Publish or Perish

- What Publish or Perish is for
- Caveat emptor
- Metrics
- Download and install Publish or Perish

If you are using Publish or Perish in one of your research articles, please refer to it in the following way:
Harzing, A.W. <year> Publish or Perish, version <number>, available at www.harzing.com/pop.htm

Are you applying for tenure, promotion or a new job? Do you want to include evidence of the impact of your research? Is your work cited in journals which are not ISI listed? Then you might want to try Publish or Perish, designed to help individual academics to present their case for research impact to its best advantage.

Publish or Perish is a software program that retrieves and analyzes academic citations. It uses Google Scholar to obtain the raw citations, then analyzes these and presents the following statistics:

- Total number of papers
- Total number of citations
- Average number of citations per paper
- Average number of citations per author
- Average number of papers per author
- Average number of citations per year
- Hirsch's h-index and related parameters
- Egghe's g-index
• The contemporary h-index
• The age-weighted citation rate
• Two variations of individual h-indices
• An analysis of the number of authors per paper.

The results are available on-screen and can also be copied to the Windows clipboard (for pasting into other applications) or saved to a variety of output formats (for future reference or further analysis). Publish or Perish includes a detailed help file with search tips and additional information about the citation metrics. Anne-Wil Harzing welcomes user feedback to help her improve the program.

If you are using Publish or Perish in one of your research articles, please refer to it in the following way:
Harzing, A.W. <year> Publish or Perish, version <number>, available at www.harzing.com/pop.htm

What Publish or Perish is for

Publish or Perish is designed to empower individual academics to present their case for research impact to its best advantage. We would be concerned if it would be used for academic staff evaluation purposes in a mechanistic way.

When using Publish or Perish for citation analyses, we would like to suggest the following general rule of thumb:

• If an academic shows good citation metrics, it is very likely that he or she has made a significant impact on the field.

However, the reverse is not necessarily true. If an academic shows weak citation metrics, this may be caused by a lack of impact on the field, but also by one or more of the following:

• Working in a small field (therefore generating fewer citations in total);
• Publishing in a language other than English (LOTE - effectively also restricting the citation field);
• Publishing mainly (in) books.

Although Google Scholar performs better than the Web of Science in this respect, it is still not very good in capturing LOTE articles and citations, or citations in books or book chapters. As a result, citation metrics in the Social Sciences and even more so in the Humanities will always be underestimated as in these disciplines publications in LOTE and books/book chapters are more likely than in the Sciences.

Caveat emptor

Although for reasons discussed in detail in Google Scholar - a new data source for citation analysis the use of Google Scholar generally provides a higher citation count than ISI, this might not be true for all fields of studies.

• The Social sciences, Arts and Humanities, and engineering in particular seem to benefit from Google Scholar's better coverage of (citations in) books, conference proceedings and a wider range of journals.
• The Natural and Health Sciences are generally well covered in ISI and hence Google Scholar might not always provide higher citation counts. In addition, for some disciplines in the Natural and Health Sciences Google Scholar's journal coverage seems to be patchy. This leads to citation counts in these areas that might even be much lower than those in ISI.
As a general rule of thumb, we would suggest that using Google Scholar might be most beneficial for three of the Google Scholar categories:

a. Business, Administration, Finance & Economics;
b. Engineering, Computer Science & Mathematics;
c. Social Sciences, Arts & Humanities.

Although broad comparative searches can be done for other disciplines, we would not encourage heavy reliance on Google Scholar for individual academics working in other areas without verifying results with either Scopus or WoS.

You might also want to read the following topics on this website for a discussion of the validity, assumptions, and limitations of the underlying sources and methods used by Publish or Perish:

- Reflections on the h-index
- Google Scholar - a new data source for citation analysis
- Reflections on norms for the h-index and related indices
- Google Scholar: the democratization of citation analysis?
- A Google Scholar h-Index for Journals

Metrics

In addition to the various simple statistics (number of papers, number of citations, and others), Publish or Perish calculates the following citation metrics (see Citation metrics for more details):

**Hirsch's h-index**

Proposed by J.E. Hirsch in his paper *An index to quantify an individual's scientific research output*, arXiv:physics/0508025 v5 29 Sep 2005. It aims to provide a robust single-number metric of an academic's impact, combining quality with quantity.

**Egghe's g-index**

Proposed by Leo Egghe in his paper *Theory and practice of the g-index*, *Scientometrics*, Vol. 69, No 1 (2006), pp. 131-152. It aims to improve on the h-index by giving more weight to highly-cited articles.

**Zhang's e-index**

Publish or Perish also calculates the e-index as proposed by Chun-Ting Zhang in his paper *The e-index, complementing the h-index for excess citations*, *PLoS ONE*, Vol 5, Issue 5 (May 2009), e5429. The e-index is the (square root) of the surplus of citations in the h-set beyond $h^2$, i.e., beyond the theoretical minimum required to obtain a h-index of 'h'. The aim of the e-index is to differentiate between scientists with similar h-indices but different citation patterns.

**Contemporary h-index**

Proposed by Antonis Sidiropoulos, Dimitrios Katsaros, and Yannis Manolopoulos in their paper *Generalized h-Index for disclosing latent facts in citation networks*, arXiv:cs.DL/0607066 v1 13 Jul 2006. It aims to improve on the h-index by giving more weight to recent articles, thus rewarding academics who maintain a steady level of activity.

**Age-weighted citation rate (AWCR) and AW-index**
The AWCR measures the average number of citations to an entire body of work, adjusted for the age of each individual paper. It was inspired by Bihui Jin’s note *The AR-index: complementing the h-index*, *ISSI Newsletter*, 2007, 3(1), p. 6. The Publish or Perish implementation differs from Jin’s definition in that we sum over *all* papers instead of only the h-core papers.

**Individual h-index (original)**

The Individual h-index was proposed by Pablo D. Batista, Monica G. Campitelli, Osame Kinouchi, and Alexandre S. Martinez in their paper *Is it possible to compare researchers with different scientific interests?*, *Scientometrics*, Vol 68, No. 1 (2006), pp. 179-189. It divides the standard h-index by the average number of authors in the articles that contribute to the h-index, in order to reduce the effects of co-authorship.

**Individual h-index (PoP variation)**

Publish or Perish also implements an alternative individual h-index that takes a different approach: instead of dividing the total h-index, it first normalizes the number of citations for each paper by dividing the number of citations by the number of authors for that paper, then calculates the h-index of the normalized citation counts. This approach is much more fine-grained than Batista et al.'s; we believe that it more accurately accounts for any co-authorship effects that might be present and that it is a better approximation of the per-author impact, which is what the original h-index set out to provide.

**Multi-authored h-index**

A further h-like index is due to Michael Schreiber and first described in his paper *To share the fame in a fair way, h_m modifies h for multi-authored manuscripts*, New Journal of Physics, Vol 10 (2008), 040201-1-8. Schreiber’s method uses fractional paper counts instead of reduced citation counts to account for shared authorship of papers, and then determines the multi-authored $h_m$ index based on the resulting effective rank of the papers using undiluted citation counts.

**Download and install Publish or Perish**

The Publish or Perish software is a Microsoft Windows application that can also be installed and used on Apple Mac OS X and GNU/Linux computers, with the aid of a suitable emulator such as Wine or CrossOver Mac.

Please refer to the following pages for detailed download and installation instructions:

- Publish or Perish on Microsoft Windows
- Publish or Perish on Apple Mac OS X
- Publish or Perish on GNU/Linux

**Publish or Perish running on Microsoft Windows 7**

**Publish or Perish running on Apple Mac OS X 10.6**

**Publish or Perish running on Ubuntu GNU/Linux 10.04**

Copyright © 1997-2011 Anne-Wil Harzing. All rights reserved.

$Id: pop.htm 714 2011-01-10 08:32:38Z dave $

http://www.harzing.com/pop.htm

3/31/2011