Brief Communication

The effect of social media promotion on academic article uptake

Nicola Botting, Lucy Dipper, Katerina Hilari

Division of Language and Communication Sciences, School of Health Sciences,

City University London, London, UK

Corresponding Author:

Dr Katerina Hilari

Division of Language and Communication Sciences

City University London

Northampton Square

London EC1V 0HB

k.hilari@city.ac.uk

Abstract

Important emerging measures of academic impact are article download and citation rates. Yet little is known about the influences on these and ways in which academics might manage this approach to dissemination.

Three groups of papers by academics in a centre for speech-language-science (available through a university repository) were compared. The first group of Target Papers were blogged and the blogs were systematically tweeted. The second group were non-blogged papers that we carefully matched for author, topic and year of publication. The third group were papers by different staff members on a variety of topics.

Results suggest an effect of social media on download rate, which was limited not just to Target Papers but also generalised to Connected Control Papers. Unrelated Control Papers showed no increase over the same amount of time [main-effect of time ($F(1,27) = 55.6$, $p < 0.001$); significant group*time interaction ($F(2,27) = 7.9$, $p = 0.002$)]. Effect on citation rates was less clear but followed the same trend. The only predictor of 2015 citation rate was downloads after blogging ($r = 0.450$, $p = 0.012$).

These preliminary results suggest that promotion of academic articles via social media may enhance download and citation rate and this has implications for impact strategies.

Introduction

In recent years an increasing number of journals and academic institutions have been providing free online access to peer reviewed scholarly journal articles. Alongside this shift of journal articles to digital platforms, is the boom of social media utilisation in promoting research. Social media provide researchers with opportunities to increase the impact of their findings, by reaching broader audiences of other researchers, policy makers, journalists and the general public, and potentially increasing the citation of their work and application of their ideas.

There are two main ways for researchers to make their articles Open Access (OA): to publish in Open Access journals (Gold Open Access); or to publish in any peer-reviewed journal but make a pre-print of their article available free by archiving it in an online repository, typically their institution's online research archive (Green Open Access) (Harnad et al., 2004). A main limitation of top Gold Open Access journals is that they charge authors publication fees; additionally only a small percentage of high quality peer-reviewed journals (indexed by Thomson-Reuters-ISI) are Gold Open Access: about 5% (Gargouri, Larivière, Gingras, Carr, & Harnad, 2012). In contrast, the only constraint of Green Open Access is that authors self-archive. There is evidence that OA articles are more highly cited, with one study finding Open Access articles were twice as likely as non-OA articles to be cited (odds ratio = 2.1 [1.5–2.9]) in the first 4–10 months after publication, with the odds ratio increasing to 2.9 (1.5–5.5) 10–16 months after publication (Eysenbach, 2006).

Citations are an important indicator of research reach, reflecting the impact of a piece of research primarily in its field and on other researchers, in the long term. Yet, research can have a greater impact when it spreads faster and to a broader audience. Social media can achieve both these aims. In health care, social media have been used in various ways to allow researchers to communicate with health professionals, patients and the public, including: to spread timely information, e.g. the World Health Organization used Twitter (with about 12,000 followers) during the influenza A (H1N1) pandemic (McNab, 2009); and to inform the clinical practice of health professionals (Giustini, 2006). Researchers are therefore using social media like Twitter and blogs increasingly to communicate their findings and promote their research.

We are interested in the impact of social media on the dissemination of research findings, in the area of speech and language disorders and sciences. In health care / biomedical research only a small proportion of scientific articles is tweeted (< 10%) and correlations between tweets and citations are low, suggesting that Twitter impact metrics (like downloads) capture a different type of research impact to citations (Haustein,Peters, Sugimoto, Thelwall, & Larivière, 2014) although this has increased to 22% in recent years (Haustein, Costas, & Larivière, 2015). This is despite the fact that researchers are active users of social media platforms (Rowlands, Nicholas, Russell, Canty, & Watkinson, 2011; Van Noorden, 2014). Evidence from other disciplines suggests that promoting Green Open Access papers through Twitter and blogging can increase their impact, as demonstrated by higher downloads (Shuai,

Pepe, & Bollen, 2012; Terras, 2012) and higher early citations (Shuai et al., 2012). Though promising, these studies did not control for the quality of the papers or their level of interest that could have led to higher social media buzz and higher download rates and citation counts. To control for this, Terras went on to upload on an Open Access repository four similar papers from one project, but only tweeted about three of them. The three papers she tweeted about got 142, 209, and 297 downloads. The paper she did not tweet about got 12 downloads (Terras, 2012). This result supports the link between social media activity and downloads, however, it is still limited, as it compared papers of only one author and project; and only looked at downloads at one time point rather than before and after blogging/ tweeting. There is less research on the relationship between downloads and citations, but recent work by Schlögl, Gorraiz, Gumpenberger, Jack & Kraker (2013) suggests a strong correlation between the two ($r = 0.77$).

This study aimed to explore the link between promotion of research papers through social media (blogging and tweeting) and impact of the research, as evidenced by two different types of measures: downloads and citations. We aimed to do this in an experimental design, by looking at these measures before and after social media promotion and comparing them to two different controls (not promoted papers): Connected Control Papers, which were by the same authors (therefore of similar quality and of similar interest) and where we expected an impact spread as they were easily accessible in the online repository; and Unrelated Control Papers, which were by different authors and on different topics, where we did not expect an impact spread.

Method

Design

The 'cases' in this study were 30 empirical papers, all available on our institutional repository, City Research Online (CRO), that fell into three groups: The first were Target Papers (TP) which we blogged about and then tweeted these blog accounts. The second group were Connected Control Papers (CCP) carefully matched for year of publication, topic and author group. The third set consisted of Unrelated Control Papers (UCP) which were also in the area of speech and language science but had different topics and a different set of authors. See Table 1 for the specific article details. This design was adopted to examine impact-spread because once papers are accessed via the blogs, it is relatively easy to then search for the same authors or topics using CRO. Although 30 papers is a relatively small sample, the groups are sufficiently powerful to achieve statistically significant results.

Papers

The papers were all empirical articles published between 2005 and 2013 in peer-reviewed journals. They were all authored by staff members from our own division and were freely available via Green Open Access on the CRO archive.

Blogs and tweets

The project ran for two consecutive years (2012 and 2013) when two student researchers (one for each year of the study) wrote blogs containing the Target Papers (five papers promoted in each year). The blogs were written to be as naturalistic as possible, and the students were encouraged to build in personal experience around the papers which were mentioned with hyperlinks to the OA versions on CRO. A monthly blog was written for three consecutive months. Links to these blogs were also tweeted three times a day for the first three days and then once a week for the remainder of the month by the student; and freely retweeted by the university division account, staff members in the division and others.

Outcome measures

We were interested in two types of impact. The first was number of full-text downloads, comparing the four months before the blog activity, and the four months during and after the blogs. This information is stored by CRO routinely and could be extracted directly.

The second metric we took was regarding citations. This is more difficult to measure accurately for papers published over a range of years and in journals with different lags, however since the groups were matched we considered two different values: i) For the Target and Connected Control papers only we used Scopus to determine change in citations from the year before the blog began (2012 for phase 1 papers or 2013 for phase 2 papers) to the year after (2013 phase 1 or 2014 phase 2). This analysis would detect any immediate effects of blogging on citation; ii) For all three groups of papers we examined change in citations from 2012 to 2015 to give the most inclusive picture of change and effect. For 2015 data, which was collected in June of that year, citations were pro-rated by a multiple of 2 to account for the missing half of the year.

Results

The papers were well matched and there were no significant differences (all $p > 0.4$) between the three paper groups on year of publication or on the number of full-text downloads in the four months before the blog activity (TP = 4.6; CCP = 6.2; UCP = 4.7) or on citations in 2012 (i.e. before the experimental period began; TP = 4.3; CCP = 3.1; UCP = 3.7). See table 1 for details.

Table 1: Papers used in the evaluation

Key authors	Title	Topic	Year of publication
Target papers			
Botting	Emotional health in adolescents with and without a history of specific language impairment (SLI)	SLI; Quality of Life (QoL); adolescents	2008
Cruice	Conceptualising quality of life for older people with aphasia	Aphasia; adults QoL	2010
Dipper	What can co-speech gestures in aphasia tell us about the relationship between language and gesture?	Aphasia; adults; gesture	2011
Hilari	The impact of stroke: are people with aphasia different to those without?	Aphasia; adults; QoL	2011
Morgan	Children are just lingual': The development of phonology in British Sign Language (BSL)	Deafness; children; communication	2006
Herman	Early vocabulary development in deaf native signers	Deafness; children; communication	2010
Marshall (C)	Identifying specific language impairment in deaf children acquiring British Sign Language: implications for theory and practice.	Deafness; children; communication	2010
Botting	Cognitive abilities in children with specific language impairment: consideration of visuo-spatial skills	SLI; cognition; children	2005
Joffe	"A place where I can be me": a role for social and leisure provision to support young people	SLI; QoL; adolescents	2011
Hilari	Stroke Social Network Scale: development and psychometric evaluation of a new patient-reported measure	Aphasia; QoL; adults	2013
Connected control papers			
Botting	The role of language, social cognition, and social skill in the functional social outcomes of young adolescents with SLI	SLI; QoL; adolescents	2008
Cruice	Reporting on psychological well-being of older adults with chronic aphasia in the context of unaffected peers	Aphasia; adults; QoL	2011
Dipper	What can iconic gestures tell us about the language system? A case of conduction aphasia	Aphasia; adults; gesture	2011
Hilari	Psychological distress after stroke and aphasia: the first six months	Aphasia; adults; QoL	2010
Morgan	The first signs of language: Phonological development in British sign language	Deafness; children; communication	2007
Herman	Lexical organization in deaf children who use British Sign Language: Evidence from a semantic task	Deafness; children; communication	2013
Marshall (C)	The acquisition of Sign Language: The impact of phonetic complexity on phonology	Deafness; children; communication	2010
Botting	Associated reading skills in children with a history of Specific Language Impairment (SLI)	SLI; literacy; children	2006
Joffe	Social, emotional, and behavioural functioning of secondary school students with low academic & language performance	SLI; QoL; adolescents	2012
Hilari	Why do people lose their friends after a stroke?	Aphasia; QoL; adults	2011

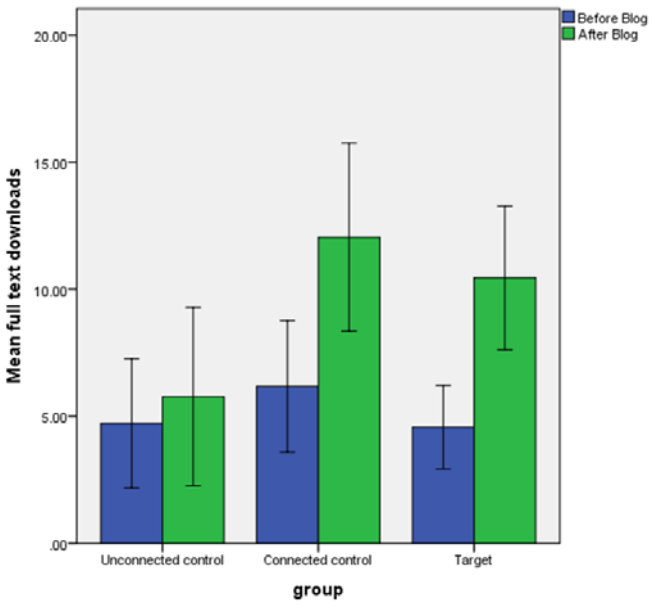
Unconnected control papers			
Chiat	Sentence imitation as a tool in identifying expressive morpho-syntactic difficulties in children with severe speech difficulties	Speech; children;	2010
Harding	Drinking speed using a valved Pat Saunders straw TM, wide bore straw and a narrow bore straw in school age children	Children; swallowing	2011
Knight	Transcribing nonsense words: The effect of numbers of voices and repetitions	Speech; adults	2010
Hasson	Discriminating disorder from difference using dynamic assessment with bilingual children	Bilingualism; children; dynamic assessment	2013
Marshall (J)	"Like déjà vu all over again": Patterns of perseveration in two people with jargon aphasia	Aphasia; adults; communication	2010
Pring	Delivering the Lee Silverman Voice Treatment (LSVT) by web camera: a feasibility study	Motor-speech; adults; intervention	2009
Kyle	Predictors of reading development in deaf children: A 3-year longitudinal study	Deafness; children; literacy	2011
Verhoeven	Intrinsic vowel F0, the size of vowel inventories and second language acquisition	Speech; adults	2011
Williams	The effects of bilingualism on speakers who stutter during late childhood	Bilingualism; Children; stuttering	2009
Kyle	Assessing the effectiveness of two theoretically motivated computer assisted reading interventions in the UK	Literacy; children; intervention	2013

Downloads

Independent paired t-tests revealed a significant rise from before blogging (July-Oct: $M = 4.6$ $SD = 2.3$) compared to after blogging (Nov-Feb: $M = 10.5$ $SD = 4.0$) for the Target Papers (t (9) = -4.5; p= 0.001) and for the Connected Control Papers (July-Oct: $M = 3.6$ $SD = 3.1$; Nov-Feb: $M = 12.1$ SD = 5.2; t (9) = -6.8, $p < 0.001$). No significant change was seen for the Unrelated Control Paper group from before (July-Oct: $M = 4.7$ $SD = 3.6$) to after (Nov-Feb: $M = 5.8$ $SD = 4.9$) blogging (t (9) = -1.5, $p = 0.17$). See Figure 1.

A mixed ANOVA was then run to confirm the difference in these patterns. There was a significant main effect of time (F (1,27) = 55.6, $p < 0.001$) and a significant interaction between group x time (F (2,27) = 7.9, $p = 0.002$). A borderline main effect of group was evident (F (2,27) =2.7, $p = 0.083$).

Fig 1: Mean full text downloads by group before and after blogging



Citations

The analyses described above were carried out to explore i) immediate change; ii) change from the years before the blogs to the present year. Figures 2a & 2b show the patterns of the different groups of papers regarding citations. Neither of the interactions were statistically significant (Figure 2a: F (1,18) = 1.03, p= 0.324; Figure 2b: F (2,27) = 0.39, p = 0.68). However there appear to be promising trends that support the full text-download data above. Namely that Unconnected Control Paper citations decline with time, but the Target and Connected Control Paper citation rate is maintained (Figure 2b).

Fig 2a: Mean citations on year of blogging and year after blogging for target and connected control papers

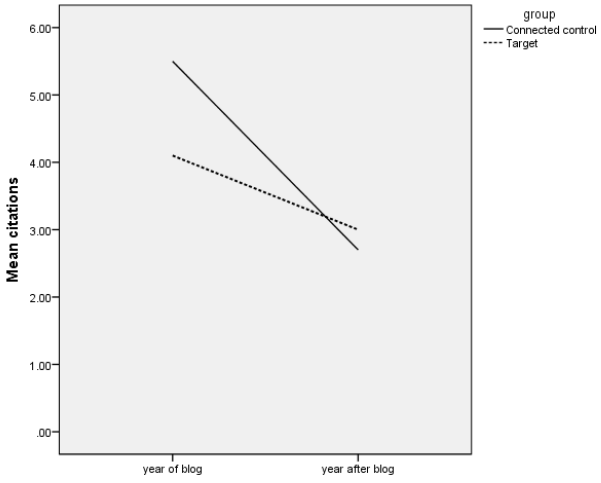
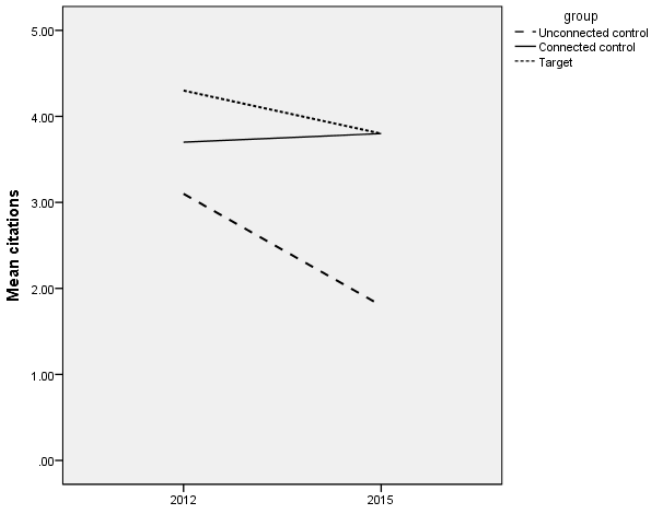


Fig 2b: Mean citations 2012 and (prorated) 2015 by group



Relationship between downloads and citations

With all groups combined, correlation analyses were run between citations 2012/2015 and downloads before and after blogging. Table 2 shows a clear relationship between 2012 citations and downloads before (r = 0.387, p = 0.035) and after blogging (r = 0.473, p = 0.008). However, the 2015 citation rate only correlated with downloads after blogging (r = 0.450, p = 0.012).

Table 2: Correlations between download and citation data

		Prorated citations 2015	Citations 2012	Average downloads before blog
Citations 2012	Pearson Correlation	.524*		
	Sig. (2-tailed)	.003		
Average downloads before blog	Pearson Correlation	.144	.387*	
	Sig. (2-tailed)	.449	.035	
Average downloads after blog	Pearson Correlation	.450*	.473*	.701*
	Sig. (2-tailed)	.012	.008	.000

Discussion

In aiming to examine the link between promotion of research papers through social media and impact of the research, this paper employed two different outcome measures (downloads and citations) and looked for a relationship between them. The experimental design, which looked at these measures before and after social media promotion and compared them to two different controls papers, allowed for tight control of the quality of the papers and their level of interest.

The significant rise in full-text downloads from before blogging and tweeting compared to after social media use clearly indicated the beneficial effects of social media on dissemination, and these were in line with findings from related studies focused on literature from other academic disciplines (Shuai et al., 2012; Terras, 2012). The comparison of Target Papers (those blogged and tweeted about) with both Connected (same authors and similar topic) and Unrelated Control Papers revealed a clear effect of impact spread from the target to the connected papers, but not to the Unrelated ones. The fact that papers were all OA in a University repository and therefore available for browsing is clearly implicated in this generalisation effect and supports the principle of OA in ensuring that a broad accessibility to science is achieved. However, though the papers were blogged by two students, the results are not attributable to student use alone since the Unconnected Control Papers were also by members of teaching staff in our own department and did not experience increased downloads.

By exploring citations rates, we were able to extend the picture emerging from the download findings to explore both immediate change and change from the years before the social media activity to the present year. Although there was no immediate effect on citation rates, the finding that both the Target and Connected Control Papers maintained their rates over the longer term (whilst the Unrelated Control Papers' rates declined) is interesting. Previous research in the healthcare and biomedical field had led us to expect the correlation between tweets and citations to be low in the short term (Haustein et al., 2014), but the longer term effect was less clear. Our results point toward an effect of social media activity on the endurance of citation rates for OA articles that remain visible in this way. Van Noorden (2014) has shown that researchers are increasingly using social media, and there now also appears to be a large number of health science professionals using this medium (although to our knowledge, no research is available on this group). Haustein and colleagues' work illustrates the growth in articles appearing on social media, reporting that the proportion of papers tweeted was 10% between 2010 and 2011, but 22% in 2012 (Haustein et al, 2014; Haustein et al, 2015).

The significant relationship between 2012 citations and downloads before and after social media promotion clearly indicated that there was a link between these two outcome measures. The fact that only downloads after blogging (and not before) correlated with 2015 citation rate suggested that blogging and tweeting were implicated in maintaining the citation rates. This supports work from other disciplines that also found correlations between downloads and subsequent citation rates (Schlögl et al, 2013).

The present study used a relatively small number of articles and a limited follow up timescale but was powerful enough to establish highly significant interaction between TP/CCP and UCP groups for downloads. It is possible that significant differences in citation rate were missed due to small numbers. Future research needs to replicate our findings to establish whether the effect is robust across both metrics. However, taken as a whole these results have supported the evidence in previous literature that a) social media promotion has a beneficial effect on the number of downloads, but that b) there is limited immediate effect on citation rates. This study adds to the evidence base through the use of both Connected and Unrelated Control Papers, which revealed an effect of impact spread from the papers promoted in social media to related papers. It also extends the evidence base by probing the relationship between two different types of impact and thereby showing a relationship between downloads and citation rates.

References

Harnad, S., Brody, T., Vallieres, F., Carr, L., Hitchcock, S., Gingras, Y, Oppenheim, C., Stamerjohanns, H., & Hilf, E. (2004) The green and the gold roads to Open Access. Nature Web Focus. http://www.nature.com/nature/focus/accessdebate/21.html

Gargouri, Y., Larivière, V., Gingras, Y., Carr, L., & Harnad, S. (2012). Green and gold open access percentages and growth, by discipline. arXiv preprint arXiv:1206.3664.

Eysenbach, G. (2006) Citation advantage of open access articles. PLoS Biol 4(5): e157. DOI: 10.1371/journal.pbio.0040157.

McNab, C. (2009) What social media offers to health professionals and citizens. http://www.who.int/bulletin/volumes/87/8/09-066712/en/ [Accessed 27/07/2015]

Giustini, D. (2006) How Web 2.0 is changing medicine. British Medical Journal, 333(7582),1283–4.

Haustein, S., Costas, R., & Larivière, V. (2015). Characterizing social media metrics of scholarly papers: The effect of document properties and collaboration patterns. PLoS ONE, 10(3), e0120495.

Haustein, S., Peters, I., Sugimoto, C. R., Thelwall, M., & Larivière, V. (2014). 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.

Rowlands, I., Nicholas, D., Russell, B., Canty, N., & Watkinson, A. (2011). Social media use in the research workflow. Learned Publishing, 24(3), 183–195.

Schlögl, C., Gorraiz, J., Gumpenberger, C., Jack, K., & Kraker, P. (2013). Download vs. citation vs. readership data: The case of an information systems journal. In *Proceedings of the 14th International Society of Scientometrics and Informetrics Conference* (Vol. 1, pp. 626-634).

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.

Terras, M. (2012). The Impact of Social Media on the Dissemination of Research: Results of an Experiment, Journal of Digital Humanities, 1(3)

Van Noorden, R. (2014). Online collaboration: Scientists and the social network. Nature, 512(7513), 126–129.