

## Deciphering impact factors

Soon after *Nature Neuroscience* was launched five years ago, we published an editorial cautioning against the misuse of impact factors (IFs)<sup>1</sup>. Now that the journal is well established, we decided to examine more closely what it means to have a high IF. Our findings illustrate the need for careful interpretation of this much-abused measurement.

We looked at the distribution of citations to individual papers in *Nature Neuroscience* (2002 IF = 14.857), and compared this to the distributions for neuroscience papers in *Nature* (overall IF = 30.432), and for samples of papers published in two larger journals, *Journal of Neuroscience* (IF = 8.045) and *Brain Research* (IF = 2.409), during the same period.

The most obvious feature of these distributions is that they are highly skewed (Fig. 1); in every case, the medians are lower than the means, reinforcing the point that a journal's IF (an arithmetic mean) is almost useless as a predictor of the likely citations to any particular paper in that journal<sup>2</sup>.

Although the distributions overlap, they are all significantly different from each other by a non-parametric test. Unsurprisingly, the peaks are systematically shifted in a direction consistent with the overall IF, such that, for example, the median paper in *Nature* would be at the 68<sup>th</sup> percentile for *Nature Neuroscience* and the 99<sup>th</sup> percentile for *Brain Research*.

What is most distinctive about the higher-impact journals is their long tails, corresponding to a relatively small number of papers that are exceptionally highly cited, and which therefore contribute disproportionately to the IF and, presumably, to the overall prestige of those journals. At the other end of the distribution, the lower-impact journals tend to publish more papers with few citations.

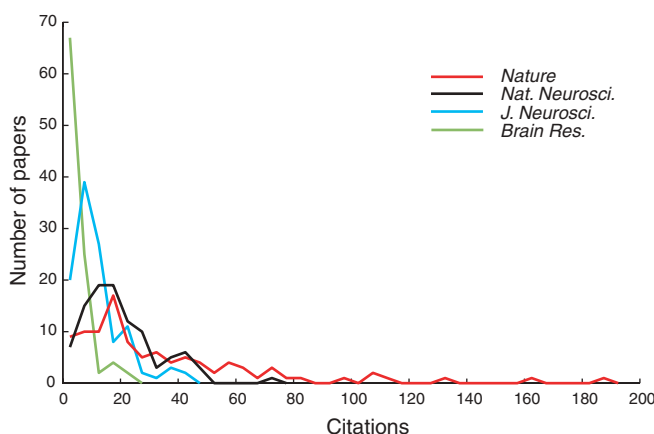
Numbers of citations are, of course, an imperfect measure of a paper's importance, not least because citation rates vary by subject;

Alzheimer's disease, for example, tends to attract more citations than cortical physiology. Some of the differences seen here may be related to subject balance, but this is unlikely to be the sole explanation, given that all four journals cover the whole of neuroscience. Given their overlapping scope and widely diverging IFs, these journals provide a reasonable test case for what might be called 'vertical stratification' of the literature. Based on these data, it seems that although there is some overlap between journals, those with the highest IF tend to be enriched for citation classics while publishing many fewer 'citation flops'.

One might argue that high citations have nothing to do with scientific quality, and are simply a consequence of the visibility conferred by publication in a top journal. This is an extreme view—it seems implausible that the journal selection process is entirely random—but it is also unlikely that citations are completely independent of where papers are published. Indeed, the main *raison d'être* of high-profile journals is to draw attention to important papers, and if this had no effect on a paper's likelihood of being cited, then one might question whether the journal system has any measurable effect or purpose. As editors, we would like to believe—although we cannot prove—that the truth lies somewhere between these two interpretations, and that the journal system serves to amplify small differences; in other words, we select good papers that would be well cited anyway, and these get an extra citation boost because they are noticed by a wider audience.

The effect of feedback loops should not be underestimated. According to a recent estimate based on propagation of citation errors, about 80% of all references are transcribed from other reference lists rather than from the original source article<sup>3</sup>. Given this finding, it is hard to escape the suspicion that many authors do not read every paper they cite, and instead tend to cite those papers that appear most often on other authors' references lists. Indeed, the distribution of citations within the literature as a whole is consistent with this model<sup>4</sup>. In a classic article 35 years ago<sup>5</sup>, the sociologist Robert Merton pointed out that science is not immune from the so-called Matthew effect, whereby the rich get richer ("For whosoever hath, to him shall be given..."). Given the current obsession with quantification and rankings, Merton's message is as timely as ever.

Properly interpreted, citation data can be a valuable tool for evaluating journals, papers, authors and perhaps even editors. But it is a blunt instrument at best, and when complex distributions are reduced to simple averages, then much of the usefulness is lost. Journal impact factors cannot be used to quantify the importance of individual papers or the credit due to their authors, and one of the minor mysteries of our time is why so many scientifically sophisticated people give so much credence to a procedure that is so obviously flawed.



**Figure 1** Cumulative citations to papers in four journals, based on data from the ISI Web of Science. For details, see **Supplementary Note** online.

1. *Nat. Neurosci.* **1**, 641–642 (1998).
2. Seglen, P.O. *Br. Med. J.* **314**, 498–502 (1997).
3. <http://arxiv.org/abs/cond-mat/0212043>
4. <http://arxiv.org/abs/cond-mat/0305150>
5. Merton, R.K. *Science* **159**, 56–63 (1968).