Type of Submission: Presentation

An *h* index for Mendeley.

Comparison of citation-based *h* indices and a readership-based h_{men} index for 29 authors

Johanna M. Askeridis
Heinrich Heine University Düsseldorf, Germany
Department of Information Science
johanna.askeridis@hhu.de

INTRODUCTION

The internet and social networking services have changed the way people communicate. Research has not been unaffected, as scientific output is not excluded from the internet's everlasting discussion. Classic indicators like the *h* index (Hirsch, 2005) do not do justice to the newest ways of expressing our interest anymore. As a result, altmetrics emerged.

Aduku, Thelwall and Kousha (2017) state: "Currently, Mendeley readership statistics seem to be the most closely related to citation counts, in comparison to other altmetrics" (p. 575). And as many classic informetric methods base on citation counts, this paper proposes an indicator based on Mendeley readership counts as an additional indicator of author-based academic evaluations as the citation-based *h* indices alone do not cover the whole story of scientific activity and impact (Teixeira da Silva & Dobránszki, 2018a, 2018b). Since positive correlations between the two have been noted in previous research, the treatment of Mendeley readers like citations is proposed here in order to make the huge reader numbers comparable. Taking the *h* index's approach and molding it to fit the case of Mendeley, a new indicator which shall be called $\boldsymbol{h_{men}}$ **index**, as in *h* index for Mendeley, emerges: **An author has an index h_{men} if h_{men} of its N_p publications have at least h_{men} Mendeley readers each and the other ($N_p - h_{men}$) publications have ≤ h_{men} Mendeley readers each.**

This paper shall investigate the comparison for *h* and h_{men} values for 29 information scientists in order to find out about the reliability of the h_{men} index as an indicator to assess scholarly impact through data provided by the altmetric data source Mendeley. The *h* and h_{men} values shall also be explored in relation to an author's scientific age by making use of the time-oriented *h* index, also called *m* index.

METHODS

Since the extent of Mendeley readership counts varies among different disciplines (Mohammadi & Thelwall, 2014), the selected authors had to publish in the same discipline, here information science. With Mendeley having been launched in only 2008, older documents are significantly less covered according to Haustein, Peters, Bar-Ilan et al. (2014). Due to this age bias, authors were selected from two groups with regard to the degree of establishment of an author in the discipline in order to be able to check for correlations between *h* index and h_{men} index within different age groups of authors. The first is a group of established authors consisting of the nine top-ranking authors from Cronin and Meho's (2006) similar study on the *h* index and citation counts, and of nine members of the International Society for Scientometrics and Informetrics (ISSI) Scientific Committee, following the selection by Dorsch, Askeridis, and Stock (2018). The second group of eleven young authors consists of academic staff from the Department of Information Science at Heinrich Heine University, Germany.

As the documents covered and the number of citations change depending on the information service, Web of Science (WoS) and Scopus have both been used to generate two different *h* values. Each author list was manually checked for erroneously included documents. Mendeley readers were obtained via Webometric Analyst 2.0 (Thelwall, 2009). From these lists, only documents with a matching probability between author name query and document of 1.0 were considered. Each list was manually checked for erroneously included documents and duplicates were merged via a title check. All index values were manually derived. The data was collected from March 17 to March 20, 2018.

The scientific age of an author was determined by checking for the oldest publication on an author's personal publication list as taken from his or her institutional or personal website. If no such list was available, the date of the oldest publication available on WoS was taken into account instead.

Pearson correlations were calculated via the statistical software R.

RESULTS

Table 1 shows the obtained *h* and h_{men} values for all authors, as well as their scientific ages. The young author group has significantly smaller values for each index, which is to be expected due to their young scientific age. In the established author group, there is a greater variance in values due to varying levels of scientific age. Each author has a higher *h* value or at least the same on Scopus in comparison to WoS. In the young author group, all authors have higher h_{men} than *h* values. This is also the case for most established authors, though some have an h_{men} value as high as their Scopus *h* value and four of the authors have a lower h_{men} than *h* value. Here, the effect of the age bias on Mendeley as described by Haustein, Peters, Bar-Ilan et al. (2014) might be showing.

In table 2, the correlations calculated between the three index values for both author groups taken together are shown, as well as the values for the young and established authors respectively when taking into consideration only one author group at a time. The correlation values for all authors taken together are very strong, positive and statistically highly significant. The strongest correlation of 0.989 is and was to be expected between the Scopus and WoS *h* values since both are based on citation counts. Both Scopus and WoS show similarly significant positive correlations with the h_{men} index, with Scopus having the slightly stronger correlation of 0.954. For the young authors, the highest correlation was found between the Scopus *h* index and the h_{men} index, 0.869, significant at the 0.001 level, which is a stronger and more significant correlation than the one between Scopus and WoS *h* indices. The weakest and least significant correlation could be established between WoS *h* index and h_{men} index with only 0.686. The established author group shows only correlations significant at the 0.001 level. They behave similarly to the correlations for all authors taken together, the strongest being between Scopus and WoS *h* indices, followed by Scopus *h* and h_{men} index with 0.869 and the slightly weaker correlation between WoS *h* and h_{men} index. In all three cases, the correlations are stronger and of higher or the same significance as the results for the young authors.

The *h* index has been criticized for making comparisons between authors of different scientific ages impossible. In accordance with Stock and Stock (2013, p. 383) following Hirsch (2005), each author's index values were divided by his or her scientific age, leading to the results shown in table 3. The smallest value of 0 can be found for young authors only, since six out of eleven have not been cited in at least one of the databases yet. On Mendeley, however, six out of eleven have values greater than 1 which is quite strong. While the WoS values for young authors are still quite weak, the values for Mendeley can compete with those of the established authors. In the established author group, values greater than 2.0 have only been achieved by the youngest three authors: Lutz Bornmann with an age of 15 and Stefanie Haustein and Cassidy R. Sugimoto with an age of 11.

DISCUSSION

When all authors are taken together, correlations in each case were found to be very strong, positive and highly significant. The correlations with the h_{men} index for all authors were 0.954 for the Scopus *h* index and 0.948 for the WoS *h* index, significant at the 0.001 level respectively. Cronin and Meho (2006) reported a correlation of 0.9, significant at only the 0.01 level, for citation counts and *h* index and deemed the *h* index "a reliable indicator of scholarly impact and influence" (p. 1278). Following this line of interpretation, the h_{men} index is reliable for scholarly impact assessment as well.

The values for the time-oriented *h* and h_{men} indices show in general no big differences between indices, but they reveal exceptional authors who were found to rather have a young scientific age. Since the young authors are able to compete with the established authors on Mendeley, the h_{men} index shows different proportions than the *h* indices. In general, the results hint towards a faster reception of scientific work via Mendeley readers than via citations.

For the young authors group, the small size of the sample and the small values with little variety render the results unreliable to a certain degree, which is a clear limitation of this study. Furthermore, Mendeley readership lists might be inaccurate as all lists were in severe need of correction. Shrivastava and Mahajan (2016) argue that citation counts and Mendeley readers are different indicators and not similar in nature, for example because they show impact with regard to different user groups, but that should not be seen as an argument to discard the h_{men} index.

It might have limitations, but it also promises new insights. The h_{men} index seems to be similar to the *h* index and thus does indeed a good job of measuring impact of scholarly documents. And if it is indeed different in nature, yet shows similarly reliable results as the *h* index, it should be seen as a means to capture an even broader impact of research than the *h* index, while at the same time staying close to traditional means of impact assessment in comparison to most other altmetrics. The h_{men} index thus comes one step closer to broader impact measurement as intended by the emergence of altmetrics. Possibly, the h_{men} index might also predict future citation values since altmetrics, including Mendeley readers, are much faster than citations. It could also be possible that a young researcher's h_{men} value might hint towards their future career development.

REFERENCES

Aduku, J. K., Thelwall, M., & Kousha, K. (2017). Do Mendeley reader counts reflect the scholarly impact of conference papers? An investigation of computer science and engineering. *Scientometrics, 112*(1), 573-581.

Cronin, B., & Meho, L. (2006). Using the h-index to rank influential information scientists. *Journal of the American Society for Information Science and Technology, 57(9)*, 1275-1278.

Dorsch, I., Askeridis, J. M., & Stock, W. G. (2018). Truebounded, overbounded, or underbounded? Scientists' personal publication lists versus lists generated through bibliographic information services. *Publications, 6*(1), 1-9.

Haustein, S., Peters, I., Bar-Ilan, J., Priem, J., Shema, H., & Terliesner, J. (2014). Coverage and adoption of altmetrics sources in the bibliometric community. *Scientometrics, 101*(2), 1145–1163.

Hirsch, J. E. (2005). An index to quantify an individual's scientific research output. *Proceedings of the National Academy of Sciences of the United States of America, 102*(46), pp. 16569-165772.

Mohammadi, E., & Thelwall, M. (2014). Mendeley readership altmetrics for the social sciences and humanities: Research evaluation and knowledge flows. *Journal of the Association for Information Science and Technology, 65*(8), 1627–1638.

Shrivastava, R, & Mahajan, P. (2016). Relationship between citation counts and Mendeley readership metrics: A case of top 100 cited papers in Physics. *New Library World, 117*(3/4), 229-238.

Stock, W. G., & Stock, M. (2013). *Handbook of Information Science*. Berlin, Boston, MA: De Gruyter Saur.

Teixeira da Silva, J. A., & Dobránszki, J. (2018a). Multiple versions of the *h*-index: cautionary use for formal academic purposes. *Scientometrics, 115*(2), 1107-1113.

Teixeira da Silva, J. A., & Dobránszki, J. (2018b). Rejoinder to "Multiple versions of the *h*-index: cautionary use for formal academic purposes". *Scientometrics, 115*(2), 1131-1137.

Thelwall, M. (2009). *Introduction to Webometrics: Quantitative Web Research for the Social Sciences*. San Rafael, CA: Morgan & Claypool.

LIST OF TABLES

Table 1: *h* Values on Scopus and WoS, h_{men} Value and Scientific Age for Each Author

Author group	Author Name	Scopus *h* index	WoS *h* index	Mendeley h_{men} index	Scientific age
Young	Beutelspacher, Lisa	3	0	5	9
	Dorsch, Isabelle	1	1	2	4
	Fietkiewicz, Kaja J.	3	1	5	5
	Göretz, Julia	0	0	1	4
	Henkel, Maria	2	1	3	4
	Ilhan, Aylin	1	0	5	3
	Mainka, Agnes	5	2	13	8
	Meschede, Christine	1	0	4	3
	Scheibe, Katrin	0	0	2	2
	Siebenlist, Tobias	3	1	5	8
	Zimmer, Franziska	1	0	2	1
Established	Bar-Ilan, Judit	30	25	30	29
	Bates, Marcia J.	23	22	25	47[a]
	Belkin, Nicholas J.	32	22	30	44
	Borgman, Christine L.	27	25	32	43
	Börner, Katy	24	21	38	26
	Bornmann, Lutz	38	36	51	15[a]
	Cronin, Blaise	28	27	29	39[a]
	Egghe, Leo	27	27	22	39[a]
	Fidel, Raya	18	17	20	42
	Haustein, Stefanie	14	11	26	11
	Ingwersen, Peter	23	21	25	30
	Leydesdorff, Loet	59	54	71	48
	Marchionini, Gary	27	21	31	32[a]
	McCain, Katherine W.	22	22	20	38[a]
	Saracevic, Tefko	25	23	27	55
	Schlögl, Christian	11	10	15	26
	Spink, Amanda	41	32	39	28[a]
	Sugimoto, Cassidy R.	21	19	37	11

Note: [a] value derived from oldest WoS publication

Table 2: *h* and h_{men} Value Correlations for All, Young and Established Authors Respectively

All authors	Scopus *h* index	WoS *h* index	Mendeley h_{men} index
Scopus *h* index	-		
WoS *h* index	0.989***	-	
Mendeley h_{men} index	0.954***	0.948***	-
Young authors	Scopus *h* index	WoS *h* index	Mendeley h_{men} index
Scopus *h* index	-		
WoS *h* index	0.765**	-	
Mendeley h_{men} index	0.869***	0.686*	-
Established authors	Scopus *h* index	WoS *h* index	Mendeley h_{men} index
Scopus *h* index	-		
WoS *h* index	0.966***	-	
Mendeley h_{men} index	0.872***	0.862***	-

Note: *p < 0.05, **p <0 .01, ***p <0 .001

Table 3: Time-oriented *h* Values on Scopus and WoS and h_{men} Value for Each Author

Author group	Author name	Time-oriented Scopus *h* index	Time-oriented WoS *h* index	Time-oriented Mendeley h_{men} index
Young	Beutelspacher, Lisa	0.3	0.0	0.6
	Dorsch, Isabelle	0.3	0.3	0.5
	Fietkiewicz, Kaja J.	0.6	0.2	1.0
	Göretz, Julia	0.0	0.0	0.3
	Henkel, Maria	0.5	0.3	0.8
	Ilhan, Aylin	0.3	0.0	1.7
	Mainka, Agnes	0.6	0.3	1.6
	Meschede, Christine	0.3	0.0	1.3
	Scheibe, Katrin	0.0	0.0	1.0
	Siebenlist, Tobias	0.4	0.1	0.6
	Zimmer, Franziska	1.0	0.0	2.0
Established	Bar-Ilan, Judit	1.0	0.9	1.0
	Bates, Marcia J.	0.5	0.5	0.5
	Belkin, Nicholas J.	0.7	0.5	0.7
	Borgman, Christine L.	0.6	0.6	0.7
	Börner, Katy	0.9	0.8	1.5
	Bornmann, Lutz	2.5	2.4	3.4
	Cronin, Blaise	0.7	0.7	0.7
	Egghe, Leo	0.7	0.7	0.6
	Fidel, Raya	0.4	0.4	0.5
	Haustein, Stefanie	1.3	1.0	2.4
	Ingwersen, Peter	0.8	0.7	0.8
	Leydesdorff, Loet	1.2	1.1	1.5
	Marchionini, Gary	0.8	0.7	1.0
	McCain, Katherine W.	0.6	0.6	0.5
	Saracevic, Tefko	0.5	0.4	0.5
	Schlögl, Christian	0.4	0.4	0.6
	Spink, Amanda	1.5	1.1	1.4
	Sugimoto, Cassidy R.	1.9	1.7	3.4