Journal of Applied Research in Memory and Cognition 7 (2018) 209–216

Contents lists available at ScienceDirect

Journal of Applied Research in Memory and Cognition

journal homepage: www.elsevier.com/locate/jarmac

Evaluating Merit Among Scientists

Robert J. Sternberg*

Cornell University, United States

What makes some scientists distinguished and others hardly noticeable? How does one know who's who? In this article, I consider various means for judging the scholarly merit of a scientist and, further, some of the psychological issues involved when judgments are made regarding the distinction of scientists. I emphasize in particular four criteria: quality, productivity, visibility, and impact. The main ways of judging scientists on the basis of these criteria are peer judgments, quantity of publications, quantity while controlling for impact factors of journals, number of scholarly citations, the *h*-index, the *i10*-index, grants and contracts, editorships of important refereed journals, service on major grant panels, scholarly awards, election to elite academies and societies, and honorary doctorates. Underlying these indices and the distinguished scientific work about which they make judgments are creative, analytical, practical, wisdom-based, and ethics-based skills and attitudes. Although quantitative indices have limitations, they generally offer many advantages over qualitative evaluations that are susceptible to various kinds of biases and factors irrelevant to scientific merit.

General Audience Summary

Science today perhaps provides the best example of how knowledge can change our world for the better. But not all scientific articles, or scientists, are created equal. What are the characteristics of better versus not as strong scientists? And how do we distinguish among the scientists who have made more of a difference in their careers, and those who have made less of a difference? I consider these issues in this article. The general conclusion I draw is that there is no one "magic bullet" measure for evaluating scientists. What works best is to use a variety of convergent sources of information. Confidential letters of recommendation from esteemed scientists can be helpful, but of course scientists have their biases. Citation statistics, such as the number of times a scientist has been cited, also are helpful, but they tend to favor older investigators and those in fields with more scientists in them to do the citing. Awards can be helpful, but it always helps to have friends on the awards committees. The general point is that although no one measure is perfect, using multiple measures reduces error and enables evaluators to come to a reasonable conclusion about the merit of a scientist's contributions.

Keywords: Creativity, Analytical thinking, Common sense, Wisdom, Ethics

Author Note

This article is based in part on observations in Sternberg (2016a, 2016c, 2017a, 2017b, 2017c).

* Correspondence concerning this article should be addressed to Robert J. Sternberg, Cornell University, Department of Human Development, B44 MVR, Ithaca, NY 14853, United States. Contact: rjs487@cornell.edu

How can scientists (or anyone else) judge the scholarly merit of scientific work and the scientists who produce the work? In recent years, great progress has been made in providing sound criteria for making such judgments.

Judging Excellence in Scientific Work

Four criteria are especially important for assessing scientific merit: *quality*, *productivity*, *visibility*, and *impact* (Sternberg, 2016a). *Quality* is judged by peer-based assessments of the merit of a scientist's work. *Productivity* is judged by the volume of scholarly work that a scientist does. *Visibility* is judged by the extent to which a scientist and the scientist's work are familiar to his or her peers. And perhaps most importantly, *impact* typically is judged by the extent to which scholars cite the given scientist's work. How does one assess a scientist's contribution based upon these four criteria (and other criteria that may be relevant as well)?

Peer Referee Evaluations from Distinguished Referees

Traditionally, the most widely used basis for assessing the work of scientists has been written or sometimes oral evaluations by peers in the general field of the scientist's work. Peer evaluations may comment, usually qualitatively, on the nature of the work, the quality or quantity of work, the reasons the work is (or is not) important, and why the work deserves a particular recognition (such as promotion to tenure). Historically, letters have been the most important basis for judging scientific work, but they come with certain limitations. First, the evaluators may be biased. For example, evaluators may be more favorably disposed toward others whose work is more similar to their own, or at least toward scientists who share their paradigms or particular theoretical presuppositions (Sternberg, 1987). Although committees charged with promotions and making awards typically are looking for thought leaders (Antonakis et al., 2004), they sometimes end up with favorable recommendations from evaluators who do not feel threatened by the work or the scientist they are judging (Sternberg, 2003a; Sternberg & Lubart, 1992). That is, the judged scientists pose no professional risk to the evaluators.

Peer evaluation is different from the criteria that follow in that its results are usually confidential. Letters of evaluation are not made public. It is this confidentially that bequeaths upon the letters their unique value; but also it can sometimes lead an evaluator to make highly disparaging comments in the expectation that the target of the comments never will find out what the evaluator said.

Quantity of Publications

Quantity of publications is typically used to measure productivity, but is sometimes dicey because of the difficulty of comparing different kinds of contributions, for example, written books versus edited books versus refereed articles versus non-refereed articles versus talks, and so on. Moreover, the mere fact of publication often does not say much: scientists may find that almost any work they do can be published somewhere, as long as the scientists do not care too much about prestige or even the cost of publication.

Quantity of Publications Controlling for Impact Factors of Journals

In order to take into account that, especially today, some journals are journals in name only, evaluators may control for the impact factors of journals in which a particular scientist publishes. Impact factors are indices of the extent to which, on average, articles in given journals are cited in the scientific literature. Of course, the potential monkey wrench here is the term "on average." The fact that one publishes in a high-impact journal does not mean that one's own particular article(s) will be highly cited. Indeed, most articles in high-impact journals are not all that often cited. Often, it is the few most cited articles that largely determine a journal's impact factor. Furthermore, more prestigious journals may be afraid to publish work that is viewed as too highly speculative or different from the work of the rest of the field. That is, the journals often are conservative in what they publish. They may focus on articles being free of flaws rather than being stunning new contributions (which, because they are so new, may indeed have more flaws than more conservative articles).

An alternative to controlling for impact factors is to control for the selectivity of journals. In this case, one looks at rejection rates for submissions rather than at citation rates. But even highly selective journals sometimes publish "dogs," so that any criterion considering quality of journals is likely to be more useful averaged over many evaluations (i.e., for large groups of scholars being evaluated) rather than for the evaluation of a single scholar.

Number of Citations

Number of citations (which can be found in profiles on Google Scholar that are public) takes into account a scientist's contributions over the entirety of the scientist's career. The idea is that number of citations is prima facie evidence of impact, as citations mean that others are using the work. The number of citations is a gross index of impact but it needs to be understood in the context in which it occurs. For example, scientists who do research in popular or "hot" areas, whose work is especially controversial or even questionable, or who study problems that have widespread appeal, are at an advantage. To some extent, then, the measure may reward scientists who work in hot areas and penalize scientists who do important work in areas that are more remote from the center of the field's interest.

The h-index

The *h*-index, like number of citations, is shown on public Google Scholar profiles. The *h*-index is the number of publications, *h*, cited at least *h* times. Thus, if a scholar has an *h*-index of 20, then the scholar has 20 publications cited at least 20 times. The *h*-index is generally considered to be one of the most useful of the quantitative measures of impact, perhaps because it considers both quantity and quality of publication. A scientist