Co-authorship Credit Allocation Methods in the Assessment of Citation Impact of Chemistry Faculty*

Jongwook Lee (이종욱)**
Kiduk Yang (양기덕)***

Contents

1. Introduction
2. Prior Research and Methods
   2.1 Authorship Credit Allocation Models
   2.2 Citation Indicators
   2.3 Study Data
3. Results
   3.1 Citation Scores by Models
   3.2 Citation Rankings by Models
4. Discussion and Conclusions

ABSTRACT

This study examined changes in citation index scores and rankings of thirty-five chemistry faculty members at Seoul National University using different co-authorship credit allocation models. Using 1,436 Web of Science papers published between 2007 and 2013, we applied the inflated, fractional, harmonic, network-based allocation, and harmonic+ models to calculate faculty's *h*-, *R*-, and normalization of *h*- and *R*- index scores and rankings. The harmonic+ model, which is based on our belief that contribution of primary authors should be the same regardless of collaboration, is designed to minimize the penalty for research collaboration imposed by harmonic and NBA models by boosting the contribution of collaborating primary authors to be on the equal footing with single authors. Although citation rankings by different models are correlated with each other within the same type of citation indicator, rankings of many faculty members changed across models, suggesting the importance of an accurate and relevant authorship credit allocation model in the citation assessment of researchers. The study also found that authorship patterns in conjunction with citation counts are important factors for robust authorship models such as harmonic and NBA, and harmonic+ model may be beneficial for collaborating primary authors. Future research that reexamines the models with updated empirical data would provide further insights into the robustness of the models.

Keywords: Authorship Credit Allocation, Harmonic Model, Network-Based Allocation Model, Chemistry, Citation Assessment

* This research was supported by Kyungpook National University Research Fund, 2013.
** Doctoral Candidate, School of Information, Florida State University (nadoopro@gmail.com) (제1저자)
*** Associate Professor, Department of Library and Information Science, Kyungpook National University (kiyang@knu.ac.kr) (교신저자)
논문접수일자: 2015년 7월 16일 최초심사일자: 2015년 7월 21일 게재확정일자: 2015년 8월 7일
한국문헌정보학회지, 49(3): 273-289, 2015. 〔http://dx.doi.org/10.4275/KSLIS.2015.49.3.273〕

1. Introduction

An increasing number of multi-authored research papers in scientific communication (Cronin 2001; Price 1981; Regaldo 1995) suggest the need for allocating credit to co-authors in an accurate manner for effective bibliometric research evaluation. Many scholars have proposed different methods for co-authorship credit allocation (Hagen 2013; Kim and Diesner 2014; Tol 2011). Of many different models, Hagen (2010; 2013; 2014) argued simplicity, accuracy, and flexibility of the harmonic counting model, demonstrating its better performance in comparison with some other models, such as inflated counting, fractional counting, Liu and Fang's model (2012a; 2012b), Lukovits and Vinkler's model (1995), and Trueba and Guerrero's model (2004). Recently, Kim and Diesner (2014) proposed a network-based co-authorship credit allocation (NBA) model and showed the robustness of their model over other models including the harmonic counting model.

There has been very little research that applied the harmonic and NBA models together and compared their performances in citation evaluation. In addition, although some prior studies showed the effects of different credit allocation models on publication and citation scores (Hagen 2008; 2014; Jian and Xiaoli 2013), relatively few researchers applied such models to the author-level research assessment in the field of chemistry, where research collaboration is frequent (Cronin, Shaw and La Barrer 2004) and the authorship order in the bylines of papers typically reflects the relative contributions of authors (Kim and Diesner 2014; Vinkler 1993; 2000). Furthermore, existing models often reduce the credit assigned to the first author as the number of co-authors increases despite his/her significant contribution in the research process.

To address this dearth of research, we applied major co-authorship credit allocation models to the calculation of citation index scores and rankings of 35 chemistry faculty members in one of the most prestigious universities in South Korea and observed the effects of the models on the citation indicators. We further compared the performance of the harmonic and NBA models and tested our modified model that gives full ('1') credit to the first author regardless of the number of co-authors.

2. Prior Research and Methods

2.1 Authorship Credit Allocation Models

We calculated citation scores of a paper by multiplying citation count by the author credit after applying different co-authorship allocation credit models. The allocation models used in our study are: (a) inflated counting, (b) fractional counting, (c) harmonic counting (Hagen 2010) and (d) network-based allocation (NBA) (Kim and Diesner 2014). The inflated and fractional counting are two common methods used in research assessment (Hagen 2008). In the inflated counting model, full credit ("1") is allocated to all coauthors, and in the fractional counting method, one credit is equally divided to coauthors, which generate *inflationary* and *equalizing* biases respectively (Hagen 2008; 2014). To address the inflationary and equalizing biases, Hagen has advocated the use of harmonic counting model that allocates authorship credit by the order and number of authors (Hodge and Greeberg 1981; Hagen 2008; 2014). In harmonic counting, the ith author of a paper authored by N coauthors can be calculated as follows (Hagen 2008):

$$i^{th} \text{ author contribution} = \frac{\frac{1}{i}}{\left[1+\frac{1}{2}+\cdots+\frac{1}{N}\right]} \quad (1)$$

The harmonic approach, however, can suffer from a lack of flexibility in its application to different disciplines where there are different academic norms for collaboration and authorship (Frandsen and Nicholaisen 2010; Kim and Diesner 2014; Maciejovsky, Budescu and Ariely 2009). In response to this, Kim and Diesner (2014) proposed a network-based allocation (NBA) model that is adaptable to different disciplines through alternation of distribution factor values. The formulas for the NBA model are shown below (Kim and Diesner 2014, 591):

$$\begin{array}{ll} 1^{st} \text{ author:} & v + v_t \sum_{n=1}^{N-r} \frac{1}{(N-n)} \\ \text{Mid-author au(r):} & \text{au(i)} = (v - v_t) + v_t \sum_{n=1}^{N-i} \frac{1}{(N-n)} \quad (2) \\ \text{Last author:} & v - v_t \end{array}$$

where N = author count (i.e., number of coauthors)
r = author order
V = paper value
d = distribution factor
$v = \frac{V}{N}$ $\quad v_t = d \times v$

Both Hagen (2010) and Kim and Diesner (2014) demonstrated performance of their models by showing smaller lack of fit values against the empirical data (i.e., expert judgment) from chemistry than other models. However, the empirical data used in two studies were different; that is, while Kim and Diesner (2014) referenced Vinker's 1993 paper, Hagen (2010) referenced Vinker's 2000 paper (see Table 1).

〈Table 1〉 Authorship credit empirical data by Vinkler (1993; 2000)

Paper	Vinkler (1993)					Vinkler (2000)					
Number of authors	Authorship order					Authorship order					
	1	2	3	4	5	1	2	3	4	5	6
2	0.71	0.29				0.65	0.35				
3	0.61	0.26	0.13			0.55	0.25	0.20			
4	0.54	0.31	0.09	0.06		0.50	0.25	0.15	0.10		
5	0.34	0.24	0.17	0.14	0.11	0.40	0.25	0.15	0.10	0.10	
6						0.35	0.25	0.15	0.10	0.10	0.05

In our study, we recalculated lack of fit (LOF) values against the more recent empirical data to find the best distribution factor for the NBA model. While Kim and Diesner (2014) reported the minimum lack of fit values with a distribution factor of 0.51 for chemistry using Vinkler's 1993 empirical data, our study found that a distribution factor of 0.48 produces the minimum lack of fit values (0.00482) for Vinker's 2000 empirical data. The LOF formula is shown below (Hagen 2010; Kim and Diesner 2014).

$$\text{Lack of fit} = \frac{1}{(n-1)}\sum\frac{(E-C)^2}{C} \qquad (3)$$

n = total number of empirical observations
E = empirical author scores (Table 1)
C = NBA author scores (d = 0.01..0.99)

Table 2 shows the authorship credit distributions of the harmonic counting model and NBA model. We measured the coefficient of determination (R^2) to see how well two methods explain the variation in the empirical data (Hagen 2013). The harmonic model and the NBA model explained 97.9% (R^2=0.9788) and 95.1% (R^2=0.9510) of variation in the empirical data respectively, showing that the harmonic model performs slightly better than the NBA model when Vinkler's 2000 data is used.

〈Table 2〉 Authorship credit of the harmonic counting and the NBA models

Method	Harmonic						NBA (d=0.48)					
Num. of authors	Authorship order						Authorship order					
	1	2	3	4	5	6	1	2	3	4	5	6
2	0.67	0.33					0.74	0.26				
3	0.55	0.27	0.18				0.57	0.25	0.17			
4	0.48	0.24	0.16	0.12			0.47	0.23	0.17	0.13		
5	0.44	0.22	0.15	0.11	0.09		0.40	0.21	0.16	0.13	0.10	
6	0.41	0.20	0.14	0.10	0.08	0.07	0.35	0.19	0.15	0.12	0.10	0.09

In the harmonic and the NBA models, however, the credit assigned to the first author decreases significantly as the number of co-authors increases, that is, the first author can be penalized more than other co-authors for collaboration despite his/her critical role in conducting research and writing the paper. In reaction to this, several scholars suggested giving a whole credit to the first author (Tscharntke et al. 2007; Zhang 2009) even though it makes a total authorship credit greater than one. When the unit of analysis is the author, as is the case in our study, reducing the first author credit in multi-authored papers essentially penalizes the first author for collaboration, which we believe is an unfair practice. Thus, we adapted Zhang's approach (2009) to extend the harmonic (harmonic+) and NBA (NBA+) models. In the harmonic+ model, the first author always receives one credit regardless of the number of co-authors, and the credits of non-first authors sum up to one except the case where there are two authors (Table 3).

〈Table 3〉 Authorship credit of the harmonic+ and NBA+ counting

Method	Harmonic+						NBA+ (d=0.48)					
Num. of authors	Authorship order						Authorship order					
	1	2	3	4	5	6	1	2	3	4	5	6
2	1.00	0.70					1.48	0.52				
3	1.00	0.67	0.33				1.15	0.51	0.35			
4	1.00	0.50	0.33	0.17			0.94	0.46	0.34	0.26		
5	1.00	0.40	0.30	0.20	0.10		0.80	0.42	0.32	0.26	0.21	
6	1.00	0.33	0.27	0.20	0.13	0.07	0.70	0.38	0.30	0.25	0.21	0.17

In the NBA+ model, setting V to 2 to make the sum of authorship credits to 2 produces problematic first author credit; for this reason, we excluded the NBA+ model in the analysis. The harmonic+ model can be formalized as:

1st author: $au(1) = 1$

Others: $au(i) = \frac{2*(N-i+1)/(N+1)(N-2)}{\sum_{n=2}^{N} 2*(N-n+1)/(N+1)(N-2)}$ (N>2) (4)

$au(2) = 0.7$ (N=2)

where N = author count (i.e., number of coauthors)

2.2 Citation Indicators

The citation indicators used in the study are *h*-index, *R*-index, and normalization of (*h*, *R*), which were used by Jian and Xiaoli (2013) who calculated citation scores of researchers in medical and health institutes in China, applying the concept of harmonic counting. The *h*-index, proposed by Hirsch (2005), is defined as "the number of papers with citation number ≥ *h*" (p. 16569). To address the weakness of *h*-index, which ignores the exact number of citation counts of papers, Jin, Liang, Rousseau, and Egghe (2007) proposed the *R*-index that takes a square root of the total citation counts of *h* papers (*h*-core), calculated by following formula (p. 857):

$$R = \sqrt{\sum_{j=1}^{h} cit_j} \quad (5)$$

where h = number of papers in h-core

Jin et al. (2007) further suggested combining *h*-index and *R*-index, as *R*-index is sensitive to articles with high number of citations. Jian and Xiaoli (2013) supported this normalization method of *h* and *R*-index (hereafter, *hR*-index), as *h*-index reflects the "number of papers" while *R*-index shows the "impact of papers" in the *h*-core. The formula of *hR*-index is:

$$Norm\,(h, R)_i = \frac{h_i}{\sum h_i} + \frac{R_i}{\sum R_i} \quad (6)$$

2.3 Study Data

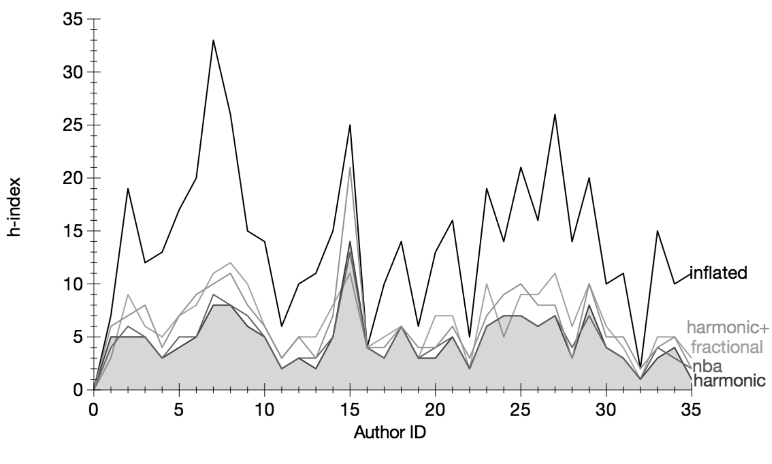
The data of this study covers faculty publications (2007-2013) in the chemistry department at one of leading universities in South Korea: Seoul National University (SNU). In the field of chemistry, SNU has been top-ranked in South Korea and ranked 16th by British Quacquarelli Symonds World University Ranking (QS Top Universities 2014). A recent bibliometric study reported that SNU has been the most productive university in chemistry in Korea (Magnone 2014), and Kim and Kim (2000) showed faculty research productivity of the chemistry department at SNU.

For the study, we collected the publication list of SNU chemistry faculty members from the department website (https://chem.snu.ac.kr), where 35 out of 36 faculty members' publication lists are available. The citation counts of the papers were gathered in February 2015 from Web of Science (WoS) databases (Web of Science Core Collection). Thirty-five SNU faculty members published 1,452 papers during seven years (2007-2013) of which 1,436 (98.9%) papers were indexed by WoS. After excluding two papers having more than 30 authors, we analyzed 1,434 WoS papers that have been cited 35,132 times in WoS databases.

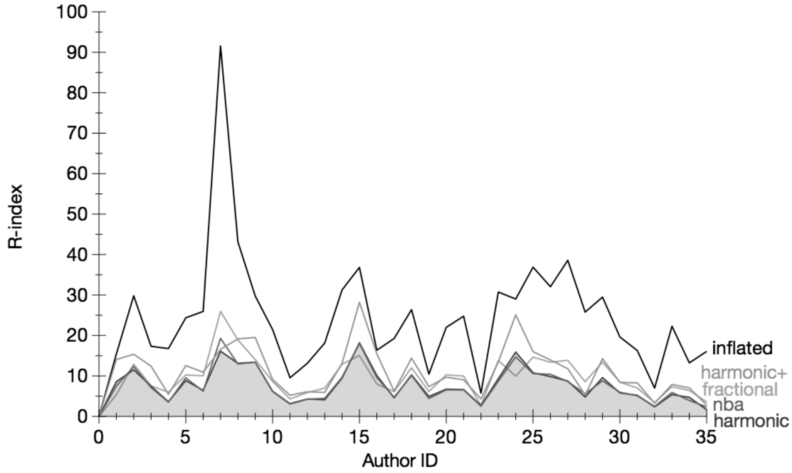
3. Results

3.1 Citation Scores by Models

We first compared the *h*-, *R*-, and *hR*- scores of 35 faculty members, calculated by four authorship credit allocation models (i.e., inflated, fractional, harmonic, and NBA) (see Appendix A). As expected, the inflated *h*- and *R*-scores were always greater than those by other models (Figure 1 and 2), except for one author (ID=16) who had a constant *h*-score across models resulting from a small number of papers authored as the first or the second author only. The author had published 4 papers (three as the first, one as the second author), all of which were cited enough times so as to nullify the effect of citation score reduction on *h*-index by different models.

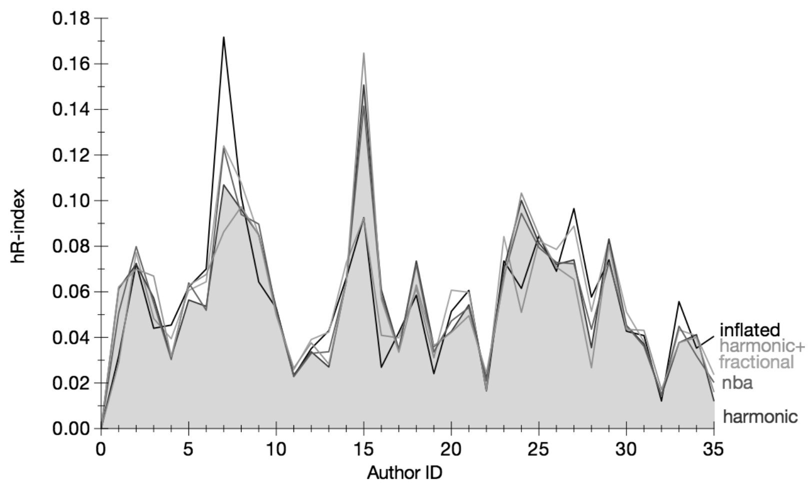


〈Figure 1〉 *h*-index scores of chemistry faculty



〈Figure 2〉 *R*-index scores of chemistry faculty

The inflated *hR*-scores were not considerably different from other *hR*-scores due to the normalization process (i.e., authors' *h*-/*R*-index scores are divided by the sum of all *h*-/*R*-index scores) (see Figure 3). Rather, some faculty members' inflated *hR*-scores were even smaller than their fractional, harmonic, or NBA *hR*-scores. In other words, some faculty members had greater inflated *h*- and *R*- scores than their harmonic or NBA *h*-/*R*- scores, but their inflated *hR*-scores were smaller than other *hR*-scores, which may be caused by a combination of factors as discussed below.



〈Figure 3〉 *hR*-index scores of chemistry faculty

The first possible factor for causing such outcome is a large proportion of the first- and second-authored papers. In the harmonic and NBA models, which assign credits to authors by their order on the byline, *h*-/*R*- scores of faculty members with many first- and second authored papers will be greater than those of faculty members with later authorship orders. This means the harmonic and NBA models' numerator to denominator ratio of the *hR*-score would tend to be greater than the inflated model for authors with high-order authorships. In our study, faculty members who authored more than 30% of papers as the first- or second- author (ID=1, 3, 9, 11, 15, 16, 19, 24), had the harmonic or NBA *hR*-scores greater than or equal to their inflated scores. Table 4 illustrates this effect of authorship order on *hR*-score computations.

〈Table 4〉 Effects of authorship order on *hR*-scores[1)]

author	average authorship order	*h*-index		*R*-index		*hR*-index	
		inflated	harmonic	inflated	harmonic	inflated	harmonic
A	1	10	5	20	10	0.67	1.00
B	3	10	3	20	6	0.67	0.60
C	5	10	2	20	4	0.67	0.40

The second factor that can influence the computation of *hR*-score is the citation count. Given that *hR*-index combines both number and impact of papers (Jian and Xiaoli 2013), the number of citations an author receives, especially the average citation count per paper after applying authorship credit adjustments, affects how *hR*-scores differ across the models. Among the faculty members with high average citation counts who authored less than 30% of papers as the first- or second- author, those with relatively less severe authorship credit adjustment (ID=2, 18, 26) show harmonic and NBA *hR*-scores greater than the inflated score (Table 5). The severity of authorship credit reduction by harmonic and NBA models is more pronounced for faculty members who tend to publish as auxiliary authors (ID=7, 14, 25, 27).

When it comes to the fractional model, the fractional *h*- and *R*- scores were typically greater than the harmonic or NBA *h*- and *R*- scores, except in a few cases where faculty members (e.g., ID=1, 15, 24) had a large proportion of first-/second- authored papers or a high citation rate. Similarly, if faculty members had the smaller fractional *hR*-scores than the harmonic or NBA

1) *h*-/*R*-index scores in Table 4 are estimated based on the average authorship order. For the high-order authorship (author A), numerator to denominator ratios of *hR*-score are 1 to 1 (5/10 + 10/20) for the harmonic model and 2 to 3 (10/30 + 20/60) for the inflated model.

hR-scores, they were more likely to have many first-/second-authored papers or have high numbers of papers or citations.

〈Table 5〉 Effects of citation counts on *hR*-scores

author ID	citation count	citation count per paper			*hR*-index		
		inflated	harmonic	NBA	inflated	harmonic	NBA
2	1122	27.4	5.5	6.0	0.072	0.072	0.079
18	267	27.9	5.0	5.0	0.059	0.074	0.072
26	932	33.5	4.3	4.7	0.069	0.072	0.073
7	8865	113.7	5.0	6.8	0.172	0.107	0.123
14	1023	46.5	5.8	6.3	0.066	0.065	0.064
25	1542	38.6	5.1	5.1	0.085	0.081	0.079
27	1893	30.5	3.0	3.2	0.097	0.074	0.072

The difference between the harmonic and NBA models were quite minor in the *h*-scores; for instance, 22 out of 35 faculty members' *h*-scores were the same using two models, while thirteen faculty members had one-score difference between them. In terms of *R*-scores, all faculty members have different scores depending on the allocation method (harmonic or NBA) used. While most faculty members have differences in the harmonic and NBA *R*-scores in the range of 1.00, a faculty member's (ID=7) *R*-score was different by 3.17 because of the authorship credit difference amplified by two papers that have been cited more than 2,000 times.

The harmonic+ *h*-scores were greater than the harmonic *h*-scores because of the increased author credits, especially to the high-order author. For the faculty member (ID=15) who authored majority of papers (73 out of 112 papers) as the first author, the advantage of using the harmonic+ model is quite obvious since his *h*-index increased from 14 (harmonic) to 21 (Table 6). Even with just a few first-authored papers, faculty members authoring highly cited papers as the second of two-author papers (ID=3, 6, 8, 25) also increase their *h*-scores by 3 or more with the harmonic+ model since the authorship credit allocation is boosted from 0.33 to 0.7 (Table 2 and 3). In regards to the *R*-scores, the increase by the harmonic+ model is predictable for those faculty members with increased *h*-scores (e.g., ID=3, 6, 8, 15, 25). Even a minor increase in *h*-score can result in a noticeable increase in *R*-score by relatively high citation counts of newly added *h*-core papers. For *hR*-scores, 17 faculty members got harmonic scores greater than harmonic+ scores because of the normalization process, calculating the relative scores of faculty members within an allocation method.

〈Table 6〉 Citation index scores of authors with well-cited high-order authored papers

author ID	citation count	*h*-index			*R*-index			*hR*-index		
		harnomic	NBA	harmonic+	harnomic	NBA	harmonic+	harnomic	NBA	harmonic+
3	404	5	5	8	7.46	7.12	12.40	0.0573	0.0551	0.0669
6	1027	5	5	9	6.44	6.23	10.92	0.0536	0.0519	0.0676
8	2470	8	8	11	13.12	12.95	19.16	0.0960	0.0939	0.0973
15	2224	14	13	21	18.20	17.98	28.20	0.1507	0.1414	0.1647
25	1542	7	7	10	10.78	10.56	15.99	0.0814	0.0794	0.0849

3.2 Citation Rankings by Models

We examined changes in citation rankings of faculty using different authorship allocation credit methods (see Appendix B). Kendall's rank correlation coefficient tau (τ) was computed to test associations, and there were significant correlations between all citation rankings. As reported in Table 7, the inflated citation rankings were more strongly correlated with the fractional rankings (.791 <= τ < = .818) than the harmonic (.616 <= τ <= .620) or NBA rankings (.654 <= τ <= .720). This pattern was also clear in the *hR*-index rankings even though it was difficult to compare *hR*-scores across different credit allocation models due to the normalization process. However, there seemed to be different patterns even between the inflated and the fractional rankings; some faculty members (ID=9, 20, 30) had lower inflated rankings than fractional rankings while others (ID=33, 35, 4) had lower fractional rankings than inflated rankings, due to the changes in fractional citation scores by the author count of papers.

The strong correlation between the harmonic and NBA models was expected, as both models allocate authorship credit by author contribution based on the author order. Despite very strong correlations between the harmonic and NBA rankings (.898 <= τ <= .946), two models generated somewhat different faculty citation rankings; that is, 86% of the faculty (30 authors) had different *h*-index rankings between the two models; 46% (17 authors) differed in *R*-index rankings; and 60% (21 authors) in *hR*-index rankings. In a similar vein, there were very strong correlations between the harmonic and harmonic+ rankings (.872 <= τ <= .923), although 71%, 77%, and 80% of the faculty (25, 27 and 28 authors) got ranking changes in *h*-, *R*-, and *hR*- index respectively. These findings suggest that, although similar authorship allocation models produce overall author rankings that are statistically correlated, many individual authors can experience ranking changes when using different authorship credit allocation models.

〈Table 7〉 Kendall's rank correlation coefficient

		c	*h*-index					*R*-index					*hR*-index				
			i	f	h	n	h+	i	f	h	n	h+	i	f	h	n	h+
c		-															
h	i	.854	-														
	f	.831	.791	-													
	h	.671	.616	.704	-												
	n	.787	.720	.792	.909	-											
	h+	.663	.640	.685	.872	.815	-										
R	i	.897	.784	.798	.684	.789	.616	-									
	f	.795	.738	.843	.728	.834	.670	.814	-								
	h	.588	.510	.608	.774	.821	.708	.620	.713	-							
	n	.621	.537	.622	.752	.825	.690	.654	.720	.946	-						
	h+	.544	.465	.566	.763	.775	.711	.563	.663	.923	.869	-					
hR	i	.918	.918	.844	.677	.782	.666	.869	.794	.570	.597	.526	-				
	f	.816	.760	.925	.720	.825	.655	.822	.922	.650	.664	.607	.818	-			
	h	.625	.561	.654	.885	.882	.788	.630	.730	.896	.862	.859	.620	.687	-		
	n	.691	.608	.691	.812	.906	.752	.700	.776	.898	.925	.838	.666	.727	.898	-	
	h+	.631	.571	.647	.849	.843	.887	.624	.703	.808	.782	.832	.620	.667	.872	.838	-

Note. Correlation is significant at the 0.01 level (2-tailed); c=citation count; i=inflated counting; f=fractional counting; h=harmonic counting; n=network-based model; h+=harmonic+ counting

4. Discussion and Conclusions

We applied different authorship credit allocation methods to 35 chemistry faculty members' citation data and observed some changes in their citation scores and rankings. The inflated and fractional counting models produced citation scores quite different from the scores by other models, while harmonic and NBA models generated quite similar citation scores across *h*-index, *R*-index, and *hR*-index. In terms of citation ranking, the inflated model produced more similar rankings to the fractional model than the harmonic and NBA models that eliminate inflationary and equalizing bias (Hagen 2014). However, the strong correlations between credit allocation models did not always guarantee minor ranking changes; rather, more than half of faculty members got different rankings even between the inflated and the fractional models and between the harmonic and the NBA models. Also, the majority of faculty members experienced ranking changes between the harmonic and harmonic+ counting methods. These findings suggest the importance of an accurate authorship credit allocation method in the citation assessment of researchers.

While Kim and Diesner (2014) showed a better performance of their NBA model in relation to Vinker's 1993 data, our study using Vinker's 2000 data suggests that the harmonic model explains 3% more variation in the empirical data set. The strength of the NBA model lies in its adaptability to different disciplines despite its relative complex formula. Our study findings based on the 35 faculty members may show that at least in chemistry, the harmonic model may be better for measuring citation scores and rankings of authors. Nevertheless, the empirical data that have been cited in co-authorship research were measured more than ten years ago, warranting future research that revisits chemists' perceptions of credit allocations.

We also proposed and tested the harmonic+ counting model, which is designed to minimize the penalty for research collaboration imposed by harmonic and NBA models by boosting the contribution of collaborating primary authors to be on the equal footing with single authors. By setting the sum of authorship credit to be 1 for multi-authored papers, harmonic and NBA models essentially penalize primary researchers for collaboration, thus introducing the "deflationary" bias for collaborating primary authors. Based on our belief that contribution of primary authors should be the same regardless of collaboration, harmonic+ model addresses the deflationary bias while still preventing inflationary and equalizing biases; however, further empirical data is necessary to test this model.

This study used seven-year bibliometric data of 35 faculty members who are affiliated with the most prestigious university in South Korea, implying that their publishing, authorship, and citation behaviors are closer to international academic norms but might be influenced by cultural factors. Although future work adding data sets from different countries could lead to more interesting findings, our study not only demonstrated how different authorship credit allocation strategies can result in different research assessment outcomes but also revealed that authorship (e.g., proportion of first- or second- authored papers), publishing (e.g., number of papers), and citation (e.g., highly cited papers) behaviors as well as features of citation indicators are important factors in determining the efficacy of authorship credit allocation models. Future research that reexamines the models with updated empirical data would provide further insights into the robustness of the models.

References

[1] Cronin, B. 2001. "Hyperauthorship: A Postmodern Perversion or Evidence of a Structural Shift in Scholarly Communication Practices?" *Journal of the American Society for Information Science and Technology,* 52(7): 558-569.

[2] Cronin, B., Shaw, D. and Barre, K. L. 2004. "Visible, Less Visible, and Invisible Work: Patterns of Collaboration in 20th Century Chemistry." *Journal of the American Society for Information Science and Technology,* 55(2): 160-168.

[3] Frandsen, T. F. and Nicholaisen, J. 2010. "What is in a name? Credit Assignment Practices in Different Disciplines." *Journal of Informetrics,* 4(4): 205-231.

[4] 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.

[5] Hagen, N. T. 2010. "Harmonic Publication and Citation Counting: Sharing Authorship Credit Equitably - not Equally, Geometrically or Arithmetically." *Scientometrics,* 84: 785-793.

[6] Hagen, N. T. 2013. "Harmonic Coauthor Credit: A Parsimonious Quantification of the Byline Hierarchy." *Journal of Informetrics,* 7: 784-791.

[7] Hagen, N. T. 2014. "Counting and Comparing Publication Output with and without Equalizing and Inflationary Bias." *Journal of Informetrics,* 8: 310-317.

[8] Hirsch, J. E. 2005. "An Index to Quantify an Individual's Scientific Research Output." *Proceedings of National Academy Sciences of the United States of America,* 102(46): 16569-16572.

[9] Hodge, S. E. and Greenberg, D. A. 1981. "Publication Credit." *Science,* 213(4511): 950.

[10] Jian, D. and Xiaoli, T. 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: 277-295.

[11] Jin, B. et al. 2007. "The R- and AR-indices: Complementing the h-index." *Chinese Science Bulletin,* 52(6): 855-863.

[12] Kim, Jinseok and Diesner, J. 2014. "A Network-based Approach to Coauthorship Credit Allocation." *Scientometrics,* 101: 587-602.

[13] Kim, Mee-Jean and Kim, Byong-Ju. 2000. "A Bibliometric Analysis of Publications by the Chemistry Department, Seoul National University, Korea." *Journal of Information Science,* 26(2): 111-119.

[14] Liu, X. Z. and Fang, H. 2012a. "Fairly Sharing the Credit of Multi-authored Papers and

its Application in the Modification of h-index and g-index." *Scientometrics,* 91(1): 37-49.

[15] Liu, X. Z. and Fang, H. 2012b. "Modifying h-index by Allocating Credit of Multi-authored Papers Whose Author Names Rank Based on Contribution." *Journal of Informetrics,* 6(4): 557-565.

[16] Lukovits, I. and Vinkler, P. 1995. "Correct Credit Distribution: A Model for Sharing Credit among Coauthors." *Social Indicators Research,* 36: 91-98.

[17] Magnone, E. 2014. "An Overview on the South Korean Scientific Production in the Field of Chemistry (1993-2012)." *Information,* 5(2): 285-304. doi:10.3390/info5020285

[18] Maciejovsky, B., Budescu, D. V. and 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.

[19] Price, D. S. and Derek, J. 1981. "Multiple Authorship." *Science,* 212(4498): 986.

[20] "QS World University Rankings: Methodology." 2014. *QS Top Universities*. October 19. [online] <http://www.topuniversities.com/university-rankings-articles/world-university-rankings/qs-world-university-rankings-methodology>

[21] Regaldo, A. 1995. "Multiauthor Papers on the Rise." *Science,* 268(5207): 25.

[22] Trueba, F. J. and Guerrero, H. 2004. "A robust formula to credit authors for their publications." *Scientometrics,* 60(2): 181-204.

[23] Tscharntke, T. et al. 2007. "Author Sequence and Credit for Contributions in Multiauthored Publications." *PLoS Biology,* 5(1): e18.

[24] Vinkler, P. 1993. "Research Contribution, Authorship and Team Cooperativeness." *Scientometrics,* 26(1): 213-230.

[25] Vinkler, P. 2000. "Evaluation of the Publication Activity of Research Teams by Means of Scientometric Indicators." *Current Science,* 79(5): 602-612.

[26] Zhang, C. T. 2009. "A Proposal for Calculating Weighted Citations Based on Author Rank." *EmboReports,* 10(5): 416-417.

[Appendix A] Citation scores of SNU chemistry faculty

		h-index					*R*-index					*hR*-index				
ID	c	i	f	h	n	h+	i	f	h	n	h+	i	f	h	n	h+
1	266	7	3	5	4	6	15.40	5.37	8.59	7.51	14.04	0.0318	0.0289	0.0614	0.0507	0.0621
2	1122	19	9	5	6	7	29.80	12.90	11.52	12.30	15.36	0.0724	0.0772	0.0721	0.0798	0.0699
3	404	12	6	5	5	8	17.32	7.53	7.46	7.12	12.40	0.0440	0.0483	0.0573	0.0551	0.0669
4	375	13	5	3	3	4	16.76	5.98	3.61	3.52	5.46	0.0454	0.0394	0.0312	0.0303	0.0316
5	847	17	7	4	5	7	24.37	10.23	8.86	9.52	12.55	0.0621	0.0606	0.0564	0.0639	0.0628
6	1027	20	8	5	5	9	25.94	10.03	6.44	6.23	10.92	0.0700	0.0644	0.0536	0.0519	0.0676
7	8865	33	11	8	9	10	91.55	25.95	16.11	19.28	16.56	0.1717	0.1240	0.1069	0.1228	0.0863
8	2470	26	12	8	8	11	43.15	18.87	13.12	12.95	19.16	0.1018	0.1078	0.0960	0.0939	0.0973
9	946	15	10	6	7	8	29.73	14.03	13.40	13.39	19.50	0.0643	0.0849	0.0850	0.0897	0.0848
10	791	14	6	5	5	6	21.45	8.64	6.22	6.12	9.11	0.0528	0.0516	0.0528	0.0515	0.0497
11	97	6	3	2	2	3	9.54	4.31	3.15	3.05	5.23	0.0230	0.0258	0.0235	0.0228	0.0265
12	292	10	5	3	3	5	13.12	5.87	4.29	4.24	6.08	0.0351	0.0391	0.0337	0.0330	0.0376
13	444	11	5	2	3	3	18.11	6.98	4.11	4.45	5.92	0.0429	0.0423	0.027	0.0337	0.0283
14	1023	15	8	5	5	7	31.26	12.90	9.47	9.68	13.11	0.0661	0.0728	0.0646	0.0644	0.0642
15	2224	25	11	14	13	21	36.82	15.05	18.20	17.98	28.20	0.0925	0.0923	0.1507	0.1414	0.1647
16	267	4	4	4	4	4	16.34	7.99	10.11	9.60	15.47	0.0269	0.0409	0.0609	0.0583	0.0568
17	494	10	5	3	3	4	19.26	6.11	4.57	4.74	6.23	0.0422	0.0398	0.0347	0.0348	0.0335
18	754	14	6	6	6	6	26.38	12.02	10.26	10.12	14.42	0.0585	0.0614	0.0735	0.0719	0.063
19	114	6	3	3	3	4	10.44	6.14	4.92	4.43	7.34	0.0241	0.0311	0.0360	0.0336	0.0363
20	557	13	7	3	4	4	22.00	10.26	6.73	6.55	9.68	0.0514	0.0607	0.0426	0.0472	0.0422
21	844	16	7	5	5	6	24.76	9.96	6.63	6.56	9.04	0.0606	0.0598	0.0543	0.0531	0.0495
22	51	5	2	2	2	3	5.75	2.74	2.68	2.52	4.30	0.0166	0.0168	0.0218	0.0209	0.0242
23	1093	19	10	6	6	7	30.73	13.81	9.31	8.72	13.48	0.0735	0.0842	0.0701	0.0668	0.0652
24	932	14	5	7	7	9	29.00	9.95	15.90	14.70	25.09	0.0615	0.0510	0.1001	0.0944	0.1033
25	1542	21	9	7	7	10	36.88	14.64	10.78	10.56	15.99	0.0846	0.0823	0.0814	0.0794	0.0849
26	1138	16	9	6	6	8	32.05	13.39	9.85	10.41	13.94	0.0690	0.0786	0.0720	0.0729	0.0708
27	1893	26	11	7	7	8	38.58	13.86	8.73	8.62	11.79	0.0965	0.0888	0.0740	0.0723	0.0654
28	819	14	6	3	4	3	25.77	8.58	4.79	5.57	5.27	0.0577	0.0514	0.0355	0.0436	0.0267
29	1314	20	10	8	7	10	29.48	13.40	9.58	8.79	14.26	0.0740	0.0831	0.0831	0.0729	0.0805
30	477	10	6	4	4	5	19.62	8.46	5.80	5.96	8.41	0.0427	0.0511	0.0452	0.0451	0.0435
31	350	11	4	3	3	5	16.28	7.03	5.21	5.10	8.28	0.0408	0.0381	0.0371	0.0361	0.0431
32	51	2	1	1	1	2	7.00	3.39	2.35	2.45	3.27	0.0121	0.0143	0.0146	0.0147	0.0171
33	652	15	5	3	4	4	22.27	7.33	5.39	5.92	7.88	0.0557	0.0433	0.0377	0.0449	0.0377
34	222	10	5	4	3	5	13.23	6.41	4.70	3.93	7.04	0.0353	0.0407	0.0412	0.0319	0.0400
35	375	11	3	1	2	2	16.06	3.56	1.65	2.34	2.83	0.0405	0.0236	0.0120	0.0202	0.0160

Note. c=citation count; i=inflated counting; f=fractional counting; h=harmonic counting; n=network-based model; h+=harmonic+ counting

[Appendix B] Citation rankings of SNU chemistry faculty members

		h-index					*R*-index					*hR*-index				
ID	c	i	f	h	n	h+	i	f	h	n	h+	i	f	h	n	h+
7	1	1	3	3	2	4	1	1	2	1	5	1	1	2	2	4
8	2	2.5	1	3	3	2	2	2	5	5	4	2	2	4	4	3
15	3	4	3	1	1	1	5	3	1	2	1	4	3	1	1	1
27	4	2.5	3	6	6	9.5	3	6	15	15	17	3	4	8	10	12
25	5	5	9	6	6	4	4	4	7	7	6	5	8	7	7	5
29	6	6.5	6	3	6	4	11	8	11	13	10	6	7	6	8.5	7
26	7	11.5	9	9.5	10.5	9.5	6	9	10	8	12	10	9	11	8.5	8
2	8	8.5	9	15	10.5	13.5	9	10.5	6	6	8	8	10	10	6	9
23	9	8.5	6	9.5	10.5	13.5	8	7	13	14	13	7	6	12	12	13
6	10	6.5	11.5	15	15.5	6.5	14	15	20	20	18	9	12	19	18	10
14	11	14	11.5	15	15.5	13.5	7	10.5	12	10	14	11	11	13	13	14
9	12	14	6	9.5	6	9.5	10	5	4	4	3	12	5	5	5	6
24	13	17.5	24	6	6	6.5	12	17	3	3	2	14	20	3	3	2
5	14	10	14	20.5	15.5	13.5	17	14	14	12	15	13	15	17	14	16
21	15	11.5	14	15	15.5	17.5	16	16	19	18	21	15	16	18	17	20
28	16	17.5	18	26.5	21.5	31.5	15	19	26	24	31	17	18	27	24	31
10	17	17.5	18	15	15.5	17.5	20	18	21	21	20	19	17	20	19	19
18	18	17.5	18	9.5	10.5	17.5	13	12	8	9	9	16	13	9	11	15
33	19	14	24	26.5	21.5	26.5	18	23	23	23	24	18	22	24	23	25
20	20	20.5	14	26.5	21.5	26.5	19	13	18	19	19	20	14	22	21	23
17	21	27.5	24	26.5	28	26.5	22	28	28	26	27	25	26	28	26	28
30	22	27.5	18	20.5	21.5	21.5	21	20	22	22	22	24	19	21	22	21
13	23	24	24	32	28	31.5	23	25	30	27	29	23	23	31	27	30
3	24	22	18	15	15.5	9.5	24	22	17	17	16	22	21	16	16	11
4	25.5	20.5	24	26.5	28	26.5	25	29	31	31	30	21	27	30	31	29
35	25.5	24	31.5	34.5	33	34.5	28	33	35	35	35	27	33	35	34	35
31	27	24	28.5	26.5	28	21.5	27	24	24	25	23	26	29	25	25	22
12	28	27.5	24	26.5	28	21.5	31	30	29	29	28	29	28	29	29	26
16	29	34	28.5	20.5	21.5	26.5	26	21	9	11	7	31	24	15	15	18
1	30	30	31.5	15	21.5	17.5	29	31	16	16	11	30	31	14	20	17
34	31	27.5	24	20.5	28	21.5	30	26	27	30	26	28	25	23	30	24
19	32	31.5	31.5	26.5	28	26.5	32	27	25	28	25	32	30	26	28	27
11	33	31.5	31.5	32	33	31.5	33	32	32	32	32	33	32	32	32	32
22	34.5	33	34	32	33	31.5	35	35	33	33	33	34	34	33	33	33
32	34.5	35	35	34.5	35	34.5	34	34	34	34	34	35	35	34	35	34

Note. c=citation count; i=inflated counting; f=fractional counting; h=harmonic counting; n=network-based model; h+=harmonic+ counting