

Making impact

Strongly held views about impact factors continue to provide a steady and well-read stream of commentaries to many a journal. All too often they are attacked for providing an over-simplistic assessment of scientific merit. Indeed, the quantitative nature of the famed factor is certainly beguiling. Absolute numbers attached to articles, authors, journals or even whole fields allow for deceptively easy intra- or cross-field comparisons. Although the necessity of such assessments is undeniable, it becomes problematic when impact-factor-based rankings become significant criteria for faculty hiring, grant selection and indeed the standing of an individual in their community (see *Nature* 415, 726–729 (2003) for a more in-depth discussion). Ironically, the greater the divergence between fields being compared, the more impact factors tend to be used, while at the same time the actual value of an impact-factor-based comparison decreases. Often the more important the department or journal, the wider the range of subjects covered and the less informative the collective impact factor.

However, it is not reasonable to blame the ISI database (the “Web of Science”) for all of these problems. As Eugene Garfield, the initiator of the ‘Science Citation Index’, points out, the database was not set up with the central premise of providing a generic system of value judgement. Rather, it was designed to allow searches of literature links to provide informed in-roads into topics and identify colleagues new to the user. At the heart of the database is the quite rational assumption that a correlation exists between the volume of cross-referencing and the inherent ‘impact’ of a paper. The lack of differentiation between ‘positive’ and ‘negative’ referencing is not relevant to the definition of ‘impact’. This simple logic forms the basis for what remains an extremely powerful means to explore the increasingly intricate scientific literature. However, it also facilitated the more famous use of the database for impact factor analysis. Its use as an over-simplistic measure of scientific performance should be avoided, although comparisons within fields and judicious comparisons between fields can be informative indicators.

The 2002 journal impact factors have just been released, with NCB achieving a value of 20.7 (<http://npg.natureny.com/ncb/info/impact/>). These numbers represent one of the more notorious manifestations of the impact factor, not least because in some countries they are used in algorithms to calculate funding and salaries. Furthermore, inter-journal comparisons as presently practised are at best tenuous. Below we suggest several ways to make inter-journal (and by the same token

inter-faculty) comparisons more meaningful:

1. **Separate impact factors for reviews and primary research.** Reviews tend to be cited more frequently than primary research. The gap is likely to widen as total publication volume grows, because references to reviews are often used as a short-hand way of citing primary material. Comparisons between primary and review journals are not meaningful. In addition, they are also problematic for composite journals such as this, which feature both primary research papers and reviews. At worst, such flawed comparisons may affect editorial and authorship strategy towards maximizing impact factors. As a remedy, we suggest assigning journals (and authors) two independent impact factors, one for primary papers and one for reviews.
2. **Subject-based breakdowns.** Comparisons of impact factors within subject areas are most meaningful. While clearly not as neat as a single number, a field-by-field comparison between journals would be much more informative and again it would alleviate any potential pressure to adapt editorial criteria to boost the journal impact factor.
3. **Reclassification of journal subject areas.** ISI does classify journals by area. *Nature Cell Biology*, for example, is sensibly included in the cell biology section, but not in the equally relevant biochemistry and molecular biology section. Surprisingly, *Nature Medicine* is included in both groups, but *Genes & Development* is not included in either. These subject areas are too general to be meaningful and should be refined considerably. General journals cannot be classified at all and we suggest that subject breakdowns within each journal are more meaningful.

Finally, given the inherent caveats, why not dispense with the customary three decimal places cultivated for impact factors, which suggest an unrealistic level of accuracy.

The attraction of numerical ranking systems cannot be underestimated and all too often areas far more subjective than science — be it wine, restaurants or sport — are on the verge of succumbing entirely to their straightjacket. The citation index remains a powerful resource, but derived rankings must always be viewed within the relevant constraints. Although other databases, such as Crossref, will also allow bibliometric comparisons that increase confidence in such data, clearly it remains of paramount importance to let the scientific content of a paper speak for itself. If all this leaves a slightly bitter aftertaste, it may be worth rethinking percentile-based wine lists at the same time. □