

According to *INTERCIENCIA*'s policy of carrying out free discussions of ideas and opinions, our pages are open to individuals and institutions that wish to express their points of view, although they do not necessarily coincide with those published in the journal.

The Editor

/CARTAS AL DIRECTOR/
/LETTERS TO THE EDITOR/
/CARTAS AO DIRETOR/

CONTRASTING VIEWS ON MEXICO'S NATIONAL SYSTEM OF RESEARCHERS

Dear Sir,

As current/ex-members of Mexico's "National System of Researchers" (SNI) evaluation committee in Area II (Biology & Chemistry), we can state that Ricker and coworkers' assertions regarding the researcher evaluation process are factually flawed. To illustrate our concern, we provide six examples of these flaws.

1. *The SNI does not evaluate all Mexican scientists.* Ricker *et al.* state that "a centralized federal agency uses such bibliometric statistics for evaluating the performance of all Mexican scientists". In fact, the SNI only evaluates those scientists who are interested in obtaining a fellowship based on their performance and are willing to be subjected to a strict peer review process based on publicly stated rules. Only a small proportion of the Mexican scientific community belongs to the SNI.

2. *Evaluations are not performed hastily.* The authors state that each evaluator has to assess the performance of each researcher quickly and is under high pressure to automate the evaluation process. This is not true. In the area of Biology & Chemistry, in 2009 each evaluator reviewed ~200 researchers over approximately 4.5 months, representing an average of one or two CV's a day during this period. The committee met weekly and each case was presented by two evaluators and discussed among all committee members. This can be a brief discussion for researchers who clearly meet the criteria for continuance in the SNI, or may be a protracted discussion in cases where the merits of the researcher's performance are less clear cut. Committee members dedicate hundreds



of hours each year to performing these evaluations with great attention to detail. The process is neither high-pressured nor automated, although it is performed using an online database.

3. *Journal impact factors are not used to determine researcher performance.* The authors argue that the SNI uses journal impact factors as a quantitative measure of the value of a researcher's work. It is a truism to say that different areas of science have different trends in publishing and citation, yet the authors explain this at great length with a number of examples. The SNI is divided into seven areas based on discipline, each with its own specialized committee, and an additional expert committee makes recommendations concerning technological advances across all areas. The evaluators are selected from the highest ranked SNI members and are fully aware of differences in journal im-

pact factors, types of academic products, etc. in the different disciplines covered by each area. It is clear that similar evaluation criteria are not suitable for researchers working in areas as distinct as diatom taxonomy and HIV antiviral therapy, which is precisely why evaluations are tailored according to the academic products relevant to each discipline.

4. *Evaluations are not competitive or largely based on bibliometric indicators.* Ricker *et al.* state that SNI evaluation committees use bibliometric indicators, principally citations, to determine a scientist's performance and that this is detrimental to small guilds of scientists working on little-studied organisms, because their work will attract few citations and will not be able to compete statistically with scientists from large research communities, such as biomedicine. Both these beliefs are fallacious. The relevant criteria are easily available on the SNI website (www.conacyt.mx/SNI/Criterios/2009/CRITERIOS-INTERNOS-AREA-II.pdf) and include a diversity of academic products appropriate to each discipline, one of which is the citation record. The performance of scientists from each discipline is evaluated according to each individual's merits and compared with the average performance of researchers in that discipline; scientists do not compete with one another at any stage. Incidentally, we notice that *Systematic Entomology* is one of the most highly ranked entomological journals, which abrogates the authors' argument that descriptive biology, such as insect taxonomy, habitually attracts few citations.

5. *Scientists are evaluated by experts from the same or closely related disciplines.* The authors suggest that the performance

of taxonomists is evaluated by chemists. This is not true; as far as possible the cases of researchers from each discipline are assigned to evaluators from the same or a closely related discipline. Reflecting this diversity, in 2009 the 14 member committee in the area of Biology & Chemistry comprised experts in taxonomy, physiology, cell physiology, embryology, genetics, molecular biology, immunology, ecology, environmental science and biochemistry, as well as materials science, analytical chemistry and two organic chemists. If a researcher disagrees with the committee's judgment, that person can appeal to an "Appeals Committee" comprising a new body of members who have never seen the case before. Every possible effort is made to be fair and objective.

6. *Evaluations focus on the quality, consistency and coherence of research activities, international recognition and leadership.* The authors state that the work of a scientist should focus on quality over quantity, and assessing the value of a scientist's performance "must be broadened

beyond the number of published articles and indicators of citation frequencies", and that "it is nonsensical to search for a single criterion for the evaluation of all scientists". This is precisely the posture of the SNI. Specifically, evaluations are based on indexed papers and citations, prestigious monographs of flora and fauna, books and book chapters by established editorial houses, the involvement in teaching and training students, technology transfer and outreach activities, authorship of patents, dissemination of science to a wider audience, national and international scholarly awards, and participation in the academic development of scholarly institutions. These are precisely the criteria that Ricker *et al.* propose for their "rule-based peer review system" which effectively renders their proposal as one supporting the *status quo* in the SNI evaluation process. However, unlike Ricker *et al.*, the SNI promotes an integrated approach to scientific research that includes multiple facets of student training, teaching, outreach and written products with emphasis on consistency, coherence and international recognition in

lines of research and evidence that each scientist has contributed significantly to the products being reported in the period of each evaluation and during the course of their academic career.

In conclusion, the authors' statements contain serious factual errors and misinterpretations of evaluation procedures that verge on being defamatory. As such, we are deeply concerned that the spurious arguments used by Ricker *et al.* may fuel the debate over the possible abolishment of the SNI, because such an act would have catastrophic consequences for the scientific community in Mexico. Without question, the SNI has had a marked positive influence on the quality of Mexican science. To entice the new generation of Mexican scientists to seek membership of this prestigious institution we need to promote ideas that strengthen, not weaken, the SNI.

TREVOR WILLIAMS
trevor.williams@inecol.edu.mx
AND MARTÍN ALUJA
martin.aluja@inecol.edu.mx

ANSWER BY RICKER, HERNÁNDEZ AND DALY

Evaluation in academia can become an emotion-driven topic. We have no intentions to diminish the merits of scientists like Drs. Trevor Williams and Martín Aluja. Rather, we argue that performance evaluation needs to be inclusive of many forms of conducting science, and evaluating institutions need to be explicit and consistent about how they accept, score, and compare different academic products and activities. We do agree that Mexico's *Sistema Nacional de Investigadores* (SNI) "has had a marked positive influence on the quality of Mexican science" since its foundation in 1984, principally by stimulating Mexican scientists to attain and expand the international research frontier in their respective fields. That achievement should not be lost, of course, but a single centralized evaluation system like the SNI needs to remain open to different forms of working in different fields and different institutions, both in basic and applied science. If not, the system as it stands provides disincentives for scientists to conduct research in non-mainstream fields, or to pursue innovative research that will not necessarily result in articles in ISI-registered publications. Ultimately, an adequate approach to evaluation in public institutions will have to take social priorities into consideration: Does the SNI as the principal federal reward system for scientists in Mexico apply the correct filters from society's viewpoint? It is only natural that

those who have been highly successful in the SNI (some of whom become evaluators) would tend to answer yes, and those who have had difficulties in the SNI (and cannot be evaluators) would tend to answer no. The requirement of reaching the highest SNI level before becoming an evaluator causes a tendency for the system to perpetuate itself and resist reform. Currently there is a situation in which the SNI's evaluation process is potentially imbalanced and unfair, at least in practice, and the criteria stated in the SNI's regulation are not applied in a satisfactory way. If one looks at the presidential decree that created the SNI, it asks in its first article to stimulate the participation of scientists in Mexico's development, according to Mexico's National Development Plan. One may ask if that is achieved with the SNI's current overwhelming emphasis on ISI-registered publications (ISI: Thomson Reuters' Institute for Scientific Information). Below, we respond to the six points of criticism raised by Williams and Aluja:

1- It is true that the SNI evaluates only Mexican scientists who submit an application for inclusion in the system (we state this on p. 831). It is not true, however, that the result is merely a fellowship for a small proportion of scientists who happen to favor being evaluated. In fact, membership in the SNI not only includes a monthly monetary "stimulus award" from the federal govern-

ment, it is usually also a prerequisite for being hired or promoted at Mexican universities or for receiving governmental research grants from CONACyT, Mexico's federal science agency.

2- Williams and Aluja write that "In the area of Biology & Chemistry, in 2009 each evaluator reviewed ~200 researchers over approximately 4.5 months, an average of 1-2 CVs a day." Expressing it that way, as a daily average of applications spread out over several months, does not reduce what is a large workload that may not allow detailed reading of all the appended material produced over a period of three years (or more) by the applicant. Evaluation is not a full-time activity, and evaluators themselves are expected to continue their own research activities.

3- Journal impact factors are definitely used to determine researcher performance in the SNI. Williams and Aluja cite the SNI's web page of the internal criteria for 2009 in the area of biology and chemistry. If one reads the text, the criteria clearly state that articles shall be published in journals that are indexed with impact factors (points 2.1 and 2.2). In practice a researcher is rejected from the SNI if he does not have at least three ISI-registered publications over a period of three years, and the number of ISI-registered publications is the key element

for determining the SNI level reached. For those with fewer than three ISI-registered publications, the SNI does not accept the possible equivalence of one or two articles that are highly innovative or represent the results of many years of research, and it does not factor in thesis supervisions or other products such as book chapters or technical reports. When applying to the SNI, one has to enter online the journal name of each published article, and subsequently the online system reports automatically the corresponding bibliometric indicators, including the impact factor.

4- Williams and Aluja state that the SNI evaluations are not largely based on bibliometric indicators, but the SNI's insistence on ISI-registered journals with high impact factors contradicts this statement. Another bibliometric indicator taken into account by the SNI is the number of citations of each article published by the applicants. The use of bibliometric indicators is a worldwide phenomenon that has also penetrated the SNI, the aim being to maintain objectivity in the process of centralized evaluation of large numbers of applications. Unfortunately, fields like descriptive taxonomy fare badly in that system, and even *Systematic*

Entomology has an impact factor of only 1.8 for 2008, despite its strong emphasis on molecular and phylogenetic analysis.

5- Contrary to the opinion of Williams and Aluja, we believe that there is an urgent need for reform in the SNI to ensure that the expertise of evaluators and the areas of study of the applicants are compatible. The SNI divides all scientific activities into seven areas which, despite the SNI's insistence on publishing in journals with high impact factor, do not correspond to bibliometrically defined fields. Thomson Reuters assigns 11,196 journals to 22 so-called broad fields in a "master journal list" (<http://sciencewatch.com/about/met/journallist/>). What in the SNI is a single scientific area for evaluating all applicants in biology and chemistry, is divided into eight broad fields in the master journal list: Biology & Biochemistry, Chemistry, Environment/Ecology, Microbiology, Molecular Biology & Genetics, Multidisciplinary (e.g., *Nature* and *Science*), Neuroscience & Behavior, and Plant & Animal Science. The SNI assigns only 14 evaluators for covering these eight broad fields: Compare this approach with the evaluation process for journal articles, in which peers are selected specifically

for each case from a worldwide pool of scientists!

6- It is true that the SNI's regulations and internal criteria mention many academic products and activities, and they contain statements that qualitative assessments should prevail over quantitative assessments. But it is also true that quality is not defined, except in the context of considering bibliometric indicators, plus vague references about prestige (e.g., in the case of book publishers). There are no indications of how to compare the "value" of different products and activities. This is why we posed a number of questions to help improve the evaluation process (p. 834). For example, how should scientific innovation be valued? This is currently left unanswered in the SNI; our scoring system of "rule-based peer review" would address these questions and thus represent an improvement.

MARTIN RICKER
mricker@ibiologia.unam.mx
HÉCTOR M. HERNÁNDEZ
hmhm@ibiologia.unam.mx
and DOUGLAS C. DALY
ddaly@nybg.org

Note from the Editor

The controversy brought up by the letter received being highly relevant to the scientific community of the region, we have considered the suggestion of requesting a very brief closing commentary from a respected member of the Mexican scientific community.

COMMENTARY

Debating about SNI, the Mexican Federal Government system to incentive "productive scientists", sheds light on important and controversial aspects of science evaluation. Both sides in the debate, Ricker, Hernández and Daly (RHD), and Williams and Aluja (WA) agree in that the SNI, created in 1984, provided a major incentive for Mexican scientists to publish their research regularly, and in international journals. Mexico has, after Brazil, the largest (and growing) scientific productivity of Latin America (Glanzel *et al.*, 2006, *Scientometrics* 67: 67-86). However, RHD highlight some problems of: i) implementation of the SNI, and ii) unintended and negative second order effects of defining productivity mostly by the number and type of publications (peer reviewed international journals) and citation rates. RHD correctly identify what was an implementation problem associated to unsophisticated uses of scientometric indicators, but WA are probably right in suggesting that RHD make too much of the acknowledged fact that different disciplines have different distributions

in the values of ISI indicators. Everybody knows this, including the SNI committees. Taxonomy is not alone in this problem. Pure mathematics has a similar if not worse problem (Batista *et al.*, 2006, *Scientometrics* 68: 179-189). It is probably true that at the beginning of the SNI this problem was more prevalent, but in the last 15-20 years some of the most prestigious taxonomists in Mexico regularly participate in the committees, so the point may be moot.

Still, there may be second order, unintended negative effects of evaluations based on scientometric indices, as RHD point out. WA respond convincingly in their sixth section: SNI uses universally accepted criteria of evaluation. Doubtless, an incentive program like SNI, which has effects on a scientist's prestige, career and payment, has the potential of creating distortions in the way science is performed (Lawrence, 2007, *Current Biology* 17: 83-85). RHD and WA disagree about the extent of these distortions. In my experience the distortions exist, but to a much lesser degree than RHD suggest. If

convincingly documented, distortions can be self-corrected by the scientific community of Mexico. However, I see far greater risks in attempting, as RHD propose, that SNI "will have to take social priorities into consideration" and should "apply the correct filters from society's viewpoint." This proposal has truly deep (and in my view disastrous) consequences for science evaluation. The debate about this in the context of SNI is barely started by RHD and it is a very important one. In practice, SNI creates incentives for publishing refereed papers in international journals (although as WA state, the evaluation is far more comprehensive). To a first approximation, this indicator correlates well (not perfectly) with *quality* of science, on the other hand, it is highly debatable that "society's points of view" would do better to promote high-quality science. I know, from first hand experience, how difficult it is to explain to non-experts why research with titles like '*Reassessment of the enigmatic Lepidopteran family Lypusidae (Lepidoptera: Tineoidea; Gelechioidea)*', or '*A lambda-lemma for normally hyperbolic invariant manifolds*' should be supported by their taxes. RHD are on a very slippery slope with this proposal. It would be interesting to see other readers debating the point.

JORGE SOBERÓN M.
jsoberon@ku.edu