

# ICMI Column

Merrilyn Goos (University of Limerick, Ireland), Vice President of ICMI

## ICMI Statement on Evaluation of Scholarly Work in Mathematics Education. A call for comments

At the ICMI executive committee meeting held in Geneva in March 2017, it was noted that ICMI had been approached to inquire whether our organisation has an official stance regarding use of citation indices as the basis for evaluation and promotion of scholars in academic positions. A suggestion arising from that meeting was that ICMI could refer to the recommendation on the evaluation of individual researchers in the mathematical sciences that had been issued by the International Mathematical Union (IMU) (available at [https://www.mathunion.org/fileadmin/IMU/Report/140810\\_Evaluation\\_of\\_Individuals\\_WEB.pdf](https://www.mathunion.org/fileadmin/IMU/Report/140810_Evaluation_of_Individuals_WEB.pdf))

A similar document based on the same considerations has now been developed by ICMI. We invite all members of the ICMI community to read this document (see below) and send us any comments by 30 September 2019. Please email comments to ICMI vice president Merrilyn Goos at [merrilyn.goos@ul.ie](mailto:merrilyn.goos@ul.ie). The final version of this document will then be published on the ICMI website.

### Evaluation of scholarly work in mathematics education

Evaluating the quality and impact of scholarly work in all academic disciplines has become an increasing concern of universities as well as many national governments. However, generic evaluation processes do not always take into account discipline-specific norms for conducting and publishing research and other forms of scholarly work undertaken to influence practice or policy. Even within the global field of educational research there exist various sub-fields that take different approaches to theory, method and dissemination of findings.

Concerns about the need to improve the evaluation of scholarly work have led to the formulation of various statements and recommendations that are either specific to a discipline<sup>1</sup> or applicable to all research fields.<sup>2</sup> The purpose of the present document is to consider the question of how to evaluate scholarly work in the specialised educational sub-field of mathematics education. It sets out ICMI's position on evaluation of individual researchers in mathematics education.

This document is organised around three questions, with brief responses set out below that are elaborated in subsequent sections:

<sup>1</sup> See the IMU (2014) statement on evaluation on researchers in the mathematical sciences.

<sup>2</sup> See the San Francisco Declaration on Research Assessment (DORA, n.d.) – a worldwide initiative covering all scholarly disciplines and all key stakeholders including funders, publishers, professional societies, institutions and individual researchers.

### 1. What is being evaluated and for what purpose?

- Individuals or institutions? Research output or other forms of scholarly work?
- For decisions about hiring, promotion and tenure?
- For decisions about institutional resource allocation and continuation or cessation of funding for research centres or institutes?

### 2. What problems arise in evaluating scholarly work in mathematics education?

- Mathematics education research journals are not adequately represented in citation databases.
- Journal citation metrics are improperly used as an indicator of article quality.
- Predatory publishers exploit inexperienced researchers.
- Evaluation focuses on too narrow a range of scholarly work.

### 3. What solutions can be proposed?

- Promote alternatives to citation-based evaluation systems.
- Develop ways of evidencing research impact as well as research quality.
- Broaden the scope of evaluation to include scholarly activity that influences educational practice and policy.

### A. What is being evaluated and for what purpose?

Academics employed in universities are expected to devote some of their time to evaluating the scholarly work of other *individuals*, for example, by reviewing journal manuscripts, conference papers and grant applications, examining research students' theses, or assessing academic performance to inform decisions about hiring or promotion. Expert peer review is universally recognised as being fundamental to research evaluation, since only experts in a field can judge the significance and originality of a piece of research or the quality and relevance of the publication outlets in which the findings are disseminated.

Research evaluation can also be used to judge the performance of higher education *institutions* with the goal of providing accountability for public spending on research. Some countries (e.g., the UK, Australia, New Zealand) conduct regular national research evaluation exercises that typically place most emphasis on publication quality, with scores or ratings being assigned to either individual academics or discipline-based units of assessment within each institution.<sup>3</sup> Judgments about

<sup>3</sup> For more information, see <https://www.ref.ac.uk/about/> (UK), <https://www.arc.gov.au/excellence-research-australia> (Australia), <https://www.tec.govt.nz/funding/funding-and-performance/funding/fund-finder/performance-based-research-fund/> (New Zealand).

research quality may be made on the basis of expert peer review or bibliometric data, or some combination of these.

Evaluation of the scholarly work of individuals or institutions is a high-stakes enterprise with significant implications for career progression and academic reputation, and sometimes for the selective allocation of institutional research funding. It is therefore essential to use valid measures that not only capture the distinguishing features of quality in a specific discipline, but also avoid perverse consequences that might lead to “gaming” of the evaluation system and thus distortion or undermining of research goals.

### **B. What problems arise in evaluating scholarly work in mathematics education?**

Research evaluation depends largely on assessment of the quality of research outputs. In mathematics education, papers in peer-reviewed journals are typically the most highly regarded form of publication. Evaluation of such outputs can be either quantitative, relying on various forms of bibliometric analysis using citation data, or qualitative, relying on expert peer judgment.

A major limitation of citation-based systems for evaluating journal quality is the limited coverage they give to mathematics education journals. Nivens and Otten (2017) compiled a list of 69 journals that have an explicit focus on mathematics education research, but found that only six appeared in the Web of Science database from which journal impact factors are calculated. They concluded that Web of Science is of little value to mathematics education, despite its widespread use to measure scholarly output in other disciplines. A further limitation of all three major journal ranking systems – Web of Science (Impact Factor, IF), Scopus (Scopus Journal Ranking, SJR), and Google Scholar (h5-index) – is that they only trace citations within their own databases, thus excluding the vast majority of mathematics education journals.

Nivens and Otten (2017) warn of a further problem: when journal citation metrics are improperly used to draw conclusions about the impact of articles published in particular journals. They show that there is little correlation between a journal’s citation-based measures of impact (such as IF) and the number of citations received by articles published in that journal. Yet journal impact measures and rankings are often used – inappropriately – in making decisions about tenure and promotion of individual academics.

Evaluations based on so-called “objective” quantitative methods are not inherently more reliable than expert human judgments. Williams and Leatham (2017) cautioned against giving too much credence to citation analysis in mathematics education, noting that “at a minimum, the literature raises questions of whether citation-based indices are valid and meaningful in our field and how they compare with other ranking methods” (p. 372).

Despite the significant problems outlined above, citation-based measures are increasingly being used to

compare and rank individual academics or even entire academic departments and disciplines. Such ill-advised evaluation practices can have perverse consequences. For example, researchers whose universities evaluate their performance on the basis of journal impact factors or quantitatively derived rankings can be exploited by predatory publishers that promise fast peer-reviewing without the full editorial and publishing services of a legitimate journal. Early career researchers, doctoral students and academics in developing countries are especially vulnerable to these unethical practices.

A different kind of problem that arises from attempts to evaluate scholarly work in mathematics education concerns the practice-engaged nature of our field (Nivens & Otten, 2017). Thus citations in scholarly journals are not the only way of measuring impact; in addition, researchers in mathematics education value dissemination of their scholarship in practitioner journals, through teacher education and professional development work and by influencing education policy development.

### **C. What solutions can be proposed?**

#### *Recommendation 1*

ICMI does not support reliance on only quantitative measures of research quality, and in particular citation analyses, to evaluate scholarly work in mathematics education. ICMI supports the IMU’s (2014) argument that “nothing (and in particular no semi-automatised pseudo-scientific evaluation that involves numbers or data) can replace evaluation by an individual who actually understands what he/she is evaluating”. Education in general and mathematics education in particular are grounded in diverse cultures and social contexts. Yet the richness and effectiveness of the mathematics education communities worldwide depend on this diversity.

Evaluating the contributions of individual researchers to advancing knowledge therefore requires different and complementary approaches in order to do justice to these complexities. At the very least, any quantitatively-based rankings of journals should be supplemented with qualitative judgments informed by the expert survey of journals conducted by Williams and Leatham (2017).

#### *Recommendation 2*

Analysis of journal citation data leads to flawed measures of academic impact. Alternative impact measures are being developed in some countries, where impact is defined in terms of “the demonstrable contribution that research makes to the economy, society, culture, national security, public policy or services, health, the environment, or quality of life, beyond contributions to academia” (Australian Research Council, 2012). These broader measures of impact should be included in any evaluation of scholarly work in mathematics education.

#### *Recommendation 3*

Following on from the previous recommendation, ICMI supports broadening the scope of evaluation of schol-

arly work to recognise academic activities that influence practice and policy in mathematics education.

### References

- Australian Research Council (2012). Research impact principles and framework. Retrieved 8 November 2016 from <http://www.arc.gov.au/research-impact-principles-and-framework#Definition>
- Declaration on Research Assessment (DORA) (n.d.). San Francisco Declaration on Research Assessment. Retrieved 4 May 2019 from <https://sfedora.org/read/>
- International Mathematical Union (2014). Recommendation on the evaluation of individual researchers in the mathematical sciences. Retrieved 3 March 2019 from [https://www.mathunion.org/fileadmin/IMU/Report/140810\\_Evaluation\\_of\\_Individuals\\_WEB.pdf](https://www.mathunion.org/fileadmin/IMU/Report/140810_Evaluation_of_Individuals_WEB.pdf)
- Nivens, R. A., & Otten, S. (2017). Assessing journal quality in mathematics education. *Journal for Research in Mathematics Education*, 48, 348–368.

- Williams, S. R., & Leatham, K. R. (2017). Journal quality in mathematics education. *Journal for Research in Mathematics Education*, 48, 369–396.



*Merrilyn Goos is professor of STEM Education and director of EPI\*STEM, the National Centre for STEM Education, at the University of Limerick, Ireland. Before taking up this position she worked for 25 years at The University of Queensland, Australia. She was formerly editor-in-chief of educational studies in mathematics and is currently vice-president of the International Commission on Mathematical Instruction.*