A COLUMN ON THE HISTORY AND CULTURE OF GEOPHYSICS
AND SCIENCE IN GENERAL

Christopher L. Liner, University of Houston

“It has become appallingly obvious that our technology has exceeded our humanity.”
—Albert Einstein (1879–1955)

H-Score. Having spent almost exactly half of my working life as an academic, it has been my dubious pleasure to serve on many faculty search committees. Being academics, candidates submit voluminous CVs listing every paper, presentation, abstract, book, book chapter, contribution to a book chapter, etc. These documents fall into three important categories: peer reviewed, not peer reviewed, and the gray area in which one cannot tell if something is peer reviewed or not. Of course, the CV also contains information about awards, funding, teaching, graduate students, and we cannot forget reference letters. But it is the issue of peer-reviewed publications that carries the most weight, going to the heart of what it means to be a scientist.

Faced with many such CVs, the question naturally arises: How do we quantitatively rank candidates? This, in turn, is complicated by several things, including how to give relative rank to papers where our candidate was not the lead author. Do we count two second-author pubs equivalent to one first author? What about being the third author on three papers, again equivalent? What about being buried in a long list of authors? In our field of applied geophysics, it is rare to have more than, say, five authors on one paper. But in some areas of study there can be many more. A recent classic in particle physics had no less than 629 authors!

Lurking behind all of this is the amazing growth of scientific publications. It all began when the Royal Society was founded in England in 1660. Submitted papers were read at a Society meeting and the inevitable and boisterous feedback was the origin of peer review. By 1665, The Royal Society published the results of investigations in the Proceedings. Peer review in those early days often took the form of “letters to the editor” and could be brutal, so much so that no less a genius than Isaac Newton virtually stopped publishing for several years. In that first year of 1665, the complete world output of scientific papers would fit easily in a single printed book. It was a slow start, but exponential growth has resulted in over 10 million peer-reviewed science papers being published between 1996 and 2006. The current rate is well over one million per year.

Back to the question of publishing credentials, how are we to navigate this ocean of literature to determine who is the more accomplished candidate? A 2005 paper by J. E. Hirsch posted on the online Cornell physics archive (arXiv) suggested a method. It is based on impact and relevance rather than paper count. The key is citation information about an author’s work. This used to be hard to come by unless you were a librarian or on the editorial staff of a large scientific journal. But Google Scholar has brought this information within easy reach of anyone.

Hirsch associates an author with an index, h, defined as “the number of papers with a citation number higher or equal to h.” Let’s say, for example, an author has published 23 peer-reviewed papers, and we go out to Google Scholar to check him out. We find the most referenced paper has been cited 243 times, the second 122, the third 65, and so on. We continue down the list till we get to the 18th paper that has been cited 16 times. This author has an h-score of 17. The 18th paper does not count because it has been cited less than 18 times. Get it?

As of 2005, the highest known h-score in physics was 110, attributed to Ed Witten, a mathematician and string theorist at the Institute for Advanced Study in Princeton. I make it 158 as of August 2008. Stephen Hawking was a 62 (now 78). In 2008, it was pointed out in another arXiv paper that there is great variability between disciplines and a normalized measure is needed. But within a given field the original definition is fine.

Hirsch recognized that time was also an important factor. He concluded that in physics he could distinguish a successful scientist (h=20 after 20 years), an outstanding one (40/20), and a truly exceptional and unique individual (60/20). We also realize that great physicists of the past have an h-score that is fading into oblivion. Einstein and Witten, or Newton and Einstein, cannot be compared on h-scores. It is a thoroughly modern measure.

All of this got me thinking about the h-scores of authors in our field of applied geophysics. To keep the study within reasonable bounds, I selected people who have published in petroleum or applied seismology. The choice of who to include and exclude is completely subjective, although I think it is unlikely that there is a really big h-score out there that I missed—unless it was someone who generally works in another field (say particle physics) and just published a paper or two in ours. I included books but not patents, so again I may have passed over some prolific and important patent authors.

The result of my limited study (conducted in August 2008) is shown in Figure 1. The vertical axis is h-score as described above, and the horizontal axis is author arranged in increasing score. For lack of a better starting point, I began with the humbling exercise of finding my own h-score (which, I think, Ed Witten passed in the ninth grade). But, as you can see, I am in good company.

There are, of course, scores lower than mine, but we are interested in the high end. I made an effort to test every candidate who seemed to have a reasonable chance for the big prize. My final conclusion is that, in our field, Stuart
Crampin holds the highest h-score at 35, barely edging out Amos Nur. Plotting the actual citation numbers (Figure 2) for each of these outstanding scientists, we see it is very close indeed. Hirsch defined the h-score as the point where the citations equal the paper count. But if he had defined it as, say, the area under the curve, Nur would be our winner.

To further emphasize, however, that h-score is not everything, consider the top citation for Sven Treitel (h=13) is 290, that of Norm Bleistein (h=18) is 702, and Leon Thomsen (h=11) has a staggering 736.

Finally, let me say that h-scores have so far not come into any search committee deliberations that I have observed. Maybe it is too new, or a bit too sterile when deciding who will be a colleague for the next decade or so. Of course, we all want good scholarship and loads of research funding, but one also has to get along with students, peers, and administrators. In that regard, a kindergarten report card might be a useful indicator.


Corresponding author: chris.liner@gmail.com

Figure 1. H-scores of selected geophysicists arranged in order of increasing value. My apologies to any esteemed authors who are omitted.

Figure 2. Detail of citations Crampin and Nur, contenders for the top h-score slot.