

Within-Journal Demonstrations of the Open-Access Impact Advantage : PLoS, Pipe-Dreams and Peccadillos

Stevan Harnad
Chaire de recherche du Canada
Centre de neuroscience de la cognition (CNC)
Institut des sciences cognitives (ISC/CSI)
Université du Québec à Montréal
Montréal, Québec, Canada H3C 3P8
<http://www.crsc.uqam.ca/>

and
Department of Electronics and Computer Science
University of Southampton
Highfield, Southampton
SO17 1BJ UNITED KINGDOM
<http://www.ecs.soton.ac.uk/~harnad/>

ABSTRACT: [Eysenbach's \(2006\)](#) study in *PloS Biology* on 1492 articles published during one 6-month period in one journal ([PNAS](#)) found that the Open Access (OA) articles were more cited than the non-OA ones. The online bibliography on the OA citation advantage <http://opcit.eprints.org/oacitation-biblio.html> records a number of prior within-journal comparisons that found exactly the same effect: freely available articles are read and cited more. Eysenbach's further finding that the OA advantage (in this particular 6-month, 3-option, 1-journal PloS/PNAS study) is greater for articles that have paid for OA publication than for those that have merely been self-archived will require replication on much larger samples as most of the prior evidence for the OA advantage comes from self-archived articles and is based on sample sizes four orders of magnitude larger for both the number of articles and the number of journals tested.

I applaud and welcome the results of the [Eysenbach \(2006\)](#) study on 1492 articles published during one 6-month period in one journal ([PNAS](#)), showing that the Open Access (OA) articles were more cited than the non-OA ones. I also agree fully that the findings are unlikely to have been an artifact of [PLoS](#)'s "strong and vested interest in publishing results that so obviously endorse our existence," nor of the fact that "the author of the article is also an editor of an open-access journal" (all quotes are from the [PLoS Biology](#) editorial by [MacCallum & Parthasarthy, 2006](#)).

However, I am less sure that PloS's and the author's vested interests are not behind statements (in both the accompanying editorial and the article itself) along the lines that: "solid evidence to support or refute... that papers freely available in a journal will be more often read and cited than those behind a subscription barrier... has been surprisingly hard to find." The online bibliography '*The effect of open access and downloads ('hits') on citation impact*' <http://opcit.eprints.org/oacitation-biblio.html> records a growing number of studies reporting precisely such evidence as of 2001, including studies based on

data from much larger samples of journals, disciplines and years than the PloS study on PNAS– and they all find exactly the same effect: freely available articles are read and cited more.

There can be disagreement about what evidence one counts as “solid,” but there can be little dispute that prior evidence derived from substantially [larger](#) and [broader-based](#) samples showing substantially the same outcome can hardly be described as “surprisingly hard to find.”

In fact, the only new knowledge from this small, journal-specific sample was (1) the welcome finding of how early the OA advantage can manifest itself, plus (2) some less clear findings about differences between first- and last-author OA practices, plus (3) a controversial finding that will most definitely need to be replicated on far larger samples in order to be credible: “The analysis revealed that self-archived articles are also cited less often than OA [sic] articles from the same journal.”

The latter (3) is a within-journal (one journal, PNAS) finding; the overwhelming majority of self-archived [OA \(sic\)](#) articles today (on which the prior large-sample OA citation advantage findings are based) do not appear in journals with a paid-OA option. Hence on the present evidence I have great difficulty in seeing this secondary advantage as any more than a paid-OA publisher’s pipe-dream at this point.

The following, however, is not a pipe-dream, but a peccadillo: “no other study has compared OA and non-OA articles from the same journal.” To be fair, this observation is hedged with “[a]s far as we are aware” (but the [OA-advantage bibliography](#) is surely public knowledge – or should be among advocates of public access to science) and the observation is further qualified with: “and [also] controlled for so many potentially confounding factors.”

But it has to be stated that of these “potentially confounding” variables -- “number of days since publication, number of authors, article type, country of the corresponding author, funding type, subject area, submission track (PNAS has three different ways that authors can submit a paper)... previous citation record of the first and last authors... [and] whether authors choosing the OA option in PNAS chose to do so for only their most important research (they didn't)” – many are peculiar to this particular short-interval, 3-option, single-journal PloS study. And several of them (country, subject, year) had already been analyzed in papers that had been published before this 2006 article and were not taken into account despite the fact that both their preprints and their postprints had been freely accessible since well before publication, and that at least one of them (Brody et al. 2005) had been explicitly drawn to the author’s attention based on a preprint draft well before the article was submitted to PloS.

[Brody et al. \(2005\)](#) had found that, alongside the OA citation advantage, more downloads in the first six months after publication are correlated with more citations 18 months later in physics; and [Hajjem et al. \(2005\)](#) had found higher citations for OA articles – *comparing always within the very same journal and year* -- for 1,307,038 articles published across 12 years (1992-2003) in 10 disciplines (Biology, Psychology, Sociology, Health, Political Science, Economics, Education, Law, Business, Management).

REFERENCES

Brody, T., Harnad, S. and Carr, L. (2005) Earlier Web Usage Statistics as Predictors of Later Citation Impact. Journal of the American Association for Information Science and Technology (JASIST) 56. <http://eprints.ecs.soton.ac.uk/10713>

Eysenbach, G, (2006) Citation Advantage of Open Access Articles. PLoS Biology 4(5): e157 DOI: 10.1371/journal.pbio.0040157 <http://dx.doi.org/10.1371/journal.pbio.0040157>

Hajjem, C., Harnad, S. and Gingras, Y. (2005) Ten-Year Cross-Disciplinary Comparison of the Growth of Open Access and How it Increases Research Citation Impact. IEEE Data Engineering Bulletin 28(4) pp. 39-47. <http://eprints.ecs.soton.ac.uk/11688/>

MacCallum, C.J., and Parthasarathy, H. (2006) Open Access Increases Citation Rate. PLoS Biology 4(5): <http://dx.doi.org/10.1371/journal.pbio.0040176>