

SOME MISUSES OF JOURNAL IMPACT FACTOR IN RESEARCH EVALUATION

Jesús Rey-Rocha¹, M. José Martín-Sempere², Jesús Martínez-Frías³ and Fernando López-Vera⁴

(^{1,2}Centre for Scientific Information and Documentation (CINDOC), Spanish Council for Scientific Research (CSIC), Madrid, Spain; ³Department of Geology, National Museum of Natural Sciences, Spanish Council for Scientific Research (CSIC), Madrid, Spain; ⁴Department of Agricultural Chemistry, Geology and Geochemistry, Faculty of Sciences, Universidad Autónoma de Madrid, Madrid, Spain)

Since it was first defined by Garfield in 1979, the use of the impact factor (IF) of scientific journals as an indication of research performance and quality has been discussed extensively. Impact factors of a selected set of approximately 5,000 of the world's leading journals in a broad range of disciplines are published yearly by the Institute for Scientific Information (ISI) of Philadelphia, in the Journal Citation Reports (JCR).

In many countries, IF is one of the criteria applied to evaluate not only the status of scientific journals, but also the publication output of scientists. In these evaluation exercises, IF is frequently considered as an indicator of research quality and scientific excellence. Sometimes, publication in "mainstream journals" or "impact journals" – defined as those with an IF, i.e. those covered by the JCR – is used as the only evaluation criterion, in such a way that scientific tribunals pay more attention to the IF of the journal than to the quality of the scientific contribution itself.

Many problems arise from the use of IF in this sense, which are broadly discussed in the extensive bibliography available (see, for example, Garfield, 1979b; Garfield and Welljams-Dorof, 1992; Reguant, 1995; Hansson, 1995; MacRoberts and MacRoberts, 1996; and Moed et al., 1999). Some of these problems are related to the geographical and linguistic origin of the journals and the papers that are published in them. Paris et al. (1998) consider how the frequency with which an article is cited is affected by its country of origin. Van Leeuwen et al. (2000) point out the language-bias problem caused by the non-English language journals within the Science Citation Index (SCI). On the other hand, Reguant and Casadellá (1994) allude to the lack of information derived from the absence, in valuation lists of current serials, of non-English-written journals with relevant information. Katz (2000) shows that a power law relationship exists between recognition and impact and the publishing size of a research community – nation, institution or group.

The misuse of the IF in relation to the language of publication in certain fields of work that are characterised by territoriality – for example, Earth Sciences (see also Kimley, 1994) – is a problem that is affecting some European, Latin-American and other non-English speaking countries. Rey-Rocha et al. (1999) showed how a biased evaluation of research performance and

quality, based on the analysis of research output published in journals covered by ISI's databases, is unfair to scientists working in those countries whose journals are poorly represented or not represented at all in the SCI. The study shows how this database does not properly reflect the total scientific output of Spanish researchers in Earth Sciences, as 69% of their articles are published in domestic journals, none of which are currently included in the SCI. This fact could be extended to many countries with a high scientific output but whose scientific journals are not significantly covered by the SCI. For instance, Italy, with 3.2% of world's scientific and technical articles in 1995-1997 (National Science Board, 2000, from ISI's data) had only 25 journals in the SCI in 1997 (ISI, 1998), or Spain, with 2.0% of articles and only 4 journals in the SCI.

Misuse of IF in research evaluation is also applicable to other scientific disciplines. As a demonstrative exercise, we have surveyed the Spanish scientific literature in three different scientific fields and disciplines: Earth Sciences, Physics and Neuropsychology. For this purpose, we searched the Spanish scientific output in both domestic and foreign journals, through both domestic and international databases, for the period 1990-1999. In addition to the most internationally prestigious databases in each area (Inspec for Physics, Medline for Medicine, and Georef for Earth Sciences), we searched the Spanish databases covering the literature published in Spanish journals on Medicine (Spanish Medical Index: IME) and Science and Technology (Spanish Index on Science and Technology: ICYT). The trend of Spanish scientists to publish in foreign journals is noticeably higher in disciplines of a more "general" or "international" interest than in those where the research has a high local or national interest. 92.1% of papers on Physics are published in foreign journals compared with 62.5% of Neuropsychology papers and only 28.1% of papers on Earth Sciences, the latter of which is a very geographically oriented discipline. The higher or lower presence of national journals of a particular field in the SCI can affect the ease with which researchers are able to publish their work published in impact journals. For instance, 52.8% of articles by Spanish Neuropsychologists were published in SCI journals, while 33.1% of them were published in 7 different Spanish journals.

In light of these data, would it follow that Spanish research papers in Physics are of better quality than those in Neuropsychology or Earth Sciences, or that Spanish Physicists perform better than their Neuropsychology and Earth Science colleagues, just because a higher proportion of the former publish in impact journals? If we were to reach this conclusion, we would be disregarding the idiosyncratic nature of some scientific fields.

Unfortunately, evaluation practices in Spain, as in many other countries, labour under the assumption that research papers published in impact journals are of better quality than those published in domestic, non-SCI journals. Although citation and impact measures, together with other bibliometric analyses, are only a part of the many possible indicators of research performance and quality, they have been over-emphasised in many evaluation processes in different countries. Other aspects as scientific quality, utility and societal quality of research, training of researchers, international scientific collaboration and collaboration with industry, etc., that have been increasingly considered in

countries with more advanced research evaluation systems, as the UK, the USA or Netherlands, remain under-considered.

REFERENCES:

- GARFIELD E. *Citation Indexing. Its theory and applications in Science, Technology and Humanities*. New York: Wiley, 1979a, p. 149.
- GARFIELD E. Is citation analysis a legitimate evaluation tool?. *Scientometrics*, 1, 4: 359-375, 1979b.
- GARFIELD E and WELLJAMS-DOROF A. Citation data: their use as quantitative indicators for science and technology evaluation and policy-making. *Science and Public Policy*, 19, 5: 321-327, 1992.
- HANSSON S. Impact factor as a misleading tool in evaluation of medical journals. *The Lancet*, 346: 906, 1995.
- ISI (INSTITUTE FOR SCIENTIFIC INFORMATION). *Science Citation Index Guide and List of Source Publications*. Philadelphia, USA: Institute for Scientific Information, 1998.
- KATZ JS. Scale-independent indicators and research evaluation. *Science and Public Policy*, 27, 1: 23-36, 2000.
- KIMLEY S. Limitations of Science Citation Index data in evaluating journals and scientists in Geology. *Proceedings of the 28th Meeting of the Geoscience Information Society*, October, 25-28, 1993, Boston, USA: Geoscience Information Society, 1994, pp. 23-31.
- MACROBERTS MH and MACROBERTS BR. Problems of citation analysis. *Scientometrics*, 1, 4: 359-375, 1996.
- MOED HF, VAN LEEUWEN ThN, and REEDIJK J. Towards appropriate indicators of journal impact. *Proceedings of the Seventh Conference of the International Society for Scientometrics and Informetrics*. Colima, Mexico, 1999, pp. 341-354.
- NATIONAL SCIENCE BOARD. *Science and Engineering Indicators 2000*. Arlington, VA: National Science Foundation, 2000.
- PARIS G, DE LEO G, MENOZZI P, and GATTO M. Region-based citation bias in science. *Nature*, 396: 210, 1998.
- REGUANT S. Reflections on scientific evaluation. Some comments on the 8th Conference of the International Federation of Science Editors. *Microbiología SEM*, 11: 499-502, 1995.
- REGUANT S and CASADELLA' J. English as lingua franca in geological scientific publications. A bibliometric analysis. *Scientometrics*, 29, 3: 335-351, 1994.
- REY-ROCHA J, MARTIN-SEMPERE MJ, LOPEZ-VERA F, and MARTINEZ-FRIAS J. English vs. Spanish in science evaluation. *Nature*, 397: 14, 1999.
- VAN LEEUWEN ThN, MOED HF, TIJSSEN RJW, VISSER MS, and VAN RAAN AFJ. First evidence of serious language-bias in the use of citation analysis for the evaluation of national science systems. *Research Evaluation*, 8, 2: 155-156, 2000.