

## Pesquisa FAPESP - June 2006 - Edição 124

Policies > Evaluation

### To publish is not everything

An article proposes a qualitative method for evaluating researcher's performances

What is the fairest way to evaluate a researcher's production? The debate that has been exciting the academic world for some time received new fuel in an article signed for by Edgar Dutra Zanotto, a professor at the Materials Engineering Department of the Federal University of São Carlos (UFSCar). Based on the ten years during which he has been an assistant coordinator to FAPESP's scientific board – around 48,000 projects had been screened by him during this period –, Zanotto is proposing a manner of classifying a researcher's production not just through the traditional quantitative parameters (articles published in scientific magazines and citations of these articles in other studies), but also via quality criteria. Zanotto's article was accepted for publication in the magazine *Scientometrics*, a reference in scienceometry, the discipline that looks to generate information in order to stimulate the overcoming of challenges in science. Entitled, "The scientists' pyramid", it presents a list of situations and qualities capable of placing a researcher within four proposed categories.

The main one of them, the top of the pyramid, presents demands within which few Brazilian researchers would be placed. Two of them are sacred, namely the publication of scientific papers in the most important scientific magazines, such as *Science*, *Nature*, *Cell*, *New England Journal of Medicine* and *Physical Review Letters*, and the index in the case of thousands of citations of their articles in publications in the Thomson Scientific (Thomson ISI) database. The others are qualitative, such as having received the most important awards in a particular field of knowledge; to have worked at research centers or laboratories of an international level; to be endowed with abundant resources; to belong to prestigious scientific academies and to the editorial body of important publications; to have been invited to give lectures and to have commanded round tables at congresses or international symposiums; to have been cited in text books of one's specialty and in the media. In all there are 11 parameters. "In order to belong to this top category, it would be necessary to comply with at least nine or ten of the parameters", explains Zanotto, who anticipates some discomfort should his methodology gain favor: "In Brazil, I believe that no more than a dozen researchers, those few who are in the waiting room for a Nobel Prize, would find themselves at the top of the pyramid", he suggests.

**Steps on the scale** – In the other three categories, baptized as classes A, B and C, the criteria are similar, but with decreasing rigor. Taking as an example the parameter of citations, those for class A must have at least 500 articles cited in the Thomson ISI database, those for class B at least one hundred and those for class C only a few or none at all. Those in class A must work in internationally known research centers or laboratories, those in B at reasonably well known

research centers and those in class C in day to day research centers.

But what would be the purpose of a classification of this type? Zanotto believes that the format would help researchers to better perceive their strong and weak points and to attempt to straighten out steps on the pyramid scale. He also believes that it would be useful for development agencies when directing their resources. "Each agency would establish its criteria. One of them could be, for example, that only a researcher at the top or in class A is appropriate to lead a project of size, such as FAPESP's thematic projects", he says. Professor Zanotto lived in the United States during 2005, privileged with a sabbatical period, and during that time reflected upon his experience in the search for a more trustworthy way to classify researchers' production. His article is a result of that reflection. The pyramid criteria refer to categories that he learned to identify during his work as an evaluator.

The concern in searching for a new classification of researchers' performances comes fed from other proposals. One of them is the so called 'h index', a proposal from the physicist Jorge Hirsch, from the University of California, in San Diego. The 'h index' is defined as the number "h" of scientific papers that have at least the number "h" citations of each of them. Putting it an easier way: a researcher with an 'h index' of 30 is the researcher who has published 30 scientific papers, and each of them has received at least 30 citations in other publications. This consideration excludes less cited works and also avoids distortions (if the citations were concentrated on a single paper by the author, this does not contaminate the general count). Thus, it gives the average size and impact of a researcher's academic production.

In his study, Zanotto suggests that, although it corrects distortions, the 'h index' is a long way from perfection. He applied the index to the production of four renowned Brazilian physicists (whose names are not revealed so as not to personalize the debate). All of the four had a similar 'h index', between 10 and 12. The most productive of the quartet, nevertheless, had displayed a citation index for scientific papers that was four times greater than the last on the list. "It's urgent to find a manner more holistic to evaluate quality, talent and a scientist's reputation", says the professor from UFSCar.

However, Zanotto's proposal also has its limitations. The main one, admitted to by the author himself, is its applicability in certain fields of knowledge, but not in others. In human sciences, the criteria of counting articles and citations would have to be substituted by another, since the reality is different from engineers, medical doctors and the exact sciences, and little is published in international magazines. "But it would be possible to make adaptations. The essential thing is to maintain the spirit of demonstrating the prestige that a researcher has in his field. Nobody manages to fool his or her peers", says Zanotto.

For the specialist in scienceometry Rogério Meneghini, Scielo's scientific coordinator, a retired professor from the Chemistry Institute of the University of São Paulo and also an ex-assistant coordinator to FAPESP's scientific board, the value of professor Zanotto's study lies in bringing up the discussion about the best evaluation criteria and also in demonstrating the importance of peer evaluation. "In practice this is what really takes place when a contest occurs in which we have forty candidates and only two or three positions", suggests Meneghini. "All of these qualitative criteria are contemplated by the evaluators and they produce a more just choice." Meneghini, nevertheless, believes it to be difficult to transpose such criteria to a systematic and en mass evaluation. "There're difficulties in making objective comparisons. Who's going to define if a chair in a determined scientific academic institute is valued so much on a researcher's curriculum when compared with another question that is mentioned in the

biography of another researcher?”, questions Meneghini. “There’s always a margin for some type of distortion. The quantity of funding that a researcher receives at times has more to do with the priority of a government than truly that of his or her performance.”

Zanotto defends his method. “It’s possible to have a distortion in one or other of the parameters, but not in the whole group”, he says. The researcher made the point of advising that he is not creating the methodology for his own benefit. “I wouldn’t place myself at the top of the pyramid”, he says. It would appear to be a joke, but there have been researchers who have suggested sliced up evaluation criteria in order to sparkle their very own curriculum. The famous Russian physicist Lev Landau (1908-1968) at a certain point proposed a logarithm for ranking researchers within their field of knowledge. Albert Einstein received a modest quotation: 0.5. However, Lev Landau appeared at the top with a 2.5.