Do blog citations correlate with a higher number of future citations? Research blogs as a potential source for alternative metrics[1]

Hadas Shema*, Judit Bar-Ilan* and Mike Thelwall**

*Judit.Bar-Ilan@.biu.ac.il; dassysh@gmail.com

Department of Information Science, Bar-Ilan University, Ramat-Gan, 52900 (Israel)

** m.thelwall@wlv.ac.uk

School of Technology, University of Wolverhampton, Wulfruna Street, Wolverhampton WV1 1LY (UK).

Abstract

Journal-based citations are an important source of data for impact indices. However, the impact of journal articles is not limited to other scholarly material, but extends beyond formal scholarly discourse. Measuring online scholarly impact calls for new indices, complementary to the older ones. In this article, we study a possible alternative metric source, blog posts aggregated at ResearchBlogging.org, which discuss peer-reviewed articles and provide full bibliographic references. Articles reviewed in these blogs therefore receive "blog citations". We hypothesized that articles receiving blog citations close to their publication time receive more journal citation later on than the articles in the same journal published in the same year that did not receive such blog citations. Statistically significant evidence for articles published in 2009 and 2010 support this hypothesis for 7 out of 12 journals (58%) in 2009 and 13 out of 19 journals (68%) in 2010. Based on these results, we propose blog citations as an alternative metric source.

Introduction

Traditional scholarly impact metrics live in an ivory tower made of formal publications. The citations generated from peer-reviewed publications have been, for decades, the building blocks for impact metrics, which rely on the slow accumulation of citations from one peer-reviewed publication to another. However, the age of the Web has given rise to new venues of discussion and dissemination of scholarly information. These highlight the limitations of traditional indices and the need for additional impact metrics in the bibliometric tool box to supplement the existing indices.

In the past few decades, the growing popularity of bibliometric indices has led to a thorough study of the citation process. The type of document, its subject, its publishing venue, authors and other characteristics all influence its citation impact in, statistically speaking, predictable ways.

The number of citations received by a document has been known to be affected by its type. Looking at Norwegian articles from the years 1981-1996, Aksnes (2003) found

[1] This is a preprint of an article to be published in the Journal of the American Society for Information Science and Technology

that while only 2% of the articles he studied were review articles, they formed 12% of the highly cited articles. Notes and proceeding papers, on the other hand, were less represented among the highly cited than in the general article sample.

Citations are not equally distributed – in fact, their distribution is highly skewed (Seglen, 1992). Therefore, attributing the average number of journal citations to a single journal article can be misleading. A small number of highly cited articles greatly impacts average-based indices (e.g., the Journal Impact Factor).

Scientists do not exist in a vacuum, and neither do citations. Phillips, Kanter, Bednarczyk and Tastad (1991) compared citations to New England Journal of Medicine (NEJM) articles covered by the New York Times with similar NEJM articles that were not covered by the New York Times, and found the covered articles to have a citation advantage. It was possible that the advantage came from the New York Times' ability to pick better articles for coverage, but a 12-weeks New York Times strike at 1978 allowed the authors to put the hypothesis into test by looking at articles selected during the strike (the newspaper produced an "edition of record" which was not publicly distributed). They found that articles covered by the Times received almost 73% more citations than control articles that appeared in the same NEJM issues and the same category, but were not covered by the New York Times, in the first year after publication, and continued to receive more citations in the ten years after their publication. This effect was not present for articles selected during the Times strike, providing evidence to the ability of the mass media to affect scholarly research distribution.

As scholarly communication migrated to the Web, so did citations. However, the meaning of “Web citation” remained rather vague, since the Web is made of much more than formal research discourse, and citations can appear everywhere. Vaughan and Shaw (2003) were the first to investigate Web citations to academic articles on a large scale, but did not offer a clear theory-based definition of them. Based on Vaughan and Shaw’s work, Bar-Ilan (2008) had broadly defined the term as *“an appearance of the title of a publication within a webpage (not necessarily as a link),”* (section 8.5). A more vague definition was offered by Thelwall, Vaughan and Björneborn (2005) *“how often the journal articles were mentioned in Web pages,”* (p. 101). When compared to the well-studied journal citations, Web citations, especially in the social media, are still somewhat of a mystery. If peer-reviewed journal citations are "*frozen footprints*" (Cronin, 1981, p.16) and "*signposts left behind after information has been utilized*" (Smith, 1981, p. 85) what are Web citations, since the Web can change?

Alternative metrics

"Altmetrics," short for alternative metrics, is a term to describe Web-based metrics for the impact of scholarly material, with an emphasis on social media outlets as sources of data (Priem, Taraborelli, Groth& Neylon, 2010). Microblogs (e.g., Twitter), reference managers, research blogs, post-publication peer-review and other resources have been suggested as possible alternative metrics sources.

Founded in 2006, Twitter is a popular microblogging service with more than 200 million active users and 400 million messages ("tweets") being sent each day (Wickre, 2013). Tweets are short messages with up to 140 characters at length. Eysenbach (2011) found a correlation between the number of tweets about Journal of

the Journal of Medical Internet Research (JMIR) articles and future citation counts. A study of the correlation between tweets about ArXiv repository pre-prints and Google Scholar citations showed similar results (Shuai, Pepe & Bollen, 2012). These findings show tweeting to be a promising altmetric source. However, the difficulty of archiving extremely large amounts of tweets and retrieving them might prove a challenge to researchers. Twitter does not provide freely full access to all of its tweets, but access can be obtained through a reseller for a price (Tornes, 2013). In addition, Twitter has donated tweets to the Library of Congress, which has archived tweets from the years 2006-2010 so far. However, a single search in the archive currently takes about 24 hours and requires physical presence at the library's building (Library of Congress, 2013). Other than the difficulties described above, the lack of effort required to produce a tweet might make tweet-based altmetrics an easy target of manipulation.

Looking at data from the commercial service altmetric.com and the Web of Science (WoS) database, Thelwall, Haustein, Larivière and Sugimoto (2013) studied associations between journal citations and different altmetric indicators. Their sample included articles archived by the biomedical and life science database PubMed between July 2011 and January 1st, 2013. Six out of the eleven altmetrics sources they studied (tweets, Facebook wall posts, research highlights, blog mentions, mainstream media mentions and forum posts) had significant associations between higher altmetric scores and higher numbers of citations, suggesting that multiple different types of altmetrics may be valid and useful. F1000Prime (formerly Faculty of 1000, F1000 from now on) is a commercial post-publication peer-review service. It offers reviews of published articles' scientific quality in the medical and life science by about 10,000 experts (F1000, 2012a) and more than 1,500 articles are recommended every month (F1000, 2012b). Several studies (e.g. Li & Thelwall, 2012; Mohammadi & Thelwall, 2013a and Waltman & Costas, 2013) have been conducted to study this platform and compare it to reference managers and to citations, to study the dynamics of the reviewing process, and the types of articles being reviewed.

Scholarly social bookmarking services such as Mendeley and CiteULike allow the storing and sharing of scholarly material. The number of users who saved an item by bookmarking it (called "readers" on Mendeley) is shown next to every item listed in those services. Users are also capable of tagging items with freely chosen keywords.

The easily accessible, large numbers of readers in reference managers make for a promising altmetrics source. However, the lack of context makes it difficult to determine the underlying use made of a bookmarked article. The users might be called "readers" but it is possible that they have not read the item they bookmarked or that they have read it but did not make use of it. On the other hand, it could be that they use reference managers in order to easily access important articles over and over again. Correlations between reader counts and citations have been studied in various settings: for Science and Nature articles (Li, Thelwall & Giustini, 2012) for JASIST articles (Bar-Ilan, 2012a), for articles in the fields of the social sciences and the humanities (Mohammadi & Thelwall, 2013b), for articles published by scientometricians (Bar-Ilan et al., 2012). These studies showed significant correlations of around 0.5 between Mendeley reader counts and citations. While reader counts have their limitations, the broad coverage of reference managers, especially Mendeley, and their continuous growth can make them an important altmetric source.

Science and research blogs

Science blogs publish posts related to science and review scientific developments, becoming popular with a section of the scholarly community. Respected scholarly media outlets such as National Geographic, the Nature Group, Scientific American and the PLoS journals all have science blogging networks. A Nature Medicine editorial, discussing blogs and peer review concluded that *"Online science blogs are a valuable forum for commenting on published research, but their present importance lies in complementing rather than replacing the current system of peer review"* (Perfecting peer review?, 2011, p. 1-2).

Kousha, Thelwall, and Rezaie (2010) have shown that it is possible, at least on a small scale, to calculate blog mentions for a set of published articles by using Google Blog Search. They concluded that, although blog citations were found to be far less common than academic citations, they could still be useful evidence of research impact on wider discussions, especially in the social sciences and humanities. While Kousha et al. (2010) considered every mention of scholarly article in blog as a citation, we would like to differentiate between *blog mentions* and *blog citations*. Blog mentions are any sort of reference to scholarly material in blogs, while blog citations cite scholarly materials in structured, formal styles (e.g., APA, MLA) and appear in blog posts.

Unlike authors of peer-reviewed articles, bloggers are not obligated to refer their sources in a formal way. Despite this, Kjellberg (2010) found that scientist bloggers would like to employ formal referencing norms in their blogs.

"A recurring topic in the interviews has to do with the fact that the researchers want to use references and point to their sources in the blog, in a similar way to how they do in regular communication within the scholarly environment,"(para. 21).

Further evidence for the diffusion of norms from formal academic citing to blogs is the aggregator ResearchBlogging.org (RB for short). ResearchBlogging.org (2008) aggregates blog posts referring specifically to peer-reviewed research. It is a self-selecting aggregator that allows bloggers to refer to peer-reviewed research in an academic citation format. Bloggers discussing peer-reviewed research can register with the aggregator and after they mark relevant posts in their blog, these posts appear on the aggregator site, giving one-stop access to a variety of research reviews from different authors. The site's human editors ensure that blogs submitted to the aggregator follow its guidelines and are of appropriate quality. RB already has an altmetric role, since it currently serves as one of the article level metrics (ALM) displayed for each article in the journal PLoS ONE. By the end of 2011 RB had over 1,230 active blogs and about 27,000 posts.

Groth and Gurney (2010) were the first to conduct an RB study and focused on posts tagged "Chemistry." Those posts referred to literature that was mostly up-to-date and published by top-tier journals: over 70% of the cited articles were from the top 20 journals in the field of Chemistry, and 21% were from the 60 top publications across all disciplines. Other studies (Shema, Bar-Ilan & Thelwall, 2012; Fausto et al., 2012) have indicated that the most cited publications in blog posts (in no specific order) were the multidisciplinary journals Science, Nature, Proceedings of the National Academy of Sciences of the United States of America (PNAS) and the Open-Access (OA) PLoS ONE are by far the most popular in RB, with 36% of the total. Other

popular categories were Health Sciences (15%) and Psychology (13%). Only 11.7% of the citations came from OA journals (Fausto et al., 2012).

Science bloggers link to various sources, including other blogs, mainstream media and scholarly material. A survey of SciLogs bloggers (a German blogging platform) showed that they were equally likely to have a post topic brought to their attention by the mainstream media as by scholarly publications (Puschmann & Mahrt, 2012). Looking at bloggers affiliated with research institutes from Scienceblogs.com and Scienceblogs.de Peters, Beutelspacher, Maghferat and Terliesner (2012) found that bloggers often link to their own blogs or other blogs in their platform, to social media sites and to major news sites, such as Spiegel or the New York Times.

Given the bloggers' usage of formal citation norms and scholarly sources, a connection to the research community is not unexpected. Shema et al. found that a majority (59%) of science bloggers were part of the academic community in some capacity. Puschmann and Mahrt (2012) found that 43% of the SciLogs bloggers were employed in the academy. In both studies the bloggers were highly educated, with 32% of the RB bloggers and 45% of the SciLogs bloggers having earned a PhD. Bora Zivkovic, the former editor of PLoS blogs and current editor of Scientific American blogs, estimated that *"[Blogs are] written by graduate students, postdocs and young faculty, a few by undergraduates and tenured faculty, several by science teachers, and just a few by professional journalists"* (Bonetta, 2007, p. 443).

We hypothesize that, since many of the bloggers are or were part of the academic community, they would be capable of recognizing articles that will appeal to its members. Therefore, we examined whether articles that were published in peer-reviewed journals and were reviewed in blogs aggregated by ResearchBlogging.org soon after their publication were more highly cited than articles published in the same year and in the same journal but that are not reviewed in the year of their publication in blogs aggregated by ResearchBlogging.org.

Data and methods

ResearchBlogging.org publishes an extended snippet of all the posts that it aggregates. An example of such a snippet can be seen in Figure 1. All the snippets of posts published during 2009 and 2010 were downloaded using the DownThemAll add-on on to Firefox (http://www.downthemall.net/). Altogether 4878 snippets from 2009 and 7777 from 2010 were downloaded. We developed software to automatically extract the following fields from these snippets: date of publication of the post, number of views of the post, title and URL of the blog post, name of the blogger and of the blog, and for each citation that appeared in the blog post (there are posts that contain several blog citations): author, title, year, source and DOI or URL of the specific publication. Altogether 6,927 and 11,500 blog citations were identified by this process for 2009 and 2010 respectively.

Since we were interested in blog citations which appeared soon after the publication of an article, we considered only blog posts reviewing articles published in 2009 and 2010 (4013 and 6116 items respectively). We only considered blog posts from the year of publication of the article (e.g. a 2009 blog post could only discuss a 2009 article). Next we limited the sample only to journals with 20 or more articles published in the journal and reviewed in ResearchBlogging.org during 2009 and 2010 respectively. The 20-article threshold was a compromise between the need to obtain statistically reliable results and the need to include as many journals as possible in the analysis. Editorials, letters and other document types were excluded, leaving only articles, reviews and proceedings papers to be considered. Articles which appeared numerous times in the sample were only taken into consideration once. A list of journals appears in Tables 1 and 2. Details of the articles published in these journals during 2009 and 2010 and the citations they received in 2009, 2010, 2011 and 2012 were retrieved from the Web of Science (WoS). The WoS database records articles according to their official publication date rather than the online publication date. Therefore we used the official date, even if an online version was published before it.

Table 1: Journals with more than 20 articles published in 2009 and reviewed in 2009 in blog posts aggregated by ResearchBlogging.org by decreasing number of reviewed articles in blogs

Journal	**# articles published by the journal in 2009**	**# articles reviewed by bloggers in 2009**	**articles reviewed by bloggers in 2009 in % from the number of overall articles published by each journal**
PLoS One	4403	193	4.4
PNAS	3765	166	4.4
Science	897	161	17.9
Nature	866	119	13.7
Psychological Science	234	49	20.9
Journal of Neuroscience	1542	40	2.6
Journal of the American Chemical Society	3332	34	1.02
Current Biology	357	28	7.8
PLoS Biology	195	26	13.3
New England Journal of Medicine	352	26	7.4
Pediatrics	752	23	3.1
Nature Neuroscience	208	22	10.6

Three journals (Current Biology, Journal of the American Chemical Society and Nature Neuroscience) fell below the 20-article threshold in 2010 and were removed from the list. Ten others journals passed the threshold and were added to the 2010 list.

Table 2: Journals with more than 20 articles published in 2010 and reviewed in 2010 in blog posts aggregated by ResearchBlogging.org by decreasing number of reviewed articles in blogs

Journal	**# articles published by the journal in 2010**	**# articles reviewed by bloggers in 2010**	**articles reviewed by bloggers in 2010 in % from the number of overall articles published by each journal**
PLoS One	6723	288	4.3
PNAS	3765	243	6.5
Nature	862	196	22.7
Science	861	171	19.9
Psychological Science	284	71	25.0
Journal of Neuroscience	1662	67	4.0
PLoS Biology	214	40	18.7
New England Journal of Medicine	345	38	31
Physical Review Letters	3107	37	1.2
JAMA	232	32	13.8
Proceedings of the Royal Society B-Biological Sciences	452	27	6.0
Conservation Biology	171	26	15.2
Ecological Applications	177	24	13.6
Lancet	271	24	8.9
Biological Conservation	315	23	7.3
Cell	320	23	7.2
Pediatrics	702	23	3.3
PLoS Computational Biology	406	21	5.2
Biology Letters	216	20	9.3

The results from Table 1 and Table 2 validate those of earlier studies regarding the most popular journals (PLoS ONE, PNAS, Science and Nature). In addition, they show bloggers' preferences for the biological and medical disciplines, with 8 out of the 12 journals (67%) in 2009 and 14 out of 19 journals (74%) in 2010 belonging to those fields

Results

In this section we present findings separately for the years 2009 and 2010. First, we show results for one-sample binomial tests (Table 3) ran on the aggregated results of all the journals from Tables 1 and 2. Then, the results are broken down and the medians for every journal's sample group are shown in comparison with the medians of articles from which were not covered by the bloggers (Tables 4 and 5). Results of Mann-Whitney tests are presented for each journal (Tables 6 and 7). Next, we show the number of reviews in each journal in comparison with the number of reviews covered by bloggers (Table 8). Last, we present a case study of the overlapping between the New England Journal of Medicine articles covered by bloggers and those reported by the media.

One-sample binomial test

The article population medians were calculated separately for each journal for the years 2009 and 2010 (data not shown). We ran a non-parametric one-sample binomial test to find if the aggregated citation categories below and above the journals' medians had different probabilities than the expected .5 (Table 3). The null hypothesis, that the categories occur with probability of .5 and .5 was rejected for 2009 and 2010 ($p<.001$). Hence, we have statistical evidence that citations attracted by blogged articles tend to be above the median for the journal in which they are published.

Table 3: Number of sample articles below and above each journal's median for 2009 and 2010.

	2009-11	**2010-12**
Above *Median*	507	823
Below *Median*	380	571

Median differences

It is well-known that citation distributions are highly skewed (Seglen, 1992), thus it is appropriate to consider medians instead of averages (Bar-Ilan, 2012b). Most of the journals in the sample belong to the life sciences, for which a two or three year citation window is considered adequate, due to the fast ageing of most journals and topics in the area (Glänzel & Schoepflin, 1995). In light of those past findings, we summed for each 2009 article the number of citations it received during 2009, 2010 and 2011 (Table 4).

Table 4: Median number of citations received by the reviewed and the non-reviewed articles in 2009

Journal	**Median # citations received during 2009-2011 for 2009 articles reviewed in RB blogs in 2009**	**Median # citations received during 2009-2011 for 2009 articles not reviewed in RB blogs in 2009**
PLoS One	8	6
PNAS	20	16
Science	41	40
Nature	57	49
Psychological Science	8	9
Journal of Neuroscience	22	12
Journal of the American Chemical Society	19	14
Current Biology	13.5	15
PLoS Biology	18.5	17
New England Journal of Medicine	172	56
Pediatrics	13	7
Nature Neuroscience	32.5	24

We see that the medians are higher for the articles that received blog citations except for the journals Psychological Science and PLOS Biology. The most striking difference is for the New England Journal of Medicine; the median number of citation received by articles that received early blog citations is more than 3 times the median number of citation received by the articles that were not reviewed in 2009 in blog posts aggregated by ResearchBlogging.org.

Table 5: Median number of citations received by the reviewed and the non-reviewed articles in 2010.

Journal	**Median # citations received during 2010-2012 for 2010 articles reviewed in RB blogs in 2010**	**Median # citations received during 2010-2012 for 2010 articles not reviewed in RB blogs in 2010**
PLoS One	7	5
PNAS	23	15
Nature	60.5	49
Science	47.5	40.5
Psychological Science	5	7
Journal of Neuroscience	17	12
PLoS Biology	23.5	15
New England Journal of Medicine	138	51
Physical Review Letters	19	11
JAMA	38.5	36
Proceedings of the Royal Society B- Biological Sciences	8	8
Conservation Biology	7	6
Ecological Applications	10	6
Lancet	99	50
Biological Conservation	7	6
Cell	77	43
Pediatrics	14	6
PLoS Computational Biology	6	7
Biology Letters	8.5	6

Other than for the journals Psychological Science and PLOS Computational Biology, all the medians of reviewed articles group are either equal or higher than those of the articles which were not reviewed by the bloggers (Table 5).

Table 4 and 5 show that for most journals, the median numbers of citations of articles covered in blogs is higher than those of articles which were not covered in them. However, we could not tell by medians alone whether the blogged articles had a statistically significant citation advantage over those who were not. In order to study this at the level of journals, we conducted a series of non-parametric Mann-Whitney tests.

Mann-Whitney tests

Table 6 shows the p-values of the Mann-Whitney tests for differences between the blogged and non-blogged groups from 2009 for the citation periods 2009-2011. For 7 out of the 12 journals (58%) the differences are significant at $p<.05$ (for six journals the differences are significant at $p<.01$). The results for the Journal of the American Chemical Society are at the edge of significance.

Table 6: Results of Mann-Whitney tests, 2009.

Journal	**p-values for the citation period 2010-2012**
PLoS One	.002**
PNAS	.000**
Science	.975
Nature	.044*
Psychological Science	.833
Journal of Neuroscience	.000**
Journal of the American Chemical Society	.059
Current Biology	.253
PLoS Biology	.988
New England Journal of Medicine	.000**
Pediatrics	.004**
Nature Neuroscience	.003**

*$p<.05$. **$p<.01$.

Table 7 shows the results of Mann-Whitney tests ($p<.05$) for each journal from the year 2010 for the citation periods of 2010-2012.

Table 7: Results of Mann-Whitney tests, 2010

Journal	p-values for the citation period 2010-2012
PLoS One	.000**
PNAS	.000**
Nature	.001**
Science	.040*
Psychological Science	.468
Journal of Neuroscience	.001**
PLoS Biology	.001**
New England Journal of Medicine	.000**
Physical Review Letters	.004**
JAMA	.742
Proceedings of the Royal Society B-Biological Sciences	.674
Conservation Biology	.924
Ecological Applications	.027*
Lancet	.006**
Biological Conservation	.206
Cell	.006**
Pediatrics	.000**
PLoS Computational Biology	.603
Biology Letters	.042*

*$p<.05$. **$p<.01$.

In 2010 a total of 13 out of the 19 journals (68%) have significant results for the citation periods studied (for 10 journals the results were significant as $p<.01$).

In order to be sure that blog citations in the year of publication could predict future citation, we repeated the analysis for 2009 with a 2010-2011 citation window and for 2010 with a 2011-2012 citation window (data not shown). There was no change in the statistical significance of any of the findings, showing that the bloggers' advantage, when it exists, does not come from articles which were already well-cited in the year of their publication, but from future citations that the bloggers would be unlikely to know about.

Reviews

In light of review articles' over-representation among highly cited articles, we decided to test whether the bloggers tend to over or under cover review articles from journals and if review articles are connected with a citation advantage. We searched each journal's sample for review articles from the same journal and year. Table 8 shows the overall number of articles classified as reviews in WoS published by a journal at 2009 or 2010, and the number of review articles covered by bloggers each year for the journal.

Table 8: Number of reviews published in sample journals and number of reviews from sample journals covered by bloggers in 2009 and 2010.

Journal	**# reviews published by the journal in 2009**	**# reviews covered by bloggers in 2009**	**# reviews published by the journal in 2010**	**# reviews covered by bloggers in 2010**
PLoS One	9	0	32	2
PNAS	6	0	5	1
Nature	66	2	37	3
Science	54	4	61	8
Psychological Science	0	0	0	0
Journal of Neuroscience	37	2	6	0
PLoS Biology	0	0	0	0
New England Journal of Medicine	33	1	38	2
Physical Review Letters	-	-	0	0
JAMA	-	-	18	3
Proceedings of the Royal Society B-Biological Sciences	-	-	21	1
Conservation Biology	-	-	8	0
Ecological Applications	-	-	0	0
Lancet	-	-	27	1
Biological Conservation	0	0	20	0
Cell	-	-	34	0
Pediatrics	38	0	35	1
PLoS Computational Biology	1	1	1	1
Biology Letters	-	-	0	0
Nature Neuroscience	5	1	-	-
Journal of the American Chemical Society	128	1	-	-
Current Biology	30	0	-	-

We have not found evidence for over or under use of review articles in the samples, but their small size does not allow us to test for statistical significance.

Case study: New England Journal of Medicine

The New England Journal of Medicine (NEJM) is a prestigious medical journal (number one in the WoS category "Medicine, general and internal") and is one of the leading peer-reviewed journals that science reporters rely on (Conrad, 1999).

We saw earlier that there are especially large differences between the samples' citation medians and other articles' citation medians for NEJM (Tables 4 and 5). Given the differences and the citation "boost" that NEJM articles receive when covered by the New York Times (Phillips et al., 1991), as well as bloggers linking to the New York Times (Peters et al., 2012) we decided to conduct a pilot study using the NEJM articles from the 2009 and 2010 samples. We "translated" first the medical terms to everyday language (e.g. sildenafil equals Viagra), then searched the New York Times' web site and the news agency Reuters' web site for stories covering the research published by NEJM. Twenty-one out of 26 articles in 2009 (81%) and 20 out of 38 articles in 2010 (53%) were covered by Reuters and/or the New York Times. Some articles were covered by more than one post, some posts covered more than one journal article and some news articles covered more than one journal article. In the case of an article covered by both NYT and Reuters, we calculated the date differences using the date of the first published news article. Unless otherwise specified, the differences are in favour of the news articles in the sense that they occurred earlier (Table 9). In 2009, the “other” category contains 2 studies which were covered by the media months before the NEJM articles were published (by press releases and so forth). In 2010, 2 of the three articles in the “other” category were published by the bloggers before they were reported on by either Reuters or NYT, and one study was covered by the media by press release rather than by its NEJM article.

Table 9: Time differences between blog posts and news articles reporting NEJM articles, for 2009 and 2010.

Time difference	2009	2010
Week or less	10	9
1-2 weeks	3	2
2-4 weeks	4	4
More than a month	1	2
More than two months	1	5
Other	2	3

Although many blog posts had been published shortly after the news articles reporting the NEJM article they cover, some had gaps between the news article and the blog post publication. One reason could be that some bloggers are slow to catch up with current events. Unlike professional news reporters, bloggers are usually not paid and are not pressured to be the first reporting exciting news. Another reason could be that while the bloggers use relatively current research in their posts, they do not use their blogs as an alternative for traditional news venues, but as platforms for their own agenda.

We have not searched other media outlets, but presumably some of them reported articles that were covered by the bloggers but were not reported either by Reuters or the Times, so the actual coverage percentage could be even higher. We cannot tell if the bloggers and the mass media are affected from one another, but the results show that in many cases the mass media and the bloggers have similar preferences.

Limitations

The sample time span is one of the study's main limitations. RB has only been active since 2008, and due to the slow accumulation of journal citations, we were able to use only articles from 2009-2010, because for them the citation window was sufficiently long. The year-long time frame for each sample meant, in theory, that the bloggers could have had some time to observe an article's popularity in the research community and cover it accordingly later at the same year.

The characteristics of RB were a source for some additional limitations. Being a self-selecting aggregator limited the sample to bloggers who chose to aggregate with it. The aggregator is focused on English-written blogs and is oriented towards certain disciplines, and the biological and medical sciences in particular. These limitations have much in common with the study's source of journal citations, Web of Science. The WoS database coverage focus on English-written peer-reviewed journals in the science, life science and medicine fields. Its coverage is not as robust in the social sciences and the humanities, which often publish monographs and books rather than

periodicals. Between the limitations of RB and WoS, the results might not be able to be generalized to all research blogs and scholarly disciplines.

Discussion and conclusions

This article described a potential source of alternative metrics, the research blog, using 2009 and 2010 data from the ResearchBlogging.org aggregator. We showed that the bloggers tend to prefer articles which turn out to be better cited than other articles from the same year and journal. In a one-sample binomial test, the overall proportion of blogged articles above each journal's population median was significantly higher than the expected probability of .5 in the two citation periods calculated for 2009 and 2010.

At the journal level, 7 out of 12 journals in 2009 (58%) and 13 out of 19 journals (68%) in 2010 had statistically significant results in terms of blogged articles attracting more citations. The higher number of significant results in 2010 might be due to the increasing number of blog posts and therefore articles in 2010 in comparison with 2009 ("wisdom of crowds"). The results show that for some, but not all journals, articles blogged in RB tend to subsequently receive more citations than other articles from the same journal. There are many different possible reasons for the cases of significant differences: bloggers pick better articles to write about and these attract more citations; bloggers sometimes write about articles that they use in their research and perhaps have already decided to cite when they blog about them; bloggers pick articles that are not necessarily better but are more interesting and get more read and hence more cited because of their interest; publicity from the mass media and/or blog post generates awareness of an article that leads to more citations. Whatever the reasons, it seems that, on balance, RB bloggers tend to pick articles that go on to become more highly cited than average. We note that most of the non-significant results came from the smaller journal samples. Had we chosen a higher threshold, for example 50 articles, only 2 of the 10 journals meeting this criterion in both years would have had non-significant results. The results' statistical significance remained identical for citation windows which did not include the year of publication (2010-2011 for 2009 and 2011-2012 for 2010), emphasizing that the bloggers tend to choose articles which will be better cited in later years.

The results validate those of Thelwall et al. (2013) regarding blogs and their association with citations for an unknown type of blogs and without estimating an effect size. Our study goes further than Thelwall et al.'s by associating blog coverage with *future* citations, estimates the effect size (the difference in median citations for blogged and non-blogged research), and also covers a precisely defined sample of blogs. The study in our current work offers particular insights about research blogs in which the bloggers use formal citation style. Thelwall et al.'s study of blogs only included articles mentioned in blogs, and the number of times the article was mentioned in blogs was crucial in the analysis. They compared each article to the two published nearest its publication date. Here, on the other hand, we compare between articles mentioned by blogs near their publication, irrespective of the number of times they are mentioned in the blogs, and articles in the same journal and the same publication year that were not picked up by bloggers. The study shows the differences between all the RB blog-covered articles and those which not covered in the relevant years and journals. Of the document types we included in the sample, review articles have been the type known to gather most citations. However, our hypothesis that the

bloggers' citing advantage might be due to higher use of review articles has not been proven. No evidence has been found for a connection between number of review articles in the sample in relation to the journals' population and a citation advantage or disadvantage.

Blog citations are worth pursuing as an altmetrics source, in part because of the effort put into them. Blog posts covering scholarly research which are written by humans and have real content (rather than advertisement or spam) take a great deal more time and thought than microblogging, bookmarking or downloading, even if the latter are not automated. The content of posts gives blog citations context, which is lacking in some of the other altmetrics sources. In conclusion, the bloggers showed themselves capable of choosing articles that, as a group, will become better cited than other articles in the same journal. Further research into the citing bloggers' motivations will allow a better understanding of their citations' functions and impact.

Acknowledgments

This work was funded by the EU FP7 ACUMEN project (Grant agreement: 266632).

References

Aksnes, D. W. (2003). Characteristics of highly cited papers. *Research Evaluation, 12*(3), 159-170.

Bar-Ilan, J. (2008). Informetrics at the beginning of the 21st century—A review. *Journal of Informetrics, 2*(1), 1-52.

Bar-Ilan, J. (2012a). JASIST 2001–2010. *Bulletin of the American Society for Information Science and Technology*, *38*(6), 24-28.

Bar-Ilan, J. (2012b). Journal report card. *Scientometrics*, *92*(2), 249-260.

Bar-Ilan, J., Haustein, S., Peters, I., Priem, J., Shema, H., & Terliesner, J. (2012). Beyond citations: Scholars' visibility on the social Web. In *Proceedings of the 17th International Conference on Science and Technology Indicators,* (Vol. 1, pp. 98–109), *Montréal, Canada*Bonetta, L. (2007). Scientists enter the blogosphere. *Cell, 129*(3), 443.

Conrad, P. (1999). Uses of expertise: sources, quotes, and voice in the reporting of genetics in the news. *Public Understanding of Science, 8*(4), 285-302.

Cronin, B. (1981). The need for a theory of citing. *Journal of Documentation* 37(1), 16-24.

F1000. (2012a). F1000 Faculty. Retrieved from http://f1000.com/prime/thefaculty.

F1000. (2012b). F1000 Faculty. Retrieved from http://f1000.com/prime/about/whatis/coverage.

Eysenbach, G. (2011). Can tweets predict citations? Metrics of social impact based on twitter and correlation with traditional metrics of scientific impact. *Journal of Medical Internet Research, 13*(4), e123.

Fausto, S., Machado, F. A., Bento, L. F., Iamarino., A., Nahas, T. R., & Munger, D. S. (2012). Research blogging: Indexing and registering the change in science 2.0. *PLoS ONE, 7*(12), e50109.

Glänzel, W., & Schoepflin, U. (1995). A bibliometric study on ageing and reception processes of scientific literature. *Journal of Information Science, 21*(1), 37-53.

Groth, P., & Gurney, T. (2010). Studying scientific discourse on the Web using bibliometrics: A chemistry blogging case study. Presented at the *WebSci10: Extending the Frontiers of Society On-Line.* Raleigh, NC. Retrieved from http://wiki.few.vu.nl/sms/images/9/9c/Websci10-FINAL-29-4-2010f.pdf

Kjellberg, S. (2010). I am a blogging researcher: Motivations for blogging in a scholarly context. *First Monday, 15*(8). Retrieved from http://firstmonday.org/ojs/index.php/fm/article/view/2962/2580

Kousha, K., Thelwall, M., & Rezaie, S. (2010). Using the web for research evaluation: The Integrated Online Impact indicator. *Journal of Informetrics, 4*(1), 124-135.

Li, X., & Thelwall, M. (2012). F1000, Mendeley and traditional bibliometric indicators. In *Proceedings of the 17th International Conference on Science and Technology Indicators, Montréal, Canada* (Vol. 2, pp. 451-551).

Li, X., Thelwall, M., & Giustini, D. (2012). Validating online reference managers for scholarly impact measurement. *Scientometrics, 91*(2), 461-471.

Library of Congress (January 2013). *Update on the Twitter archive at the Library of Congress* [White paper]. Retrieved from http://lib.trinity.edu/research/citing/apaelectronicsources.pdf

Mohammadi, E., & Thelwall, M. (2013a). Assessing non-standard article impact using F1000 labels. *Scientometrics*. Advance online publication. doi: 10.1007/s11192-013-0993-9

Mohammadi, E., & Thelwall, M. (2013b). Assessing the Mendeley readership of social sciences and humanities research. To be presented at the *14th International Society of Scientometrics and informetrics conference.* Vienna, Austria.

Perfecting peer review? [Editorial]. (2011, January 7). *Nature Medicine, 17*, p. 1-2. doi:10.1038/nm0111-1

Peters, I., Beutelspacher, L., Maghferat, P. and Terliesner, J. (2012). Scientific bloggers under the altmetric microscope. *Proceedings of the American Society for Information Science and Technology***,** *49*(1), 1–4. doi: 10.1002/meet.14504901305

Phillips, D. P., Kanter, E. J., Bednarczyk, B., & Tastad, P. L. (1991). Importance of the lay press in the transmission of medical knowledge to the scientific community. *New England Journal of Medicine., 325*(16), 1180-1183.

PLoS ONE. (n.d.). *Article-Level Metrics Information.* Retrieved from http://www.plosone.org/static/almInfo

Priem, J., Taraborelli, D., Groth, P., & Neylon, C. (2010). *Altmetrics: a manifesto.* Retrieved from http://altmetrics.org/manifesto/

Puschmann, C., & Mahrt, M. (2012). Scholarly blogging: A new form of publishing or science journalism 2.0? In A. Tokar, M. Beurskens, S. Keuneke, M. Mahrt, I. Peters, C. Puschmann, K. Weller, & T. van Treeck (eds.), *Science and the Internet* (pp. 171-182). Düsseldorf: Düsseldorf University Press.

Seglen, P. O. (1992). The skewness of science. *Journal of the American Society for Information Science 43*(9), 628-638.

Shema, H., Bar-Ilan, J., & Thelwall, M. (2012). Research blogs and the discussion of scholarly information. *PLoS ONE, 7*(5), e35869.

Shuai, X., Pepe, A., & Bollen, J. (2012). How the scientific community reacts to newly submitted preprints: Article downloads, twitter mentions, and citations. *PLoS ONE, 7*(11), e47523.

Smith, L. C. (1981). Citation analysis. *Library Trends*, *30*(1), 83–106.

Thelwall, M., Vaughan, L., & Björneborn, L. (2005). Webometrics. *Annual Review of Information Science and Technology*, *39*(1), 81-135.

Thelwall, M., Haustein, S., Larivière, V. & Sugimoto, C. (2013). Do altmetrics work? Twitter and ten other candidates. *PLoS ONE, 8*(5), e64841.

Tornes, A. (2013, February 21). *4 Things You Need To Know About Migrating to Version 1.1 of the Twitter API*. [Web log post]. Retrieved from http://blog.gnip.com/tag/version-1-1-of-the-twitter-api/

Vaughan, L., & Shaw, D. (2003). Bibliographic and web citations—What is the difference? *Journal of the American Society for Information Science and Technology*, *54*(14), 1313–1322.

Waltman, L., & Costas, R. (2013). F1000 recommendations as a new data source for research evaluation: A comparison with citations. *arXiv preprint arXiv:1303.3875*.

Wickre, K. (2013, March 21). *Celebrating #Twitter7* [Web log post]. Retrieved from http://blog.twitter.com/2013/03/celebrating-twitter7.html