Assessing research in the mathematical sciences

Alan L. Carey*, Michael G. Cowling** and Peter G. Taylor***

In the context of the forthcoming Research Quality Framework (RQF), we discuss the assessment of the quality of research in mathematics and statistics. The purpose of this document is two-fold: to act as a resource for RQF panel members and to help group leaders prepare context statements, as the RQF will require them to do. However, the discussion of assessment that we propose is equally valid for the purposes of hiring and of promotion.

Introduction

The recommended procedures for the Australian Research Quality Framework (RQF) were released in October 2006 [5]. Under these procedures, research in Australia will be assessed by a number of assessment panels. While the most relevant panel to the mathematical sciences is 'Panel 4: Mathematical and Information Sciences and Technology', a number of other panels are also relevant, particularly for statisticians and applied mathematicians.

This document aims to indicate the many ways in which mathematical sciences research can be and is evaluated internationally. Essentially, it is a statement about the research culture of the mathematics and statistics community. Our hope is that this will be a guide to assessors on RQF panels on how to judge the quality of research in the mathematical sciences. We also intend this document to be of assistance to researchers in drawing up context statements, as detailed in [5, Section 4.1.5].

Primarily, the mathematical sciences are described by Research Field Courses and Disciplines Classification (RFCD) Codes 2301 (mathematics), 2302 (statistics), 2399 (other math sciences) and 2804 (computation theory and mathematics). However, these codes do not cover all research output in the mathematical sciences, which also appears under the RFCD codes associated to many other disciplines, including the various branches of engineering and theoretical physics. Further, mathematical, and particularly statistical, research is also published in journals related to biological, medical and agricultural science, economics and in many parts of social science.

There are vastly different research cultures in the mathematical sciences, and even within various subdisciplines of mathematics and statistics. To give some extreme examples, Andrew Wiles published an average of one paper a year over 13 years before he proved Fermat's Last Theorem, and yet his achievement is arguably the most significant mathematical result of the late twentieth century. The great logician Gödel's research output consisted of

*Centre for Mathematics and its Applications, Mathematical Sciences Institute, Australian National University, Canberra ACT 0200 E-mail: acarey@maths.anu.edu.au
**School of Mathematics and Statistics, University of NSW, UNSW Sydney NSW 2052
E-mail: m.cowling@unsw.edu.au
***Department of Mathematics and Statistics, University of Melbourne Vic 3010, Australia
E-mail: P.Taylor@ms.unimelb.edu.au

half a dozen papers. On the other hand, Paul Erdős published over 1500 papers, the great majority in collaboration with colleagues all over the world, and the polymath Leonhard Euler wrote close to one thousand papers. However, these extremes are not representative, and in most areas of the mathematical sciences, a publication rate of one to five papers per year is considered 'normal'. There are significant variations between the norms of the many subfields.

Because of this variability, most mathematicians and statisticians argue that it is dangerous to use bibliometric data without first attempting to understand the culture of the relevant subdiscipline. Indeed, as argued comprehensively in [4], the use of crude measures of productivity based on standards of related but distinct disciplines is likely to reduce quality in the long term. Most mathematicians and statisticians agree that it is important to use a much wider range of indicators than just publications.

Publications

Most mathematicians and statisticians support the principle espoused in [5] that quality is best assessed via disinterested peer assessment of research outputs. In particular, there is widespread support for the principle that high quality publications are the primary indicators of research strength, and that assessors should read some of the papers or books that have been produced, not just rely on scientometric measures such as ISI journal impact factors.

Mathematicians and statisticians produce several types of outputs. The most common are papers in refereed journals with an international editorial process and international circulation. Articles are also published in conference proceedings, frequently as a result of mathematicians and statisticians working close to a field of application and publishing according to the culture of that field. For example, a mathematician who collaborates with engineers or computer scientists might publish in conference proceedings in these areas, which are often amongst the most prestigious outlets available.

In contrast, many mathematics conferences do not publish proceedings at all. Indeed, a common characteristic of the most prestigious conferences in mathematics and statistics is that speakers are expected to present their work and then publish it in fully-refereed journals. Scholarly books are only rarely produced.

The ordering of the authors is frequently judged to be important by those compiling quantitative measures of research output. In mathematics and statistics, it is very common, but not universal, to order authors alphabetically (this may disadvantage authors whose surnames begin with later letters in the alphabet [1], but there does not appear to have been a complete analysis of this). Other systems that are used by some people include ordering authors by the level of contribution, going from highest to lowest, and putting graduate students first. The results reported in a joint paper are frequently the product of 'brainstorming' sessions attended by all authors, and in this context, it is often argued that the result would not have been produced if there had not been a contribution from all authors, and that attempting to apportion different proportions of the idea to each of the authors is a fruitless exercise.

It is virtually unknown for mathematicians to be listed as authors on papers to which they have not made substantial intellectual contributions; this contrasts with many laboratory disciplines, where researchers' names are often included in the list of authors by virtue of their position in a laboratory. As a result, mathematicians tend to be listed as authors of fewer papers than their colleagues in the experimental sciences. Again in contrast to many other disciplines, papers in the mathematical sciences do not even attempt to compile an exhaustive bibliography of all relevant papers. Rather, a paper will be cited because a result contained therein is needed. Particularly coupled with the fact that in many subdisciplines of mathematics publication is infrequent, this means that numbers of citations of a paper in the mathematical sciences is generally lower than that of a paper in many other sciences. This in turn leads to scientometric indices such as impact factors of mathematical sciences journals being lower than those of other scientific disciplines.

In the mathematical sciences, there can be a considerable time-lag, typically between one-and-a-half and two years, between manuscript submission and subsequent publication. This should be kept in mind, especially when the performance of an early career researcher is assessed. An important consequence of this lag is that the ISI journal impact factor is not a robust measure of a journal's standing, since it only takes into account the number of citations in the two years following publication. In fact some of the most prestigious mathematics journals have low impact according to this measure, and rankings by impact can vary widely from year to year: for instance, in the 'Mathematics' ranking by impact factor of some 120 journals, the *Publications Mathématiques de l'Institute des Hautes Etudes Scientifiques* was 100th in 1989 and first in 1990.

There is, however, a generally accepted crude ranking of mathematics and statistics journals within discipline areas in terms of their quality, which is quite well correlated with impact when this is measured over decades rather than years. Expert opinion can advise on this. For more on publication patterns in mathematics and the evaluation of journals, see [2], [3].

Grant Funding

The major use of grant money in mathematics and statistics is to fund employment of postdoctoral fellows and other staff. The existence of such funding is essential for the development of the next generation of practitioners in the disciplines, so it is very important that mathematicians engage in the grant application process.

It is possible for some mathematicians to pursue their research without engaging vigorously in the grant process. Despite this, we would argue strongly that grant success should be applied as a measure of research productivity. It is a good indicator of the esteem that researchers are held in by their peers. Moreover, a person who has a long and consistent record of grant success is very likely to be a research leader in the sense that they will have supervised and mentored a number of postdoctoral fellows. The most common granting scheme accessed by Australian mathematicians is the ARC Discovery Grant Scheme. Applied mathematicians and statisticians are also able to access the ARC Linkage Grant Scheme. International funding is also becoming a prestigious source of support for research in the mathematical sciences.

One consequence of the fact that mathematicians and statisticians generally do not need expensive experimental equipment is that they generally apply for less funding per application than other scientific and technological disciplines. It is therefore inappropriate to judge grant success by looking at the total funding earned, especially if this is compared with researchers from disciplines that do use expensive equipment. A better measure is the number of successful applications, or the rate at which a researcher achieves success.

Postgraduate research student training

An important role of academics in all fields is their contribution to the development of higher degree students. A good measure of the effectiveness of a higher degree program is the proportion of students who graduate in a timely fashion and go on to employment (including further study) in the area of their study and research.

As with other measures discussed above, it is important to assess research student supervision in a manner that is appropriate to the discipline. Typically, a supervisor–student relationship in mathematics or statistics resembles those that occur in the humanities more than those in the laboratory sciences. Supervision is frequently one-to-one between student and supervisor. Joint supervision is becoming less unusual but, even then, all supervisors have to keep on top of the detail of the intellectual content.

It is rare for students to work on a research problem that is a small part of a large project that a team, including other students, is working on. On the contrary, a supervisor may simultaneously look after students who are working on several very different projects, which require different sorts of intellectual input.

These factors (and others) mean that the rate of production of PhD graduates in the mathematical sciences is lower than in many other scientific disciplines. Overall, there have been less than 1500 PhD graduates in the mathematical sciences in the entire history of the Australian university system and the number of people who have supervised more than ten students is quite small. The contribution to student supervision of such people ought to be regarded very highly. Mid-career researchers who have supervised between five and ten successful PhD students ought to be well-regarded for their contribution to supervision.

Further indicators of esteem

There are a number of other measures that are good indicators of research standing in the mathematical sciences community. These include:

- Invited conference talks at highly selective and prestigious venues. These include large conferences, such as the International Congress of Mathematicians, the International Congress of Industrial and Applied Mathematics, and the International Congress on Mathematical Physics, as well as smaller focussed events such as those held at the Mathematisches Forschungsinstitut at Oberwolfach in Germany.
- Invited fully- or partly-funded visits to leading research centres and institutes such as the Newton, Erwin Schrödinger and Mittag-Leffler Institutes in Europe or the Mathematical Sciences Research Institute, the Institute for Mathematics and its Applications and the Fields Institute in North America.
- Prizes, fellowships and awards, particularly those won in international competition.
- The quality and extent of researchers' collaborations is often taken as a good measure of their standing in the mathematical sciences community. These frequently take the form of links with international leaders of the discipline or subdiscipline. For applied research, substantial industrial collaborations provide an analogous indicator.
- Membership of editorial boards of international journals.
- Membership of the organising committee or advisory board of prestigious international conferences.
- Scholarly activity such as reviewing and refereeing.
- Assessing research theses and research grant applications, especially if in another country.

- The extent to which researchers' work influences the work of other researchers.
- Production of (documentably) high-quality and widely-used software packages.

A critique of citation analysis

In this section we shall make some points about citations in the mathematical sciences.

First, to make bibliometric assessments, citation data from many sources is needed. In particular, ISI data misses articles that appear on the web where, in many fields of research, the most intense citation activity occurs prior to actual publication. Other sources, such as Google Scholar, www.arxiv.org and MathSciNet, all provide different information.

It is noteworthy that the papers for which Tao and Perelman won Fields Medals (the mathematical equivalent of the Nobel Prize) in 2006 do not show up in ISI citation data, as they were still in electronic form at the time of the award. The paper which led to the Fields Medal for Simon Donaldson had about 80 ISI citations at the time of the award, while many Nobel Prize winning works have thousands of citations.

Second, in interpreting citation data of a research output, it is essential to understand the citation culture of the subdiscipline that provides the audience for the research. There are wide variations of culture within a single four-digit RFCD code, and agglomerating data from groups with different cultures will mean that differences of quality within a subdiscipline will be swamped by differences in culture between subdisciplines.

Third, to assess the importance of a research article, the way it influences research will ideally be considered over its full life-time. This life-time varies within the various subdisciplines of the mathematical sciences, but it is common for the 'citation half-life' of an article to be over 10 years. It is certainly hard to assess correctly the long-term value of a mathematical contribution until many years after its appearance, and the RQF proposal to judge research output from a comparatively short time window is problematic for the mathematical sciences.

Fourth, the ISI classifies the mathematical sciences into 'fundamental' and 'applied' based on the labels that journals apply to themselves. This may be misleading; for example, an assessor might need to consider the citation record of papers that are genuinely 'cross-disciplinary' where the impact within the mathematical sciences may be small but the impact in another discipline may be large. Arrow's theorem in economics, a Nobel prize winning piece of research, that has not led to substantial new mathematics, is an example of such a phenomenon.

Finally, it is essential to distinguish original research papers from surveys or review articles. This does not occur with naive bibliometric analysis.

There is concern about the use of bibliometric analysis in many disciplines other than the mathematical sciences, and even suggestions that such indices are already being manipulated. See for instance [6], [7].

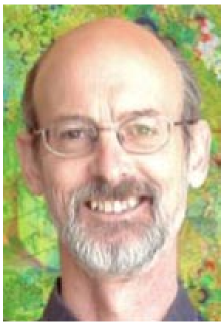
Acknowledgements

The authors thank many colleagues who have provided useful comments, information, and references, including I. Aitchison, N. Boland, P. Cerone, D. Daners, J. Denier, G. Gottwald, J. Gray, K. Landman, M. Murray, P. Pansu, F. Planchon, N. Trudinger, C. Villani and

A. Welsh, and the members of the Steering Committee of the Australian Mathematical Society. This paper has been endorsed by the Steering Committee of the Australian Mathematical Society as a statement of what matters in the mathematical sciences; it has also been endorsed by the National Committee for the Mathematical Sciences.

References

[1] Einav, L. and Yariv, L. (2006). What's in a surname? The effects of surname initials on academic success. *J. Econ. Persp.* **20,** 175–188.
[2] Ewing, J. (2006). Measuring journals. *Notices Amer. Math. Soc.* **53,** 1049–1053. Available online at http://www.ams.org/notices/200609/comm-ewing.pdf (accessed 24 April 2007).
[3] Grossman, J.W. (2005). Patterns of research in mathematics. *Notices Amer. Math. Soc.* **52,** 35–41. Available online at http://www.ams.org/notices/200501/fea-grossman.pdf (accessed 24 April 2007).
[4] Hall, P. Measuring research performance in the mathematical sciences in Australian universities. *Gazette Austral. Math. Soc.* **34,** 26–30.
[5] Research Quality Framework: Assessing the Quality and Impact of Research in Australia, Commonwealth of Australia, 2006. http://www.review.ms.unimelb.edu.au/Report.html (accessed 25 April 2007).
[6] Monastersky, R. (2005). The number that's devouring science. *Chron. Higher Ed.* October 14, 2005.
[7] Williams, T. and Wilson, J. (2005). Second thoughts on citation counts, impact factors. *OR/MS Today* August 2005. Available online at http://www.lionhrtpub.com/orms/orms-8-05/frletters.html (accessed 24 April 2007).


Alan Carey


Michael Cowling


Peter Taylor

Alan Carey studied at the Universities of Sydney and Oxford, then took up research fellowships at the University of Adelaide and the ANU. He held a continuing position at the University of Adelaide, with brief appointments to Flinders University and to ANU, from the mid-80s until 2002. Since then he has been the Dean of the Mathematical Sciences Institute at ANU. He has been a Clay Mathematics Institute Scholar on three occasions, in 2000 and again in 2001 at Harvard, and in 2006 at the Erwin Schrödinger Institute in Vienna. He was President of the AustMS from 2000 to 2002.

Michael Cowling studied at ANU and Flinders University, then spent a decade in Canada, Italy and the USA before returning to Australia as Professor of Mathematics at UNSW in 1983. Career highlights since then include the Medal of the Australian Mathematical Society, a 10-year stint as an ARC Senior Research Fellow, and two years as President of the Australian Mathematical Society. He is currently Head of the School of Mathematics and Statistics at UNSW, and Editor of the Journal of the Australian Mathematical Society.

Peter Taylor studied at the University of Adelaide in the 1980s, and then spent time at the Universities of Western Australia and Adelaide before moving to Melbourne at the beginning of 2002. In January 2003, he became the first Professor of Operations Research at the University of Melbourne. He is currently a chief investigator of the ARC Centre of Excellence in Mathematical and Statistical Modelling of Complex Systems, President of ANZIAM, Head of the Department of Mathematics and Statistics at the University of Melbourne, and the Editor-in-Chief of 'Stochastic Models'.