

How to normalize Twitter counts? A first attempt based on journals in the Twitter Index

Lutz Bornmann¹ · Robin Haunschild²

Received: 20 November 2015 / Published online: 27 February 2016
© The Author(s) 2016. This article is published with open access at Springerlink.com

Abstract One possible way of measuring the broad impact of research (societal impact) quantitatively is the use of alternative metrics (altmetrics). An important source of altmetrics is Twitter, which is a popular microblogging service. In bibliometrics, it is standard to normalize citations for cross-field comparisons. This study deals with the normalization of Twitter counts (TC). The problem with Twitter data is that many papers receive zero tweets or only one tweet. In order to restrict the impact analysis on only those journals producing a considerable Twitter impact, we defined the Twitter Index (TI) containing journals with at least 80 % of the papers with at least 1 tweet each. For all papers in each TI journal, we calculated normalized Twitter percentiles (TP) which range from 0 (no impact) to 100 (highest impact). Thus, the highest impact accounts for the paper with the most tweets compared to the other papers in the journal. TP are proposed to be used for cross-field comparisons. We studied the field-independency of TP in comparison with TC. The results point out that the TP can validly be used particularly in biomedical and health sciences, life and earth sciences, mathematics and computer science, as well as physical sciences and engineering. In a first application of TP, we calculated percentiles for countries. The results show that Denmark, Finland, and Norway are the countries with the most tweeted papers (measured by TP).

Keywords Twitter counts · Twitter percentiles · Twitter Index · Altmetrics

✉ Lutz Bornmann
bornmann@gv.mpg.de

Robin Haunschild
R.Haunschild@fkf.mpg.de

¹ Division for Science and Innovation Studies, Administrative Headquarters of the Max Planck Society, Hofgartenstr. 8, 80539 Munich, Germany

² Max Planck Institute for Solid State Research, Heisenbergstr. 1, 70569 Stuttgart, Germany

Introduction

The success of the modern science system is closely related to a functioning research evaluation system by peers: without critical assessments by peers improvements of research approaches would be absent and standards could not be reached (Bornmann 2011). With the advent of large bibliometric databases (especially the citation indexes of Thomson Reuters) and the need for cross-disciplinary comparisons (e.g. of complete universities) bibliometrics has been more and more used to supplement (or sometimes to replace) peer review. Various national research evaluation systems have a strong focus on bibliometrics (Bornmann in press) and a manifesto has been published how bibliometrics can be properly used in research evaluation (Hicks et al. 2015). Citation analyses measure the impact of science on science. Since governments are interested today not only in this recursive kind of impact, but also in the broad impact of science on the wider society, scientometricians are searching for new metrics measuring broad impact reliably and validly. The use of case studies for demonstrating broad impact in the current UK Research Excellence Framework (REF) is a qualitative approach with the typical problems of missing generalizability, great amount of work, and case selection bias (only favorable cases of impact are reported) (Bornmann 2013).

One possible way of measuring broad impact quantitatively is the use of alternative metrics (altmetrics) (Bornmann 2014a; NISO Alternative Assessment Metrics Project 2014)—a new subset of scientometrics (Priem 2014). “Alternative metrics, sometimes shortened to just altmetrics, is an umbrella term covering new ways of approaching, measuring and providing evidence for impact” (Adie 2014, p. 349). An important source of altmetrics is Twitter: It is a popular microblogging platform with several million active users and messages (tweets) being sent each day. Tweets are short messages which cannot exceed 140 characters in length (Shema et al. 2014). Direct or indirect links from a tweet to a publication are defined as Twitter mentions (Priem and Costello 2010). Twitter mentions can be counted (Twitter counts, TC) in a similar way as traditional citations and the impact of different publications can be compared.

In bibliometrics, it is standard to normalize citations (Wilsdon et al. 2015). Citations depend on publication year and subject category. Thus, for cross-field and cross-time comparisons normalized citation scores are necessary and have been developed in recent decades (Vinkler 2010). Against the backdrop of the general practice of normalizing citations, many authors in the area of altmetrics argue for the necessity to field- and time-normalize altmetrics, too (Fenner 2014; Taylor 2013). Recently, Haunschild and Bornmann (2016) have proposed methods to normalize Mendeley counts—a popular altmetrics based on data from an online reference manager. In this paper, we propose corresponding methods for TC so that Twitter impact can be fairly measured across papers published in different subject categories and publication years. Since Twitter data has other properties than Mendeley data, methods developed for Mendeley cannot simply be transferred to Twitter and new methods for Twitter data are in need.

Research on Twitter

A free account on Twitter enables users to “follow” other Twitter users. This means one subscribes to their updates and can read their “tweets” (short messages) in a feed. Also, one can “retweet” these messages or tweet new short messages which are read by own

followers in their feeds. Up to a third of the tweets may be simple retweets (Holmberg 2014). “Tweets and retweets are the core of the Twitter platform that allows for the large-scale and rapid communication of ideas in a social network” (Darling et al. 2013). Whereas at the start of the platform Twitter was mostly used for personal communication, studies have uncovered its increasing use for work-related purposes (Priem and Costello 2010; Priem and Hemminger 2010).

It is possible to include references to publications in tweets: “We defined Twitter citations as direct or indirect links from a tweet to a peer-reviewed scholarly article online” (Priem and Costello 2010). Since tweets are restricted to 140 characters, it is frequently difficult to explore why a paper has been tweeted (Haustein et al. 2014a). In most of the cases tweets including a reference to a paper will have the purpose of bringing a new paper to the attention of the followers. Thus, tweets are not used (and cannot be used) to extensively discuss papers. According to Haustein et al. (2014a) “unlike Mendeley, Twitter is widely used outside of academia and thus seems to be a particularly promising source of evidence of public interest in science” (p. 208). TC do not correlate with citation counts (Bornmann 2015) and the results of Bornmann (2014b) show that particularly well written scientific papers (not only understandable by experts in a field) which provides a good overview of a topic generate tweets.

The results of Haustein et al. (2014b) point to field differences in tweeting: “Twitter coverage at the discipline level is highest in Professional Fields, where 17.0 % of PubMed documents were mentioned on Twitter at least once, followed by Psychology (14.9 %) and Health (12.8 %). When the data set is limited to only those articles that have been tweeted at least once, the papers from Biomedical Research have the highest Twitter citation rate ($T/P_{\text{tweeted}} = 3.3$). Of the 284,764 research articles and reviews assigned to this discipline, 27,878 were mentioned on Twitter a total of 90,633 times. Twitter coverage is lowest for Physics papers covered by PubMed (1.8 %), and Mathematics papers related to biomedical research receive the lowest average number of tweets per tweeted document ($T/P_{\text{tweeted}} = 1.5$)” (p. 662). According to Zahedi, Costas, and Wouters (2014) “in Twitter, 7 % of the publications from Multidisciplinary field, 3 % of the publications from Social & Behavioural Sciences and 2 % of publications from Medical & Life Sciences are the top three fields that have at least one tweet. In Delicious, only 1 % of the publications from Multidisciplinary field, Language, Information & Communication and Social & Behavioural Sciences have at least one bookmark while other fields have less than 1 % altmetrics” (p. 1498).

The results of both studies indicate that TC should be normalized with respect to field assignments.

Methods

Data

We obtained the Twitter statistics for articles and reviews published in 2012 and having a DOI ($n_A = 1,198,184$ papers) from Altmetric¹—a start-up providing article-level metrics—on May 11, 2015. The DOIs of the papers from 2012 were exported from the in-house database of the Max Planck Society (MPG) based on the Web of Science (WoS, Thomson Reuters) and administered by the Max Planck Digital Library (MPDL). We received altmetric data from Altmetric for 310,933 DOIs (26 %). Altmetric did not register

¹ <http://www.altmetric.com>.

altmetric activity for the remaining papers. For 37,692 DOIs (3 %), a Twitter count of 0 was registered. The DOIs with no altmetric activity registered by Altmetric were also treated as papers with 0 tweets. Furthermore, our in-house data base was updated in the meantime. 12,960 DOIs for papers published in 2012 were added (e.g. because new journals with back files were included in the WoS by Thomson Reuters). We treat these added papers as un-tweeted papers. Thus, a total of 937,843 papers (77 %) out of 1,211,144 papers were not tweeted.

Normalization of Twitter counts

In the following, we propose a possible procedure for normalizing TC which is percentile-based. The procedure focusses on journals (normalization on the journal level) and pools the journals with the most Twitter activities in the so called Twitter Index (TI).

Following percentile definitions of Leydesdorff and Bornmann (2011), ImpactStory—a provider of altmetrics for publications—provides Twitter percentiles (TP) which are normalized according to the publication year and scientific discipline of papers (Chamberlain 2013; Roemer and Borchardt 2013). The procedure of ImpactStory for calculating the percentile for a given paper i is as follows²: (1) The discipline is searched at Mendeley (a citation management tool and social network for academics) from which paper i is most frequently read. “Saves” at Mendeley are interpreted in altmetrics as “reads” and Mendeley readers share their discipline. (2) All papers which are assigned to the same discipline in Mendeley and are published in the same year (these papers constitute the reference set of paper i) are sorted in descending order according to their TC. (3) The proportion of papers is determined in the reference set which received less tweets than paper i . (4) The proportion equals the percentile for paper i .

It is a sign of professional scientometrics to use normalized indicators. Compared to other methods used for normalization in bibliometrics, percentile-based indicators are being seen as robust indicators (Hicks et al. 2015; Wilsdon et al. 2015). However, the procedure used by ImpactStory has some disadvantages (as already outlined on its website): (1) There might be instances where a paper’s actual discipline doesn’t match the disciplinary reference set used for the normalization. Papers might be read in disciplines to which they do not belong. (2) The discipline for a paper might change, if the most frequently read discipline changes from one year to another. (3) If a paper does not have any readers at Mendeley, all papers within one year constitute the reference set in ImpactStory. It is clear that this change favors papers from certain disciplines then (e.g. life sciences). (4) The results of Haustein et al. (2014b) show that approximately 80 % of the articles published in 2012 do not receive any tweet. Most of the articles with tweets received only one tweet. The long tail of papers in the distribution of tweets with zero or only one tweet leads to high percentile values for papers, although they have only one or two tweets. (5) The results of Bornmann (2014b) show that many subject categories (in life sciences) are characterized by low average TC. Only very few categories show higher average counts. This is very different to mean citation rates which exhibit greater variations over the disciplines. The missing variation of average TC over the subject categories let Bornmann (2014b) come to the conclusion that TC should be normalized on a lower level than subject categories. The normalization on the journal level could be an alternative.

² <http://feedback.impactstory.org/knowledgebase/articles/400281-highly-cited-and-other-impact-badges>.

Against the backdrop of these considerations, we develop a first attempt to normalize TC properly in this study which improves the method used by ImpactStory. First of all, the normalization of TC only makes sense, if most of the papers in the reference sets have at least one tweet. Strotmann and Zhao (2015) published the 80/20 scientometric data quality rule: a reliable field-specific study is possible with a database, if 80 % of the field-specific publications are covered in this database. We would like to transfer this rule to Twitter data and propose to normalize Twitter data only then if the field-specific reference sets are covered with at least 80 % on Twitter (coverage means in this context that a publication has at least one tweet). We could use Mendeley disciplines (following ImpactStory) or—which is conventional in bibliometrics—WoS subject categories (sets of journals with similar disciplinary focus) for the normalization process. However, both solutions would lead to the exclusion of most of the fields (because they have more than 20 % papers with zero tweets). Thus, we would like to propose the normalization on the journal-level which is also frequently used in bibliometric studies (Vinkler 2010). Here, the reference set is constituted by the papers which are published in the same journal and publication year.

In this study, we use all articles and reviews published in 2012 as initial publication set. The application of the 80/20 scientometric data quality rule on the journals in the set leads to 413 journals with TC for at least 80 % of the papers (4.3 %) and 9242 journals with TC for less than 80 % (95.7 %). We propose to name the set of journals with high Twitter activity as TI. Because many TI journals have published only a low numbers of papers, we reduced the journals in the TI further on (this will be explained later on in this section). We propose to compose the TI every 12 months (e.g. by Twitter). In other words, every 12 months the journals should be selected in which at least 80 % of the papers had at least one tweet. Then, the papers in these journals are used for evaluative Twitter studies on research units (e.g. institutions or countries).

In order to normalize tweets, we propose to calculate percentiles (following ImpactStory) on the base of the tweets for every paper in a journal. There are several possibilities to calculate percentiles (it is not completely clear which possibility is used by ImpactStory). The formula derived by Hazen (1914) $((i - 0.5)/n * 100)$ is used very frequently nowadays for the calculation of percentiles (Bornmann et al. 2013b). It is an advantage of this method that the mean percentile for the papers in a journal equals 50. Table 1 shows the calculation of TP for an example set of 11 publications in a journal. If the papers in a journal are sorted in descending order by their TC, i is the rank position of a paper and n is the total number

Table 1 The calculation of Twitter percentiles for an example set of 11 publications in a journal

No.	Tweets	i	n	Twitter percentile
11	54	11	11	95.45
10	44	9.5	11	81.82
9	44	9.5	11	81.82
8	13	8	11	68.18
7	10	7	11	59.09
6	9	6	11	50.00
5	1	4	11	31.82
4	1	4	11	31.82
3	1	4	11	31.82
2	0	1.5	11	9.09
1	0	1.5	11	9.09

of papers published in the journal. Paper no. 6 is assigned the percentile 50 because 50 % of the papers in the table have a higher rank (more tweets) and 50 % of the values have a lower rank (fewer tweets). Papers with equal TC are assigned the average rank i in the table. For example, as there are two papers with 44 tweets, they are assigned the rank 9.5 instead of the ranks 9 and 10.

The TP are field-normalized impact scores. The normalization on the base of journals is on a lower aggregation level than the normalization on the basis of WoS subject categories (Bornmann 2014b). WoS subject categories are aggregated journals to journal sets. TP are proposed to use for comparisons between units in science (researchers, research groups, institutions, or countries) which have published in different fields.

Results

The results which are presented in the section “[Differences in Twitter counts between Twitter Index journals](#)” show the differences in TC between the TI journals. We test the field-normalized Twitter scores in “[Validation of Twitter percentiles using the fairness test](#)” whether the field-normalization effectively works. In “[Comparison of countries based on Twitter percentiles](#)”, we present some results on the Twitter impact of countries which is field-normalized.

Differences in Twitter counts between Twitter Index journals

For the calculation of the TP we have identified the 413 journals in 2012 with TC for at least 80 % of the papers. We further excluded 259 journals from the TI, because these journals had less than 100 papers published in 2012. For the calculation of the TP the paper set should not be too small and the threshold of 100 can be well justified: If all papers in a journal with at least 100 papers had different TC, all integer percentile ranks would be occupied. Thus, the set of journals in the TI is reduced from 413 journals with TC for at least 80 % of the papers to 156 journals which have also published at least 100 papers. Table 2 shows a selection of twenty journals with the largest average tweets per paper. A table with data for all 156 journals in the TI is located in the Appendix (see Table 6).

The results in Tables 2 and 6 (in the Appendix) reveal a large heterogeneity between the journals with respect to the average and median number of tweets. Whereas the papers published in the *New England Journal of Medicine* have on average 78.6 tweets, the papers published in the *British Dental Journal* are tweeted on average 16 times. The large differences in average tweets already for the twenty most tweeted journals might demonstrate that the normalization of TC on the journal level seems sensible. In contrast to the results of Bornmann (2014b) on the level of subject categories (see the explanation of the study above), there is a greater variation of average tweets on the journal level. In other words, the TI journals are not characterized by only a few journals with very high average TC and most of the journals with low averages or nearly zero average TC. Thus, the level of journals seems proper for the normalization of TC.

Haustein et al. (2014a) found a broad interest by the general public in papers from the biomedical research, which is also reflected in the average TC in Tables 2 and 6: Many journals form the area of general biomedical research are among the most tweeted journals in the tables.

Table 2 Twenty journals in the Twitter Index with the largest average tweets per paper published in 2012

Journal	Number of papers	Sum of tweets	Average tweets	Median number of tweets	Percentage of tweeted papers
<i>New England Journal of Medicine</i>	215	16,908	78.6	53	100.0
<i>The Lancet</i>	233	10,750	46.1	21	100.0
<i>British Medical Journal</i>	325	12,469	38.4	20.5	98.8
<i>Journal of Medical Internet Research</i>	182	6398	35.2	19	98.9
<i>Psychological Science</i>	213	7024	33.0	11	96.7
<i>PLoS Medicine</i>	114	3754	32.9	17	100.0
<i>Science</i>	826	25,262	30.6	19	99.4
<i>Nature</i>	863	26,326	30.5	13	97.9
<i>Nature Climate Change</i>	120	3008	25.1	8	99.2
<i>Archives of Internal Medicine</i>	103	2491	24.2	4	97.1
<i>Science Translational Medicine</i>	211	4995	23.7	16	99.1
<i>American Journal of Clinical Nutrition</i>	325	6421	19.8	6	92.6
<i>Health Affairs</i>	268	5165	19.3	9	94.8
<i>Journal of the American Medical Association</i>	230	4350	18.9	8.5	98.7
<i>PLoS Biology</i>	152	2799	18.4	7	95.4
<i>Nature Medicine</i>	162	2862	17.7	11	96.9
<i>British Journal of Sports Medicine</i>	198	3329	16.8	5	89.4
<i>BMC Medicine</i>	126	2048	16.3	8	99.2
<i>Scientific Reports</i>	790	12,800	16.2	3	91.9
<i>British Dental Journal</i>	133	2129	16.0	12	100.0

Validation of Twitter percentiles using the fairness test

Bornmann et al. (2013a) proposed a statistical approach which can be used to study the ability of the TP to field-normalize TC (see also Kaur et al. 2013; Radicchi et al. 2008). The approach can be named as fairness test (Radicchi and Castellano 2012) and compares the impact results for the TP with that of bare TC with respect to field-normalization. We already used this test to study field-normalized Mendeley scores (Bornmann and Haunschild 2016).

In the first step of the fairness test (made for TP and TC separately), *all papers from 2012* are sorted in descending order by TP or TC, respectively. Then, the 10 % most frequently tweeted papers are identified and a new binary variable is generated, where 1 marks highly tweeted papers and 0 the rest.

In the second step of the test, all papers are grouped by the main disciplines as defined in the OECD field classification scheme.³ The OECD aggregates WoS subject categories (journal sets composed of Thomson Reuters) to the following broad fields: (1) natural sciences, (2) engineering and technology, (3) medical and health sciences, (4) agricultural sciences, (5) social sciences, and (6) humanities. Thomson Reuters assigns many journals

³ <http://ipscience-help.thomsonreuters.com/incitesLive/globalComparisonsGroup/globalComparisons/subjAreaSchemesGroup/oecd.html>.

to more than one WoS category. Thus, many papers in the TI correspondingly belong to more than one OECD field.

In the third step of the test, the proportion of papers belonging to the 10 % most frequently tweeted papers is calculated for each OECD field—using the binary variable from the first step. If TP were fair, the proportions within the fields should equal the expected value of 10 %. Furthermore, TC should show more and greater deviations from 10 % than TP.

The results of the fairness tests are shown in Table 3: TP is compared with TC. The table shows the total number of papers (published in TI journals) within the OECD fields and the proportion of papers within a field which belongs to the 10 % most frequently tweeted papers. Since there is no paper from the TI journals published in the humanities, this field could not be considered in the analyses.

The largest deviations of the proportions from 10 % can be expected for TC, because TC is not field-normalized. In the interpretation of the proportions in Table 3, we deem deviations of less than 2 % points acceptable (i.e. proportions with greater than 8 % and less than 12 %). The used bibliometric and Twitter data are erroneous, why we cannot expect exact results in Table 3. In other words, if the deviations of the proportions are within ± 2 % around the expected value of 10 % (our range of tolerance), the normalization seems to be successful. We have printed in bold all proportions in the table with a deviation of more than ± 2 %. As the results show, TC is outside the range of tolerance in three out of five fields. The social sciences reveal the largest deviation (with 14.9 %).

The TP shows the intended results in Table 3: All OECD fields have less than 1 % point deviations from 10 %; for three fields the proportion equals the expected value of 10 %. Thus, TC seems to field-normalize TC in all fields properly. However, following the argumentations of Sirtes (2012) and Waltman and van Eck (2013), the favorable results for TP could have a simple reason: The calculation of the TP is based on a field classification scheme which is also used for the fairness tests in Table 3 (the OECD aggregates journals, which we used for the normalization of TC). Therefore, Waltman and van Eck (2013) proposed to repeat the fairness test using another field categorization scheme: they used an algorithmically constructed classification system (ACCS). In 2012, Waltman and van Eck (2012) proposed the ACCS as an alternative scheme in bibliometrics to the frequently used field categorization scheme based on journal sets. The ACCS was developed using direct citation relations between publications. In contrast to the WoS category scheme, where a paper can be assigned to more than one field, each publication is assigned to only one field category in the ACCS. We downloaded the ACCS for the papers at the Leiden Ranking homepage⁴ and used it on the highest field-aggregation level (in order to compare the results with those based on the OECD fields).

The results of the comparison between TC and TP based on ACCS are shown in Table 4. In four out of five fields TC is outside the range of tolerance. The situation improves with percentile-based field normalization, but in mathematics and computer sciences as well in social sciences and humanities the proportions are significantly above the expected value of 10 %.

Taken as a whole, the proportions for TC in Tables 3 and 4 reveal that field-normalization is generally necessary for tweets. Larger deviations from the expected value of 10 % are visible in most of the disciplines. However, the results of the study are ambivalent for TP: In natural sciences, life sciences, health sciences, and engineering TP seems to reflect field-normalized values, but in mathematics and computer science as well as in social sciences and humanities this does not seem to be the case.

⁴ <http://www.leidenranking.com/methodology/fields>.

Table 3 Number of papers and proportion of papers belonging to the 10 % most frequently tweeted papers in five main disciplines (as defined by the OECD)

OECD fields	Twitter counts (TC)		Twitter percentiles (TP)	
	Number of papers	Proportion top-10 %	Number of papers	Proportion top-10 %
Natural sciences	22,932	9.1	22,932	9.8
Engineering and technology	1838	4.5	1838	10.0
Medical and health sciences	29,192	8.9	29,192	10.2
Agricultural sciences	130	0.8	130	10.0
Social sciences	1603	14.9	1603	10.0
Humanities	0		0	

Some papers are counted more than once due to multiple field-assignment

Table 4 Number of papers and proportion of papers belonging to the 10 % most frequently tweeted papers in five fields (as defined by the ACCS on the highest level)

Field	Twitter counts		Twitter percentiles	
	Number of papers	Proportion top-10 %	Number of papers	Proportion top-10 %
Biomedical and health sciences	26,940	11.5	26,940	10.0
Life and earth sciences	4145	18.1	4145	10.8
Mathematics and computer science	289	19.0	289	13.1
Physical sciences and engineering	10,565	3.0	10,565	9.6
Social sciences and humanities	2686	20.7	2686	14.3

Comparison of countries based on Twitter percentiles

In the final part of this study, we use TP to rank the Twitter performance of countries in a first application of the new indicator. The analysis is based on all papers (from 2012) published by the countries which are considered in the TI. Since the results in the section “Validation of Twitter percentiles using the fairness test” show that the normalization of TC is only valid in biomedical and health sciences, life and earth sciences, mathematics and computer science, as well as physical sciences and engineering, we considered only these fields in the country comparison. These fields were selected on the base of the ACCS.

The Twitter impact for the countries is shown in Table 5. The table also presents the proportion of papers published by a country in the TI. As the results reveal, all proportions are less than 10 % and most of the proportions are less than 5 %. With a value of 8.1 %, the largest proportion of papers in the TI is available for the Netherlands. Thus, the calculation of the Twitter impact on the country level is generally based on a small proportion of papers. The tweets per paper vary between 16.9 (Denmark) and Taiwan (3.9). Both countries are also the most and less tweeted countries measured by TP (Denmark = 55.4, Taiwan = 45.6). The Spearman rank-order correlation between tweets per

Table 5 Number of papers, number and proportion of papers in the Twitter Index, as well as sum of tweets, tweets per paper, and median Twitter percentiles for those countries with at least 1000 papers in the Twitter Index

Country	All papers	TI papers	Proportion of papers in TI	Sum of tweets	Tweets per paper	Median TP
Denmark	23,583	1330	5.6	22,460	16.9	55.4
Finland	20,502	1278	6.2	19,336	15.1	54.4
Norway	18,848	1009	5.4	12,349	12.2	53.1
UK	141,236	9956	7.0	143,125	14.4	52.1
Canada	98,649	6230	6.3	65,958	10.6	52.0
Australia	78,780	4787	6.1	54,173	11.3	51.5
Sweden	35,080	1947	5.6	27,052	13.9	51.2
Spain	90,510	3915	4.3	38,399	9.8	50.3
Belgium	28,780	1433	5.0	13,847	9.7	50.2
USA	724,091	50,823	7.0	686,221	13.5	50.1
Netherlands	60,390	4885	8.1	51,437	10.5	49.1
Switzerland	36,838	2379	6.5	24,495	10.3	49.0
Israel	18,467	1004	5.4	10,405	10.4	48.5
Brazil	82,866	1264	1.5	9050	7.2	48.4
Germany	168,769	8888	5.3	72,967	8.2	48.0
Japan	164,930	6137	3.7	46,473	7.6	48.0
Italy	124,321	4254	3.4	34,675	8.2	47.6
China	304,361	7225	2.4	37,071	5.1	47.5
India	76,138	1195	1.6	5603	4.7	47.5
France	138,431	6567	4.7	55,649	8.5	46.6
South Korea	97,786	2303	2.4	13,474	5.9	46.6
Taiwan	61,120	1776	2.9	6893	3.9	45.6

Many papers are multiply counted, because they belong to more than one country

paper and TP is $r_s = 0.9$. Thus, the difference in both indicators to measure Twitter impact on the country level is small.

Discussion

While bibliometrics is widespread used to evaluate the performance of different entities in science, altmetrics offer a new form of impact measurement “whose meaning is barely understood” yet (Committee for Scientific and Technology Policy 2014, p. 3). The meaning of TC is especially unclear, because the meta-analysis of Bornmann (2015) shows that TC does not correlate with citation counts (but other altmetrics do). The missing correlation means for de Winter (2015) that “the scientific citation process acts relatively independently of the social dynamics on Twitter” (p. 1776) and it is not clear how TC can be interpreted. According to Zahedi et al. (2014) we thus need to study “for what purposes and why these platforms are exactly used by different scholars”. Despite the difficulties in the interpretation of TC, this indicator is already considered in the “Snowball Metrics Recipe Book” (Colledge 2014). This report contains definitions of indicators, which have

been formulated by several universities—especially from the Anglo-American area. The universities have committed themselves to use the indicators in the defined way for evaluative purposes.

In this study, we have dealt with the normalization of TC. Since other studies have shown that there are field-specific differences of TC, the normalization seems necessary. However, we followed the recommendation of Bornmann (2014b) that TC should not be normalized on the level of subject categories, but a lower level (see here Zubiaga et al. 2014). We decided to use the journal level, since this level is also frequently used to normalize citations (Vinkler 2010). It is a further advantage of the normalization on the journal level that it levels out the practice of a substantial number of journals to launch a tweet for new papers in that journal: The practice leads to larger expected values for these journals. The problem with Twitter data is that many papers receive zero tweets or only one tweet. In order to restrict the impact analysis on only those journals producing a considerable Twitter impact, we defined the TI containing journals with at least 80 % tweeted papers. For all papers in each TI journal, we calculated TP which range from 0 (no impact) to 100 (highest impact). TP is proposed to use for cross-field comparisons.

We used the fairness test in order to study the field-independency of TP (in comparison with TC). Whereas one test based on the OECD fields shows favorable results for TP in all fields, the other test based on an ACCS points out that the TP can be validly used particularly in biomedical and health sciences, life and earth sciences, mathematics and computer science, as well as physical sciences and engineering. In a first application of TP, we calculated percentiles for countries whereby this analysis show that TP and TC are correlated on a much larger than typical level ($r_s = 0.9$). The high correlation coefficient points out that there are scarcely differences between the indicators to measure Twitter impact. The high correlation might be due to the fact that most of the papers used belong to only two fields (biomedical and health sciences and physical sciences and engineering) whereby the variance according to the fields is reduced between the papers.

This paper proposes a first attempt to normalize TC. Whereas Mendeley counts can be normalized in a similar manner as citation counts (Haunschild and Bornmann 2016), the low Twitter activity for most of the papers complicates the normalization of TC. In order to address the problem of low Twitter activity we defined the TI with the most tweeted journals. For 2012, the TI only contains 156 journals. However, we can expect that the journals in the TI will increase in further years, because Twitter activity will also increase. There is a high probability that the Twitter activity will especially increase in those fields where it is currently low (e.g. mathematics and computer science). The broadening of Twitter activities will also lead to a greater effectiveness of the percentile-based field-normalization, because the variance in fields will increase.

Besides further studies which address the normalization of TC and refine our attempt of normalization, we need studies which deal with the meaning of tweets. Up to now it is not clear what tweets really measure. Therefore, de Winter (2015) speculates the following: “It is of course possible that the number of tweets represents something else than academic impact, for example ‘hidden impact’ (i.e., academic impact that is not detected using citation counts), ‘social impact’, or relevance for practitioners ... Furthermore, it is possible that tweets influence science in indirect ways, for example by steering the popularity of research topics, by faming and defaming individual scientists, or by facilitating open peer review” (de Winter 2015, p. 1776). When the meaning of TC is discussed, the difference between tweets and retweets should also be addressed. Retweets are simply

repetitions of tweets and should actually be handled otherwise than tweets in an impact analysis (Bornmann and Haunschild 2015; Taylor 2013).

Acknowledgments The bibliometric data used in this paper are from an in-house database developed and maintained by the Max Planck Digital Library (MPDL, Munich) and derived from the Science Citation Index Expanded (SCI-E), Social Sciences Citation Index (SSCI), Arts and Humanities Citation Index (AHCI) prepared by Thomson Reuters (Philadelphia, Pennsylvania, USA). The Twitter counts were retrieved from Altmetric.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

Appendix

See Table 6.

Table 6 Journals in the Twitter Index ($n = 156$) sorted by the average tweets per paper published in 2012

Journal	Number of papers	Sum of tweets	Average tweets	Median number of tweets	Percentage of tweeted papers
<i>The Lancet</i>	239	11,246	47.1	21.0	100.0
<i>New England Journal of Medicine</i>	217	17,073	78.7	53.0	100.0
<i>Journal of Orthopaedic Trauma</i>	173	444	2.6	2.0	100.0
<i>British Dental Journal</i>	133	2129	16.0	12.0	100.0
<i>Cancer Cell</i>	114	797	7.0	6.0	100.0
<i>The Lancet Oncology</i>	107	1658	15.5	5.0	100.0
<i>Journal of Glaucoma</i>	101	222	2.2	2.0	100.0
<i>Age and Ageing</i>	147	734	5.0	3.0	99.3
<i>Science</i>	831	25,386	30.5	19.0	99.3
<i>BMC Medicine</i>	126	2048	16.3	8.0	99.2
<i>Cell Stem Cell</i>	122	1093	9.0	6.0	99.2
<i>Nature Climate Change</i>	120	3008	25.1	8.0	99.2
<i>Implementation Science</i>	119	1051	8.8	6.0	99.2
<i>PLoS Medicine</i>	115	3754	32.6	17.0	99.1
<i>Science Translational Medicine</i>	211	4995	23.7	16.0	99.1
<i>RNA</i>	201	716	3.6	3.0	99.0
<i>Journal of Medical Internet Research</i>	182	6398	35.2	19.0	98.9
<i>Pancreas</i>	174	249	1.4	1.0	98.9
<i>International Journal of Gynecological Cancer</i>	254	295	1.2	1.0	98.8
<i>British Medical Journal</i>	328	12,639	38.5	20.5	98.8
<i>Journal of the American Medical Association</i>	232	4382	18.9	8.5	98.7
<i>Journal of Pediatric Hematology/Oncology</i>	220	262	1.2	1.0	98.6
<i>Genome Biology</i>	140	1375	9.8	7.0	98.6

Table 6 continued

Journal	Number of papers	Sum of tweets	Average tweets	Median number of tweets	Percentage of tweeted papers
<i>Journal of Thoracic Oncology</i>	258	399	1.5	1.0	98.1
<i>Transplantation</i>	357	522	1.5	1.0	98.0
<i>International Journal of Public Health</i>	102	370	3.6	3.0	98.0
<i>Pediatric Emergency Care</i>	246	531	2.2	1.0	98.0
<i>Nature</i>	869	26,462	30.5	13.0	97.9
<i>Journal of Chemical Theory and Computation</i>	504	621	1.2	1.0	97.8
<i>Annals of Surgery</i>	301	1487	4.9	3.0	97.7
<i>Journal of Polymer Science Part B: Polymer Physics</i>	171	231	1.4	1.0	97.7
<i>BMC Medical Informatics and Decision Making</i>	159	1056	6.6	4.0	97.5
<i>Nature Geoscience</i>	148	1637	11.1	6.0	97.3
<i>Clinical Journal of Pain</i>	109	400	3.7	2.0	97.2
<i>Obesity Surgery</i>	246	625	2.5	2.0	97.2
<i>Archives of Internal Medicine</i>	103	2491	24.2	4.0	97.1
<i>Genome Research</i>	238	1990	8.4	6.0	97.1
<i>Nature Medicine</i>	164	2890	17.6	11.0	97.0
<i>Anesthesiology</i>	219	572	2.6	2.0	96.8
<i>Psychological Science</i>	213	7024	33.0	11.0	96.7
<i>Cell</i>	415	6337	15.3	7.0	96.4
<i>Genes, Brain and Behavior</i>	106	246	2.3	2.0	96.2
<i>Obesity Reviews</i>	101	929	9.2	5.0	96.0
<i>Nature Cell Biology</i>	126	594	4.7	3.0	96.0
<i>Journal of Nuclear Medicine</i>	242	451	1.9	2.0	95.9
<i>Journal of Burn Care and Rehabilitation</i>	142	200	1.4	1.0	95.8
<i>British Journal of Psychiatry, The</i>	117	1214	10.4	6.0	95.7
<i>BMC Infectious Diseases</i>	391	1157	3.0	2.0	95.7
<i>Journal of Cardiovascular Pharmacology</i>	138	153	1.1	1.0	95.7
<i>Microbes and Infection</i>	160	463	2.9	2.0	95.6
<i>PLoS Biology</i>	152	2799	18.4	7.0	95.4
<i>Journal of Spinal Disorders</i>	105	209	2.0	2.0	95.2
<i>Journal of the National Cancer Institute</i>	124	888	7.2	4.0	95.2
<i>Nature Genetics</i>	225	1963	8.7	6.0	95.1
<i>Clinical Nuclear Medicine</i>	118	258	2.2	2.0	94.9
<i>Organometallics</i>	998	1105	1.1	1.0	94.9
<i>Health Affairs</i>	268	5165	19.3	9.0	94.8
<i>Nature Methods</i>	150	974	6.5	4.0	94.7
<i>BMC Psychiatry</i>	236	1400	5.9	4.0	94.5
<i>Critical Care Medicine</i>	363	1427	3.9	2.0	94.5

Table 6 continued

Journal	Number of papers	Sum of tweets	Average tweets	Median number of tweets	Percentage of tweeted papers
<i>Nuclear Medicine Communications</i>	160	361	2.3	2.0	94.4
<i>Journal of Pediatric Gastroenterology and Nutrition</i>	283	754	2.7	2.0	94.3
<i>Inorganic Chemistry</i>	1560	1785	1.1	1.0	94.2
<i>Neurology</i>	389	3051	7.8	4.0	94.1
<i>Journal of Clinical Gastroenterology</i>	135	258	1.9	1.0	94.1
<i>Canadian Medical Association Journal</i>	114	1897	16.6	7.0	93.9
<i>Aging Cell</i>	112	359	3.2	3.0	93.8
<i>International Journal of Applied Earth Observation and Geoinformation</i>	140	178	1.3	1.0	93.6
<i>Macromolecular Bioscience</i>	166	175	1.1	1.0	93.4
<i>Immunity</i>	164	520	3.2	2.0	93.3
<i>BMC Neurology</i>	163	578	3.5	3.0	93.3
<i>Journal of Applied Ecology</i>	156	924	5.9	3.0	92.9
<i>Advanced Energy Materials</i>	169	184	1.1	1.0	92.9
<i>American Journal of Clinical Nutrition</i>	325	6421	19.8	6.0	92.6
<i>International Journal of Behavioral Nutrition and Physical Activity</i>	148	1372	9.3	5.0	92.6
<i>Scientific Reports</i>	794	12,934	16.3	3.0	91.9
<i>The American Journal of Human Genetics</i>	223	1163	5.2	3.0	91.9
<i>Journal of Occupational and Environmental Medicine</i>	195	369	1.9	1.0	91.3
<i>Pediatric Critical Care Medicine</i>	160	318	2.0	1.0	91.3
<i>Journal of the American Society for Mass Spectrometry</i>	228	242	1.1	1.0	91.2
<i>Neuro-Oncology</i>	159	282	1.8	1.0	91.2
<i>Journal of Clinical Investigation</i>	394	1543	3.9	2.0	91.1
<i>Journal of Nuclear Cardiology</i>	101	200	2.0	2.0	91.1
<i>Ecology Letters</i>	166	1054	6.3	4.0	91.0
<i>Macromolecular Rapid Communications</i>	261	254	1.0	1.0	90.8
<i>Cyberpsychology, Behavior, and Social Networking</i>	107	1359	12.7	3.0	90.7
<i>Macromolecular Chemistry and Physics</i>	265	260	1.0	1.0	90.6
<i>Science Signaling</i>	180	1432	8.0	7.0	90.6
<i>Nature Chemistry</i>	126	804	6.4	4.0	90.5
<i>Sexually Transmitted Infections</i>	125	655	5.2	4.0	90.4
<i>Nature Neuroscience</i>	225	2715	12.1	3.0	90.2
<i>Personality and Social Psychology Bulletin</i>	131	1970	15.0	5.0	90.1

Table 6 continued

Journal	Number of papers	Sum of tweets	Average tweets	Median number of tweets	Percentage of tweeted papers
<i>Steel Research International: A journal for steel and related materials</i>	131	125	1.0	1.0	90.1
<i>Medical Care</i>	179	353	2.0	1.0	89.9
<i>Nature Photonics</i>	118	448	3.8	2.0	89.8
<i>Journal of Animal Ecology</i>	134	350	2.6	2.0	89.6
<i>Dalton Transactions</i>	1709	1783	1.0	1.0	89.5
<i>British Journal of Sports Medicine</i>	198	3329	16.8	5.0	89.4
<i>Angewandte Chemie International Edition</i>	2227	4408	2.0	1.0	89.1
<i>World Journal of Surgical Oncology</i>	276	402	1.5	1.0	89.1
<i>The Oncologist</i>	180	612	3.4	3.0	88.3
<i>Nature Chemical Biology</i>	128	458	3.6	3.0	88.3
<i>BMC Public Health</i>	1125	5718	5.1	3.0	88.3
<i>Journal of Asthma</i>	152	635	4.2	3.0	88.2
<i>Medicine and Science in Sports and Exercise</i>	310	4307	13.9	3.0	88.1
<i>Nutrition Journal</i>	114	1497	13.1	5.0	87.7
<i>Free Radical Research</i>	137	138	1.0	1.0	87.6
<i>Chemistry Central Journal</i>	177	224	1.3	1.0	87.6
<i>Virology Journal</i>	320	679	2.1	2.0	87.5
<i>Shock</i>	183	220	1.2	1.0	87.4
<i>BMC Pregnancy and Childbirth</i>	164	865	5.3	3.0	87.2
<i>Pediatric Allergy and Immunology</i>	109	650	6.0	5.0	87.2
<i>American Journal of Clinical Oncology</i>	101	103	1.0	1.0	87.1
<i>Annals of Plastic Surgery</i>	250	264	1.1	1.0	86.8
<i>BMC Family Practice</i>	128	491	3.8	3.0	86.7
<i>Pain</i>	269	1001	3.7	2.0	86.6
<i>Pediatric Anesthesia</i>	170	316	1.9	2.0	86.5
<i>Journal of Clinical Psychology</i>	110	354	3.2	2.0	86.4
<i>Allergy</i>	194	1025	5.3	4.0	86.1
<i>Spine</i>	557	1852	3.3	1.0	86.0
<i>Critical Care (BMC)</i>	298	1041	3.5	2.0	85.9
<i>Journal of Strength and Conditioning Research</i>	390	3719	9.5	4.0	85.9
<i>Journal of Nutrition</i>	311	1089	3.5	2.0	85.9
<i>Journal of Experimental Medicine</i>	190	484	2.5	2.0	85.8
<i>European Journal of Neuroscience</i>	359	1050	2.9	2.0	85.5
<i>Journal of Hospital Medicine</i>	130	671	5.2	3.5	85.4
<i>Medical Education</i>	116	476	4.1	3.0	85.3

Table 6 continued

Journal	Number of papers	Sum of tweets	Average tweets	Median number of tweets	Percentage of tweeted papers
<i>Macromolecular Materials and Engineering</i>	116	107	0.9	1.0	85.3
<i>Psychosomatic Medicine</i>	122	333	2.7	1.0	85.2
<i>Clinical and Experimental Allergy</i>	169	722	4.3	4.0	85.2
<i>Acta Crystallographica Section C: Crystal Structure Communications</i>	229	224	1.0	1.0	84.3
<i>Nature Nanotechnology</i>	120	908	7.6	2.0	84.2
<i>Journal of Organometallic Chemistry</i>	410	362	0.9	1.0	84.1
<i>Obstetrics and Gynecology</i>	307	1749	5.7	2.0	84.0
<i>Anaesthesia</i>	149	413	2.8	2.0	83.9
<i>Journal of Feline Medicine and Surgery</i>	130	212	1.6	1.0	83.8
<i>Methods in Ecology and Evolution</i>	123	535	4.3	3.0	83.7
<i>Cell Reports</i>	242	839	3.5	2.0	83.5
<i>Journal of Child Psychology and Psychiatry and Allied Disciplines</i>	121	508	4.2	2.0	83.5
<i>Journal of Ecology</i>	139	390	2.8	2.0	83.5
<i>Anesthesia and Analgesia</i>	300	641	2.1	2.0	83.3
<i>Molecular Cell</i>	296	729	2.5	2.0	83.1
<i>Coordination Chemistry Reviews</i>	136	148	1.1	1.0	83.1
<i>Therapeutic Drug Monitoring</i>	103	209	2.0	1.0	82.5
<i>Journal of Clinical Psychopharmacology</i>	103	190	1.8	1.0	82.5
<i>Annals of Pharmacotherapy, The Pediatrics</i>	181	467	2.6	2.0	82.3
<i>Pediatrics</i>	698	8318	11.9	3.0	82.2
<i>Annals of Allergy, Asthma and Immunology</i>	145	684	4.7	3.0	82.1
<i>Brain</i>	280	1160	4.1	2.0	81.8
<i>Journal of the American Medical Informatics Association</i>	186	1058	5.7	2.0	81.7
<i>Conservation Biology</i>	114	420	3.7	2.0	81.6
<i>Nature Immunology</i>	124	319	2.6	2.0	81.5
<i>Plant Journal</i>	345	769	2.2	2.0	80.9
<i>American Journal of Psychiatry, The Neuron</i>	108	451	4.2	2.0	80.6
<i>Neuron</i>	340	2346	6.9	3.0	80.3
<i>Journal of Personality and Social Psychology</i>	141	1084	7.7	3.0	80.1

References

- Adie, E. (2014). Taking the alternative mainstream. *Profesional De La Informacion*, 23(4), 349–351. doi:10.3145/epi.2014.jul.01.
- Bornmann, L. (2011). Scientific peer review. *Annual Review of Information Science and Technology*, 45, 199–245.

- Bornmann, L. (2013). What is societal impact of research and how can it be assessed? A literature survey. *Journal of the American Society of Information Science and Technology*, 64(2), 217–233.
- Bornmann, L. (2014a). Do altmetrics point to the broader impact of research? An overview of benefits and disadvantages of altmetrics. *Journal of Informetrics*, 8(4), 895–903. doi:10.1016/j.joi.2014.09.005.
- Bornmann, L. (2014b). Validity of altmetrics data for measuring societal impact: A study using data from Altmetric and F1000Prime. *Journal of Informetrics*, 8(4), 935–950.
- Bornmann, L. (2015). Alternative metrics in scientometrics: A meta-analysis of research into three altmetrics. *Scientometrics*, 103(3), 1123–1144.
- Bornmann, L. (in press). Measuring impact in research evaluations: A thorough discussion of methods for, effects of, and problems with impact measurements. *Higher Education*. Retrieved June 2015, from <http://arxiv.org/abs/1410.1895>
- Bornmann, L., de Moya Anegón, F., & Mutz, R. (2013a). Do universities or research institutions with a specific subject profile have an advantage or a disadvantage in institutional rankings? A latent class analysis with data from the SCImago ranking. *Journal of the American Society for Information Science and Technology*, 64(11), 2310–2316.
- Bornmann, L., & Haunschild, R. (2015). *t factor: A metric for measuring impact on Twitter*. Retrieved October 30, 2015, from <http://arxiv.org/abs/1508.02179>.
- Bornmann, L., & Haunschild, R. (2016). Normalization of Mendeley reader impact on the reader- and paper-side: A comparison of the Mean Discipline Normalized Reader Score (MDNRS) with the Mean Normalized Reader Score (MNRS) and bare reader counts. <https://dx.doi.org/10.6084/m9.figshare.2554957.v1>.
- Bornmann, L., Leydesdorff, L., & Mutz, R. (2013b). The use of percentiles and percentile rank classes in the analysis of bibliometric data: Opportunities and limits. *Journal of Informetrics*, 7(1), 158–165.
- Chamberlain, S. (2013). Consuming article-level metrics: Observations and lessons. *Information Standards Quarterly*, 25(2), 4–13.
- Colledge, L. (2014). *Snowball metrics recipe book*. Amsterdam: Snowball Metrics Program Partners.
- Committee for Scientific and Technology Policy. (2014). *Assessing the impact of state interventions in research—Techniques, issues and solutions*. Brussels: Directorate for Science, Technology and Innovation.
- Darling, E. S., Shiffman, D., Côté, I. M., & Drew, J. A. (2013). The role of Twitter in the life cycle of a scientific publication. *PeerJ PrePrints*, 1, e16v1. doi:10.7287/peerj.preprints.16v1.
- de Winter, J. C. F. (2015). The relationship between tweets, citations, and article views for PLOS ONE articles. *Scientometrics*, 102(2), 1773–1779. doi:10.1007/s11192-014-1445-x.
- Fenner, M. (2014). *Altmetrics and other novel measures for scientific impact*. Retrieved July 8, 2014, from http://book.openingscience.org/vision/altmetrics.html?utm_content=buffer94c12&utm_medium=social&utm_source=twitter.com&utm_campaign=buffer.
- Haunschild, R., & Bornmann, L. (2016). Normalization of Mendeley reader counts for impact assessment. *Journal of Informetrics*, 10(1), 62–73.
- Haustein, S., Larivière, V., Thelwall, M., Amyot, D., & Peters, I. (2014a). Tweets vs. Mendeley readers: How do these two social media metrics differ? *It Information Technology*, 56(5), 207–215.
- Haustein, S., Peters, I., Sugimoto, C. R., Thelwall, M., & Larivière, V. (2014b). Tweeting biomedicine: An analysis of tweets and citations in the biomedical literature. *Journal of the Association for Information Science and Technology*, 65(4), 656–669. doi:10.1002/asi.23101.
- Hazen, A. (1914). Storage to be provided in impounding reservoirs for municipal water supply. *Transactions of American Society of Civil Engineers*, 77, 1539–1640.
- Hicks, D., Wouters, P., Waltman, L., de Rijcke, S., & Rafols, I. (2015). Bibliometrics: The Leiden Manifesto for research metrics. *Nature*, 520(7548), 429–431.
- Holmberg, K. (2014). *The impact of retweeting on altmetrics*. Retrieved July 8, 2014, from <http://de.slideshare.net/kholmber/the-meaning-of-retweeting>.
- Kaur, J., Radicchi, F., & Menczer, F. (2013). Universality of scholarly impact metrics. *Journal of Informetrics*, 7(4), 924–932. doi:10.1016/j.joi.2013.09.002.
- Leydesdorff, L., & Bornmann, L. (2011). Integrated impact indicators compared with impact factors: An alternative research design with policy implications. *Journal of the American Society for Information Science and Technology*, 62(11), 2133–2146. doi:10.1002/asi.21609.
- NISO Alternative Assessment Metrics Project. (2014). *NISO Altmetrics Standards Project White Paper*. Retrieved July 8, 2014, from http://www.niso.org/apps/group_public/document.php?document_id=13295&wg_abbrev=altmetrics.
- Priem, J. (2014). Altmetrics. In B. Cronin & C. R. Sugimoto (Eds.), *Beyond bibliometrics: Harnessing multi-dimensional indicators of performance* (pp. 263–288). Cambridge, MA: MIT Press.

- Priem, J., & Costello, K. L. (2010). How and why scholars cite on Twitter. *Proceedings of the American Society for Information Science and Technology*, 47(1), 1–4. doi:10.1002/meet.14504701201.
- Priem, J., & Hemminger, B. M. (2010). Scientometrics 2.0: Toward new metrics of scholarly impact on the social Web. *First Monday*, 15(7).
- Radicchi, F., & Castellano, C. (2012). Testing the fairness of citation indicators for comparison across scientific domains: The case of fractional citation counts. *Journal of Informetrics*, 6(1), 121–130. doi:10.1016/j.joi.2011.09.002.
- Radicchi, F., Fortunato, S., & Castellano, C. (2008). Universality of citation distributions: Toward an objective measure of scientific impact. *Proceedings of the National Academy of Sciences*, 105(45), 17268–17272. doi:10.1073/pnas.0806977105.
- Roemer, R. C., & Borhardt, R. (2013). Institutional altmetrics & academic libraries. *Information Standards Quarterly*, 25(2), 14–19.
- Shema, H., Bar-Ilan, J., & Thelwall, M. (2014). Do blog citations correlate with a higher number of future citations? Research blogs as a potential source for alternative metrics. *Journal of the Association for Information Science and Technology*, 65(5), 1018–1027. doi:10.1002/asi.23037.
- Sirtes, D. (2012). Finding the Easter eggs hidden by oneself: Why fairness test for citation indicators is not fair. *Journal of Informetrics*, 6(3), 448–450. doi:10.1016/j.joi.2012.01.008.
- Strotmann, A., & Zhao, D. (2015). *An 80/20 data quality law for professional scientometrics?* Paper presented at the Proceedings of ISSI 2015—15th International society of scientometrics and informetrics conference, Istanbul, Turkey.
- Taylor, M. (2013). Towards a common model of citation: Some thoughts on merging altmetrics and bibliometrics. *Research Trends*, 35, 19–22.
- Vinkler, P. (2010). *The evaluation of research by scientometric indicators*. Oxford: Chandos Publishing.
- Waltman, L., & van Eck, N. J. (2012). A new methodology for constructing a publication-level classification system of science. *Journal of the American Society for Information Science and Technology*, 63(12), 2378–2392. doi:10.1002/asi.22748.
- Waltman, L., & van Eck, N. J. (2013). A systematic empirical comparison of different approaches for normalizing citation impact indicators. *Journal of Informetrics*, 7(4), 833–849.
- Wilsdon, J., Allen, L., Belfiore, E., Campbell, P., Curry, S., Hill, S., et al. (2015). *The metric tide: Report of the independent review of the role of metrics in research assessment and management*. Bristol: Higher Education Funding Council for England (HEFCE).
- Zahedi, Z., Costas, R., & Wouters, P. (2014). How well developed are altmetrics? A cross-disciplinary analysis of the presence of 'alternative metrics' in scientific publications. *Scientometrics*, 101(2), 1491–1513. doi:10.1007/s11192-014-1264-0.
- Zubiaga, A., Spina, D., Martínez, R., & Fresno, V. (2014). Real-time classification of twitter trends. *Journal of the Association for Information Science and Technology*. doi:10.1002/asi.23186.