception and supervision (about 50%), but far smaller contributions to work performed. The two different conditions for indicated corresponding author in five-author papers suggest that a corresponding author is perceived to have contributed substantially to initial conception and supervision, but far less so if the author is listed as the middle author rather than the last author. To a lesser degree, the perception of contribution to the category of work performed also increases if the middle author is the corresponding author. From the data no conclusions to overall contribution can be drawn, because it is unknown how the three work categories would combine to an overall contribution.

Sauermann and Haeussler (2017) study the number of listed contribution categories in the multidisciplinary journal *PLOS ONE* per author position. They point out that such statements are not informative about how much a given author contributed to the category and how important the category was for the given paper. Corresponding authors contributed on average to more categories than other authors, and this effect can be found across all authorship positions.

It must be concluded that extant studies give an inconclusive picture of the interaction of contribution, corresponding authorship, and last author position. More empirical studies are needed before discipline-specific adjustments to authorship credit assignment methods based on last or corresponding authorship can be considered. We therefore restrict our study to the information contained in author count and position and refrain from making adjustments to the studied credit methods, as it would be premature to do so at the current state of knowledge.

### **1.3. The Contribution of This Paper**

In the remainder of the paper we comparatively validate the quality of several coauthor credit assignment methods using a novel empirical data set. We collected explicit numerical



contribution statements on the micro level of authors and papers from German, Australian, and New Zealand cumulative dissertation theses. This is the first study of this specific kind, as prior comparable research did not use author contribution statements and only used aggregate data. Hence, we are able to conduct a validation study for coauthorship credit allocation methods that uses the most reliable and fine-grained micro level data available to date.

## 2. METHODS AND DATA

The empirical data used in this study is derived from contribution statements of PhD candidates for coauthored publications used in cumulative dissertation theses. Cumulative dissertation theses are theses that consist entirely or partially of published material, such as journal papers, book chapters, or conference papers.

PhD theses of German graduates were searched on Google Scholar, which indexes the full texts of theses archived in many university repositories. Searches were made by the German and English keywords “Eigenanteil”, “eigener Anteil”, “Eigenbeitrag”, “Anteilserklärung”<sup>1</sup>, and “self-contribution” in combination with keywords for PhD theses (“dissertation”, “PhD thesis”, and “doctoral thesis”).

Furthermore, we searched for universities with English language guidelines for coauthorship contribution statements in cumulative PhD theses on Google with the keyword “theses” in combination with “contribution statement” or “coauthorship statement”, which yielded a number of universities with such policies in Australia and New Zealand. For each identified university, we searched the university’s institutional repository by a full text search for such statements and inspected the found hits.

All found dissertations were checked and statements of contribution to coauthored published articles, proceedings papers, and edited book chapters given in percentages were manually extracted. If equally shared first authorships were declared, the same percentage was used for the indicated authors, as typically only that of the dissertation author was specified explicitly. For two-author publications for which the contribution percentage for one author is given explicitly we also recorded the remaining proportion to 100% for the other author as this can be inferred directly in these cases. We also used percentage figures for article coauthors other than the thesis author, where these were given. Verbal narrative declarations of contribution were not considered. Declarations giving separate percentage contribution shares for multiple work tasks but no overall estimate were also not considered. Contributions to unpublished conference talks, news items, and working papers (unless also published in a journal) were discarded. Publications that were not published at the time of thesis publications were searched and verified and only retained if they were eventually published and the final versions had the same authors and authorship order as in the contribution statement. Percentage contribution statements to single-authored publications were not collected, as these in all cases were given as 100%, if they were mentioned at all.

While this sample is by no means representative, it does have the advantage that the contribution statements are made publicly and are checked and approved by the supervisors and coauthors, which is required by the graduation regulations, and should therefore be relatively accurate and reliable. It should be stressed that all used credit figures are documented in the final published theses. Nevertheless, the data cannot be claimed to be objective, as the figures may be shaped by team-internal negotiation processes. As the sample is restricted to a specific

---

<sup>1</sup> The German terms mean “own share” (twice), “own contribution,” and “declaration of contribution.”

**Table 2.** Number of observations by author count and author position

Author count	Author position													Row percentage
	1	2	3	4	5	6	7	8	9	10	11	12	17	
2	51	45	0	0	0	0	0	0	0	0	0	0	0	12.3
3	77	34	31	0	0	0	0	0	0	0	0	0	0	18.3
4	62	27	27	23	0	0	0	0	0	0	0	0	0	17.9
5	46	19	16	15	15	0	0	0	0	0	0	0	0	14.3
6	54	11	10	11	8	8	0	0	0	0	0	0	0	13.1
7	28	10	8	6	5	4	5	0	0	0	0	0	0	8.5
8	19	4	3	1	2	1	0	1	0	0	0	0	0	4.0
9	14	1	2	2	2	1	1	1	1	0	0	0	0	3.2
10	8	2	0	0	0	1	1	0	0	0	0	0	0	1.5
11	7	1	0	0	0	2	0	0	0	0	0	0	0	1.3
12	2	2	2	1	1	2	1	1	1	1	1	1	0	2.1
13	1	2	2	1	0	0	0	0	0	1	0	0	0	0.9
14	1	1	0	0	0	0	0	0	0	0	0	0	0	0.3
15	4	0	1	0	1	0	0	0	0	0	0	0	0	0.8
16	2	0	0	0	0	0	0	0	0	0	0	0	0	0.3
17	1	1	0	2	0	0	0	0	0	0	0	0	1	0.6
18	2	0	0	0	0	0	0	0	0	0	0	0	0	0.3
21	0	0	0	1	0	0	0	0	0	0	0	0	0	0.1
28	0	1	0	0	0	0	0	0	0	0	0	0	0	0.1
48	0	0	1	0	0	0	0	0	0	0	0	0	0	0.1
74	0	1	0	0	0	0	0	0	0	0	0	0	0	0.1
Column percentage	48.7	20.8	13.2	8.1	4.4	2.4	1.0	0.4	0.3	0.3	0.1	0.1	0.1	

Downloaded from [http://direct.mit.edu/qss/article-pdf/1/2/551/1889824/qss\\_a\\_00048.pdf](http://direct.mit.edu/qss/article-pdf/1/2/551/1889824/qss_a_00048.pdf) by guest on 27 May 2024

author group, namely PhD researchers, working and publishing under their specific conditions, these conditions inevitably influence the data. In particular, PhD researchers are commonly expected to provide major contributions to coauthored publications in order for them to count towards their cumulative theses, as a demonstration of independent research ability is a regular requirement for graduation. For this reason, the authorship positions in a sample of PhDs' publications will not be distributed randomly but be concentrated toward the position indicating most contribution relatively more often; that is, first author position. Nevertheless, for the purpose of validating generally applicable authorship credit models, the specific nature of the sample is of little consequence, as the methods should be valid regardless of the specifics of the publications and authors. In other words, the nature of the data does not cause any inherent systematic distortions that would be biased towards any of the validated methods.



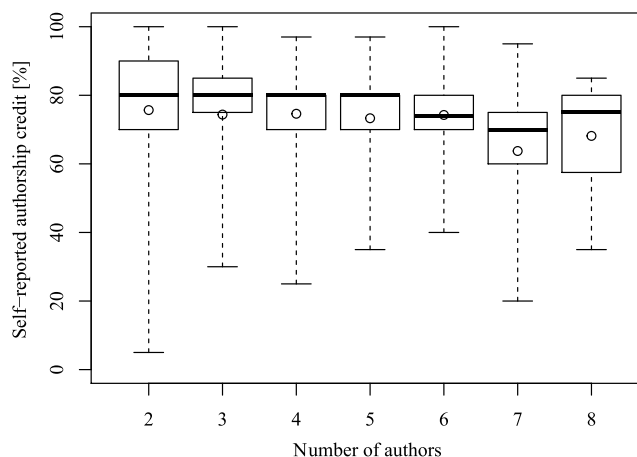
**Table 3.** Number of dissertations and publications by subjects

Subject	Dissertations	Publications
Biology	25	104
Medicine	19	71
Chemistry	19	68
Engineering	16	62
Economics and business	9	23
Psychology	7	23
Pharmaceutics	4	27
Nursing	4	18
Sport science	4	18
Nutrition	3	7
Geology	3	6
Astronomy	2	5
Agricultural science	2	4
Environmental science	2	2
Mathematics	1	4
Political science	1	3
Arts	1	1
Educational science	1	1
Sociology	1	1
Tourism studies	1	1
Veterinary medicine	1	1

As there seems to be considerable disagreement about the relative contributions of corresponding authors who are not first authors (Du & Tang, 2013) and heterogeneity in the meaning of the corresponding authorship signal between disciplines, we do not study credit allocation schemes modified to accommodate for corresponding authorship. Because statements about equal contributions or shared first authorship are not readily available in machine-readable form at the moment and hence impractical to use at scale, we do not take modifications for shared first authorships of credit allocation schemes into account either. The data set is published at <https://doi.org/10.5281/zenodo.3755227>.

### 3. RESULTS

One hundred twenty-five PhD theses completed between 2005 and 2019 at 22 different universities, fulfilling the above stated criteria, were found. Of these dissertations, 53 originated in Germany, 52 in Australia, and 20 in New Zealand. These included 465 combinations of thesis



**Figure 1.** Boxplot of distributions of claimed credit for first authors of authors with two to eight authors. Horizontal bar indicates the group median, circle the group mean.  $N$  (from two to eight authors): 51, 77, 62, 46, 54, 28, 19.

and article, from 382 unique articles. Considering all stated contribution figures (i.e., combinations of article and author), there were 778 data points, which were spread across author counts and author positions as shown in Table 2. In this table, the rows refer to the number of coauthors of papers while the columns refer to the specific positions in the author byline that an author can occupy. For example, the cell of author count 4 and author position 2 refers to data about the second listed author in papers with four authors, and there are 27 observations of this combination in the data. It can be seen that the observed cases are concentrated on the first author position across most author counts. Overall, almost half of all observations are of first author positions, confirming the expected skew due to the nature of the data. As mentioned above, PhD researchers are expected to make major contributions and major contributions are more likely to result in first authorship. We classified theses by subject, based on the title and degree-conferring department. The distribution of theses and articles across subjects is shown in Table 3. Most theses and publications are from the natural sciences, engineering, psychology, and the health sciences, whereas the social sciences, mathematics, and arts and humanities are hardly represented.

In Figure 1 we show that there is considerable variability in the claimed credit for the same position/author count combinations. Displayed are distribution summary boxplots for the credit of first authors of papers by two to eight authors, as there are reasonable numbers of observations only for these cases. For example, while the average contribution of first authors

**Table 4.** Lack-of-fit scores of authors' credit attributions and predicted values of counting methods ( $N = 778$  and  $777$ )

	Lack-of-fit index	Lack-of-fit, one outlier removed
Harmonic	0.174	0.168
First	1.926	1.917
Fractional	0.852	0.852
Geometric	15.503	0.323
Proportional	0.381	0.365

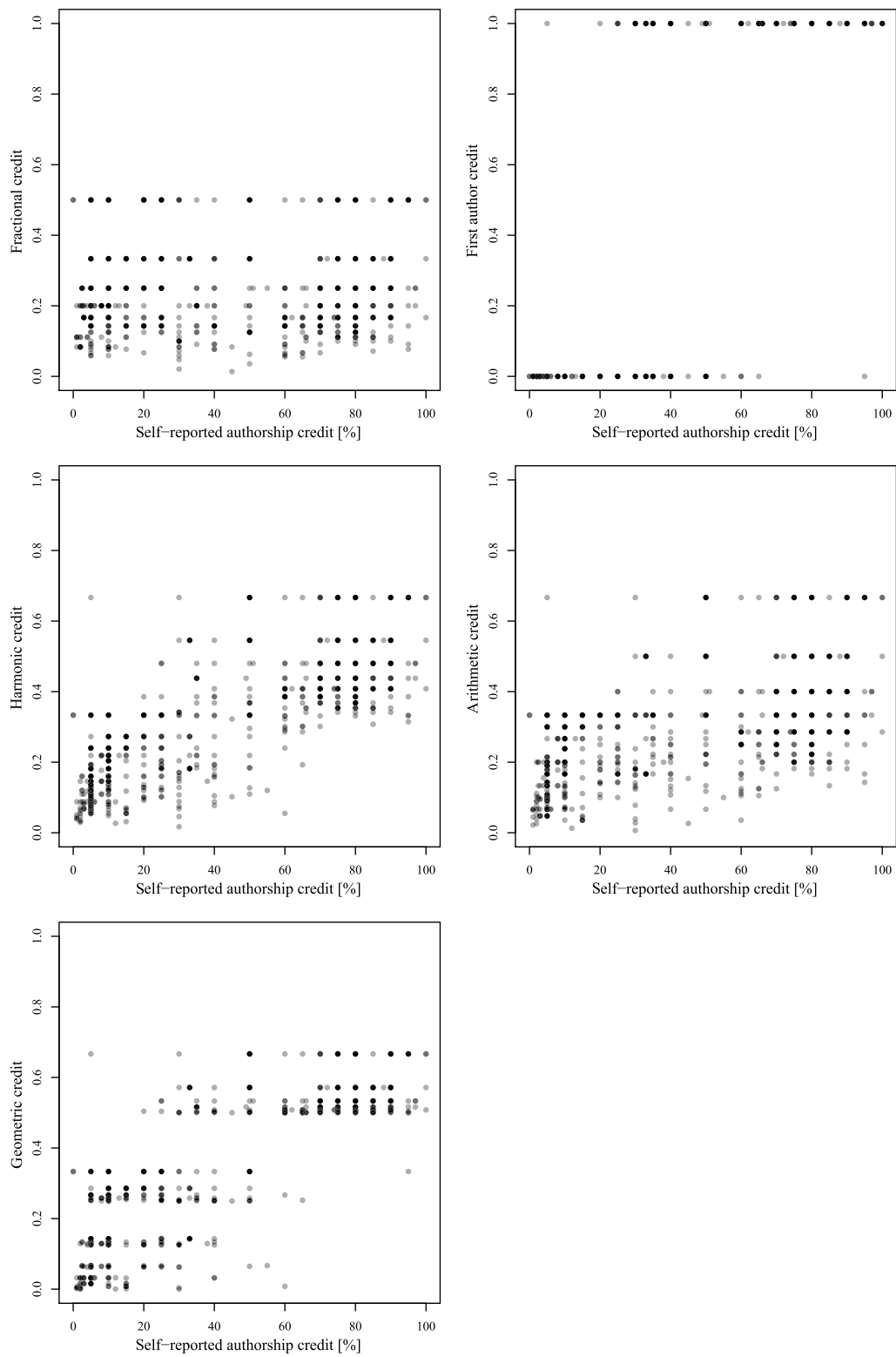


Figure 2. Scatterplots of predictions of authorship credit allocation methods against authors' claimed contributions.  $N = 778$ .

for two-author papers is 76%, the observed values range from 30% to 100%, with a standard deviation of 20%. As can be verified in the accompanying data set, there are statements claiming 100% authorship credit for first author positions of papers with multiple authors.

We follow the earlier literature by evaluating the validity of the studied credit allocation methods for our data using a lack-of-fit index. This measure is a sum of squared deviations of predictions from data, scaled by the size of the data set according to the formula

$$\text{Lack-of-fit} = \frac{1}{n-1} \sum \frac{(O-E)^2}{E}$$

where  $n$  is the total number of observations,  $O$  the observed value, and  $E$  the expected value (that is, the value predicted by the model). The values of the lack-of-fit measure, following Hagen (2010), are presented in Table 4. For this calculation, percentage values of the reported credit were divided by 100 in order to make the values comparable to the cited prior literature. For the first author count method, we changed the credit of nonfirst authors with a value of 0.0% to 0.1% to avoid division by zero. It is important to note that in contrast to the prior literature, such as Hagen (2010) and Kim and Kim (2015), we did not use the average values of contribution per combination of author count and author position. Instead we used the full data on the level of individual observations. When using the whole sample (second column), the lack-of-fit of the geometric counting method is extremely influenced by a single observation. In this observation, the claimed credit is 30% for the 17th author (who is not a corresponding author) of a paper with 17 authors. Geometric counting assigns this position only 0.0008% credit, while harmonic counting, for example, assigns 1.7%. For this reason, the third column in Table 4 shows the lack-of-fit values for all methods without this particular observation. The results are in line with the prior validation studies of Hagen (2010) and Kim and Kim (2015), the latter of which uses a similar but extended data set to the former and eliminates some inconsistencies. Harmonic credit most closely approximates the empirical data, followed by arithmetic and geometric credit. The data for the comparisons is visualized in Figure 2.

#### 4. DISCUSSION

Collaborative research published in multiauthored works is pervasive in science, as is evaluation based on published outputs. The choice of coauthorship counting method is decisive for the results of bibliometric research and evaluation studies (Gauffriau & Olesen Larsen, 2005). Therefore, in this study we have validated various proposed authorship credit allocation methods for multiauthored scientific publications by their model fit to empirical data of authorship credit statements of PhD graduates from three countries in cumulative dissertation theses. It was found that the harmonic credit method shows the highest agreement with the data with a lack-of-fit index of 0.174. Arithmetic and geometric credit performed slightly worse (lack-of-fit: 0.381; 15.503 or with outlier removed 0.323) while fractional credit and first author-only credit are clearly inferior methods (0.852 and 1.926, respectively). The results are in agreement with those of Hagen (2010) and strengthen the case for replacing fractional counting with harmonic counting in scientometric research and evaluation. However, the results presented here should be interpreted with due caution, as the data set is not representative. Most of the sampled authorship statements are from the natural sciences, the health sciences, and engineering and the data set is skewed towards first authorship position.

##### 4.1. A Research Agenda for Quantitative Author Contribution Studies

Despite being a topic of discussion since at least the 1980s, author credit attribution has made little progress. The reason is a lack of empirical studies that provide solid quantitative

evidence. Future studies should complement the results presented here by surveying all authors of coauthored publications independently and asking them to assign percentage ranges of approximate authorship credit to all authors. Ranges, and not scalars, are necessary in order to capture the uncertainty inherent in such judgments. Such studies should also take into account the different meaning of corresponding and last authorship across disciplines.

#### ACKNOWLEDGMENTS

This research was funded by German Federal Ministry of Education and Research grants 01PQ16004 and 01PQ17001.

#### AUTHOR CONTRIBUTIONS

PD: Conceptualization, Data curation, Formal Analysis, Funding acquisition, Investigation, Methodology, Software, Visualization, Writing – original draft, Writing – review & editing.

#### COMPETING INTERESTS

The author has no competing interests.

#### DATA AVAILABILITY

The complete data set is published at <https://doi.org/10.5281/zenodo.3755226>.

#### REFERENCES

- Bhandari, M., Einhorn, T. A., Swiontkowski, M. F., & Heckman, J. D. (2003). Who did what?: (Mis)perceptions about authors' contributions to scientific articles based on order of authorship. *Journal of Bone and Joint Surgery*, 85(8), 1605–1609.
- Butler, L., & Ginn, D. (1998). Canadian nurses' views on assignment of publication credit for scholarly and scientific work. *Canadian Journal of Nursing Research Archive*, 30(1), 171–183.
- Cole, J. R., & Cole, S. (1974). *Social stratification in science*. Chicago, London: University of Chicago Press.
- de Solla Price, D. (1981). Multiple authorship. *Science*, 212(4498), 986–986. <https://doi.org/10.1126/science.212.4498.986-a>
- Du, J., & Tang, X. (2013). Perceptions of author order versus contribution among researchers with different professional ranks and the potential of harmonic counts for encouraging ethical co-authorship practices. *Scientometrics*, 96(1), 277–295. <https://doi.org/10.1007/s11192-012-0905-4>
- Egghe, L., Rousseau, R., & van Hooydonk, G. (2000). Methods for accrediting publications to authors or countries: Consequences for evaluation studies. *Journal of the American Society for Information Science*, 51(2), 145–157. [https://doi.org/10.1002/\(SICI\)1097-4571\(2000\)51:2%3C145::AID-ASI6%3E3.0.CO;2-9](https://doi.org/10.1002/(SICI)1097-4571(2000)51:2%3C145::AID-ASI6%3E3.0.CO;2-9)
- Gauffriau, M., Larsen, P., Maye, I., Roulin-Perriard, A., & von Ins, M. (2008). Comparisons of results of publication counting using different methods. *Scientometrics*, 77(1), 147–176. <https://doi.org/10.1007/s11192-007-1934-2>
- Gauffriau, M., & Olesen Larsen, P. (2005). Counting methods are decisive for rankings based on publication and citation studies. *Scientometrics*, 64(1), 85–93. <https://doi.org/10.1007/s11192-005-0239-6>
- Hagen, N. T. (2008). Harmonic allocation of authorship credit: Source-level correction of bibliometric bias assures accurate publication and citation analysis. *PLOS ONE*, 3(12), e4021. <https://doi.org/10.1371/journal.pone.0004021>
- Hagen, N. T. (2010). Harmonic publication and citation counting: Sharing authorship credit equitably—not equally, geometrically or arithmetically. *Scientometrics*, 84(3), 785–793. <https://doi.org/10.1007/s11192-009-0129-4>
- Hart, R. L. (2000). Co-authorship in the academic library literature: A survey of attitudes and behaviors. *The Journal of Academic Librarianship*, 26(5), 339–345.
- Hodge, S. E., & Greenberg, D. A. (1981). Publication credit. *Science*, 213, 950.
- Huang, M.-H., Lin, C.-S., & Chen, D.-Z. (2011). Counting methods, country rank changes, and counting inflation in the assessment of national research productivity and impact. *Journal of the American Society for Information Science and Technology*, 62(12), 2427–2436. <https://doi.org/10.1002/asi.21625>
- Kalyane, V., & Vidyasagar Rao, K. (1995). Quantification of credit for authorship. *ILA Bulletin*, 30(3–4), 94–96.
- Kim, J., & Kim, J. (2015). Rethinking the comparison of coauthorship credit allocation schemes. *Journal of Informetrics*, 9(3), 667–673. <https://doi.org/10.1016/j.joi.2015.07.005>
- Laudel, G. (2002). What do we measure by co-authorships? *Research Evaluation*, 11(1), 3–15.
- Lin, C.-S., Huang, M.-H., & Chen, D.-Z. (2013). The influences of counting methods on university rankings based on paper count and citation count. *Journal of Informetrics*, 7(3), 611–621. <https://doi.org/10.1016/j.joi.2013.03.007>
- Lindsey, D. (1980). Production and citation measures in the sociology of science: The problem of multiple authorship. *Social Studies of Science*, 10(2), 145–162.
- Maciejovsky, B., Budescu, D. V., & Ariely, D. (2009). The researcher as a consumer of scientific publications: How do name-ordering conventions affect inferences about contribution credits? *Marketing Science*, 28(3), 589–598. <https://doi.org/10.1287/mksc.1080.0406>

- Moore, M. T., & Griffin, B. W. (2006). Identification of factors that influence authorship name placement and decisions to collaborate in peer-reviewed, education-related publications. *Studies in Educational Evaluation, 32*(2), 125–135. <https://doi.org/10.1016/j.stueduc.2006.04.004>
- Mowatt, G., Shirran, L., Grimshaw, J. M., Rennie, D., Flanagan, A., Yank, V., ... Bero, L. A. (2002). Prevalence of honorary and ghost authorship in Cochrane reviews. *Journal of the American Medical Association, 287*(21), 2769–2771. <https://doi.org/10.1001/jama.287.21.2769>
- Sauermann, H., & Haeussler, C. (2017). Authorship and contribution disclosures. *Science Advances, 3*(11), e1700404.
- van Hooydonk, G. (1997). Fractional counting of multiauthored publications: Consequences for the impact of authors. *Journal of the American Society for Information Science, 48*(10), 944–945. [https://doi.org/10.1002/\(Sici\)1097-4571\(199710\)48:10%3C944::Aid-Asi8%3E3.0.Co;2-1](https://doi.org/10.1002/(Sici)1097-4571(199710)48:10%3C944::Aid-Asi8%3E3.0.Co;2-1)
- Vinkler, P. (1993). Research contribution, authorship and team cooperativeness. *Scientometrics, 26*(1), 213–230. <https://doi.org/10.1007/BF0201680>
- Vinkler, P. (2000). Evaluation of the publication activity of research teams by means of scientometric indicators. *Current Science, 79*(5), 602–612.
- von Glinow, M. A., & Novelli Jr., L. (1982). Ethical standards within organizational behavior. *Academy of Management Journal, 25*(2), 417–436.
- Waltman, L. (2012). An empirical analysis of the use of alphabetical authorship in scientific publishing. *Journal of Informetrics, 6*(4), 700–711. <https://doi.org/10.1016/j.joi.2012.07.008>
- Waltman, L. (2016). A review of the literature on citation impact indicators. *Journal of Informetrics, 10*(2), 365–391. <https://doi.org/10.1016/j.joi.2016.02.007>
- Wren, J. D., Kozak, K. Z., Johnson, K. R., Deakynne, S. J., Schilling, L. M., & Dellavalle, R. P. (2007). The write position. *EMBO Reports, 8*(11), 988–991. <https://doi.org/10.1038/sj.embor.7401095>