

How do you measure how good you are as a scientist?

Synthesis: Tsirogiannis Yannis, 2013

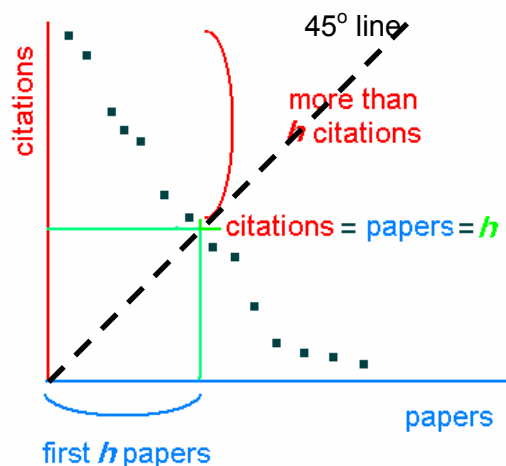
How would you compare the impacts of **different scientists in the same field** if you had to decide which would get a grant you were allocating? Measuring scientific performance is both more complicated, and important, than it might seem at first. Various methods for measurement and comparison have been proposed, none of which are perfect. And you should care about this because these metrics are increasingly being used by Funding Bodies and employers to allocate grants, jobs etc. So how could scientific performance be measured? The methods that might spring to mind at first are:

- Peer review. A good idea in principle, but it is subject to human nature so perceived performance will inevitably be affected by personal relationships. Also if a lesser known scientist publishes a ground-breaking publication he/she would likely get less recognition than if the same paper was published by a more eminent colleague.
- Number of articles published. A long publication list tends to look good on your CV, but the number of articles published gives no indication of the impact that your publications have had on the field. A small number of publications that have been well heeded by colleagues in the field (i.e. they are cited often) are better than a long list of publications cited poorly, or not at all.
- Average number of citations per article published. So if it's citations we are interested in, then surely the average number of citations per article is a better number to look at. Well, not really. The average could be skewed greatly by one highly-cited article so does not allow a good comparison of overall performance

The h-index

In 2005, Jorge E Hirsch of UCSD published this paper in PNAS in which he put forward the h-index as a metric for measuring and comparing overall scientific productivity of individual scientists.

The h-index has been quickly adopted as the metric of choice for many committees and bodies.



Conceptually, the h-index is pretty simple. You just plot papers versus the number of citations you (or someone else) have received and the h-index is the number of papers at which the 45degree line (citations=papers) intercepts the curve, as shown in the diagram. i.e. h = the number of papers that have received at least h citations.

For example if you have an h-index of 20, it means you have 20 papers with at least 20 citations. It also means that you are doing pretty well with your

science! Hirsch reckons that after 20 years of research, an h index of 20 is good, 40 is outstanding 60 is truly exceptional.

So we can ask ourselves, “Have I published one paper that’s been cited at least once?” If so, we’ve got an H-index of one and we can move on to the next question, “Have I published two papers that have each been cited at least twice?” If so, our score is 2 and we can continue to repeat this line of questioning until we can’t answer ‘yes’ anymore. Luckily, there’s no need to block off your weekend to try to figure out your stats- the computer’s got you covered (see below).

The advantage of the h-index is that it combines productivity (i.e. number of papers produced) and impact (number of citations) in a single number. So both productivity and impact are required for a high h index; neither a few highly cited papers or a long list of papers with only a handful of (or no!) citations will yield a high h index.

In that paper, Hirsch shows that successful scientists do indeed have a high h-index. A simple example is that Nobel prize winners in physics all have high h-indices (84% had an h of at least 30).

Example of how to calculate your own h-index from scratch

Calculation process

Step 1 is to create a list of all your publications and the number of times each publication has been cited. Data on the number of times cited can be taken from any database that shows/links to articles that cite a particular publication (e.g., Web of Science, Google Scholar, PsycInfo, SciFinder). Be sure to remove any duplicate citations between these sources – e.g., a specific journal article that cites your publication and is listed in both PsycInfo and Google Scholar should only be counted once.

Step 2 is to put that list in descending ORDER by the NUMBER OF TIMES CITED. The first publication on the list should have the largest number of citations and the paper(s) with the fewest numbers of citations at the bottom.

Step 3 is optional – but helpful – to number the items or use line numbers to easily identify which item in the list is #20 or #15 and so on.

To calculate the H-index – look down through the list to see the point at which the number of times a publication has been cited is equal to or larger than the line number for a given publication. For example, if there are 20 items that have 20 citations or more, but the 21st publication on the list has 19 or less citations, then your H-index is 20. Similarly, if the 8th publication on the list has 8 or more citations while the 9th item has 7 or less, then your h-index would be 8.

Example

You have 3 papers which are registered in internationally accepted data bases. One has 20 citations, one has 0 citations and one has 3 citations.

No	Paper	Citations
1	Review...	10
2	How....	3
3	Comparison of....	0

For No 2 (How...) the number of citations is equal or bigger (bigger in this case, 3) from the line number (2). So the h-index = 2

So it does not matter if you had published or not the 3rd paper which has 0 citations, or if your most cited paper had 3 or 10 citations.

Have in mind that in some cases the h-index for the last (for example) 5 years is asked for. In this case the above mentioned procedure uses only the papers that have been published the last 5 years.

Limitations of the h-index

But as attractive as it is to have a single number that measures scientific performance, it is only a rough indicator and should only be considered as such. Hirsch himself writes:

Obviously a single number can never give more than a rough approximation to an individual's multifaceted profile, and many other factors should be considered in combination in evaluating an individual. This and the fact that there can always be exceptions to rules should be kept in mind especially in life-changing decision such as the granting or denying of tenure.

Limitations of the h-index include:

- It does not take into account the number of authors on a paper. A scientist who is the sole author of a paper with 100 citations should be given more credit than one who is on a similarly cited paper with ten co-authors.
- It penalises early career scientists. Outstanding scientists with only a small number of publications cannot have a high h-index, even if all of those publications are ground-breaking and highly cited. This Wikipedia article gives a great illustration of this: "Had Albert Einstein died in early 1906, his h-index would be stuck at 4 or 5, despite his being widely acknowledged as one of the most important physicists, even considering only his publications to that date."
- Review articles have a greater impact on the h-index than original papers since they are generally cited more often.

Calculating the h-index

There are several online resources you can use to directly calculate a scientist's h-index. The most established are ISI Web of Knowledge, and Scopus, both of which require a subscription (probably via your institute), but there are free options too, one of which is Publish or Perish.

We recommend:

ISI Web of Knowledge: <http://wokinfo.com/>

and

Scopus: <http://www.scopus.com/>

If you check your own, or someone else's h-index with each of these databases, you might get a different value. This is because each uses a different database to count the total publications and citations. ISI and Scopus use their own databases, and Publish or Perish uses Google Scholar. Each database has different coverage, so will come up with different h-index values. e.g. ISI has good coverage of journal publications, but poor coverage of conferences, while Scopus covers conferences better, but has poor journal coverage pre 1992 (this was shown by a 2007 study, [click here for the abstract](#)).

References

Marnett A., 2010. H-Index: What It Is and How to Find Yours (<http://www.benchfly.com/blog/h-index-what-it-is-and-how-to-find-yours/>)

Oswald N., 2009. Does Your h-index Measure Up?. Careers (<http://writing.bitesizebio.com/articles/does-your-h-index-measure-up/>)

Unkown, How to create your own H-index from scratch (http://lgdata.s3-website-us-east-1.amazonaws.com/docs/58/218213/Creating_H-index_manually.pdf)

Wikipedia, h-index (<http://en.wikipedia.org/wiki/H-index>)

Κουτσογιάννης Δ., Καρλαύτης Μ., Σαπουντζάκης Ε., 2011. Εξελίξιριο (<http://itia.ntua.gr/dk-el/exelixirio/>)