REPORT OF THE JOURNAL WORKING GROUP

1. Executive Summary

- Journals are essential for mathematical development. In mathematical research, they provide peer validation of results and act as organs of communication.
- Mathematicians know that bibliometric data lose "crucial information that is essential for the assessment of research" [3]. However, they are widely used[1] by national research agencies, universities, and libraries as mechanisms for evaluation of journals, departments and individuals, despite their known inaccuracies for mathematics.
- Such evaluations shape mathematical development, through funding, career opportunities or library purchases. An overwhelming majority of mathematicians in the world do not have recourse to more senior mathematicians with experience at rebutting such inaccurate opinions.
- An increasing number of journals appear to be implementing practices oriented to the pursuit of rapid publication, higher impact factor, and commercial return, at the cost of appropriate mathematical judgment.

To counteract these trends and protect the mathematical literature, we have considered how the IMU and ICIAM could undertake an honest, careful rating of journals based on the judgment of expert mathematicians. We examine the benefits and the drawbacks of such a rating scheme, and describe a plan for its implementation. The committee as a whole recommends this plan; a minority are persuaded by the arguments against it in Section 10.

2. Introduction

The mathematical literature is tremendously important to the IMU and ICIAM communities. Publication of mathematical articles provides the primary method of disseminating and validating new mathematical results, and the corpus of proven and reviewed results in the published literature is very much at the center of mathematics. Journals play a major role in the creation and maintenance of the mathematical literature. They carry out numerous functions (printing, posting, distributing, indexing, etc.), but their most important function is to maintain and certify the quality (correctness, significance and originality) of mathematical work through a careful, fair, and expert peer review process.

There are many journals (MathSci now reviews about 680 cover-to-cover and about 2,000 selectively[2], the database includes reference lists from about 450[3]; the Zentralblatt MATH database indexes papers from more than 3,500 current or historical journals[4]). These differ in a variety of ways, from specialized to general coverage, from regional to global, and from highly selective to more accepting. Journals in the mathematical sciences range from those that are clearly at the top of any scale of excellence to others that are far from satisfactory.

The mathematical community is large and diverse, and there is need for a diverse set of journals, and for clearer understanding of how different journals serve the community. Unfortunately, there are also journals which are published with dubious motivations, which do not carry out acceptable peer-review, and which contribute nothing of value to the mathematical literature [2]. There appears to have been an explosion of these in recent years. They have the potential to do significant damage to the mathematical publishing enterprise, making the literature less reliable, more difficult to search, and less respected. By undermining

Date: 30 June 2011.

[1]See for example the discussion recorded on 26 August 2010 at 1800 hours, and viewable at `http://www.icm2010.org.in/from-the-venue/online-streaming-archives`

[2]`http://www.dialogselect.com/sources/0239.html`

[3]`http://www.ams.org/mathscinet/mrcit/journal_list.html`

[4]`http://www.zentralblatt-math.org/zmath/en/about`

the confidence in and prestige of mathematical publications, this situation could lead to a loss of financial support for mathematical research.

3. Why a rating of mathematical journals is desirable

There are numerous stakeholders who have a legitimate need to assess the quality of mathematical journals. Institutions that provide financial support, such as universities through library subscriptions, need such information to make wise decisions in allocating limited resources. Those involved with journal production, such as editorial boards and publishers, need accurate feedback on their journals' performance in order to improve. Potential editors and referees need information to make informed decisions about where to channel their efforts. Mathematical authors need information to make informed decisions about where to submit their work.

In view of these needs, there have been numerous efforts to evaluate mathematical journals. The best known and most influential of these is the Impact Factor (IF), produced and sold by Thomson-Reuters. Other examples are the SJR and SNIP metrics from Scopus, produced and sold by Elsevier, and the Mathematical Citation Quotient (MCQ) included as part of MathSciNet. Unfortunately these are based purely on bibliometrics, and it has been amply demonstrated that they are poorly suited to mathematics and give a very inaccurate view of journal quality. See, in particular, the Citation Statistics report produced by the IMU, ICIAM, and IMS [3]. Moreover, it has also been demonstrated that IF is easily manipulated, and that large inaccuracies in the IF as an indicator of quality have been introduced intentionally by editors and others with an interest in raising a specific journal's IF [1]. In particular, a comparison of IF with an assessment of journal quality based on expert opinion found a large discrepancy.

What is needed to counteract these problems is a careful, honest effort to evaluate journal quality based on *the judgment of expert mathematicians.*

4. Why the IMU and ICIAM should lead this effort

A collaboration of IMU and ICIAM is appropriate to lead the effort to develop mathematics journal rating for several reasons. First of all, the problem described above fits squarely within their charters. The statutes of the IMU list as a primary objective the support of "international mathematical activities considered likely to contribute to the development of the mathematical science in any of its aspects," and explicitly empower IMU to "support the publication and distribution of scientific material in the field of mathematics." The ICIAM bylaws list among the purposes of the organization "to promote industrial and applied mathematics internationally" and to promote the goals of its member societies, who are in many cases significantly involved in journal publication.

Second, journal ranking is naturally an international undertaking. Mathematical journals thoroughly cross international borders, in their authors, editors, subscribers, and readers. The wide scope of IMU and ICIAM, and their reputation among a huge body of mathematical experts, covering all fields, makes them highly and almost uniquely capable of mustering the necessary expertise for a high quality rating. Moreover, their official status and prestige give IMU and ICIAM the possibility to create a ranking which will gain wide acceptance.

Various nations have responded to the lack of an accurate journal rating system by creating their own systems, but this is problematic because individual nations for the most part do not have the expertise or resources to carry this out alone. Moreover, in the cases where countries have carried out such rankings, for example Australia[5], Brazil[6], and Norway[7], this leads to a great duplication of effort. Further, no national effort is likely to be able to attain widespread acceptance.

A third reason argues in favor of IMU and ICIAM leading this effort. Unlike mathematical societies, neither union is itself directly involved in journal publication. Therefore, the same issues of conflict of interest do not arise.

[5] http://www.arc.gov.au/era/era_journal_list.htm

[6] http://qualis.capes.gov.br/webqualis/ConsultaListaCompletaPeriodicos.faces

[7] http://dbh.nsd.uib.no/kanaler/

5. Rating criteria

We believe that the most valuable way to rate journals is by expert assessment of their output, i.e., the papers they publish. Our proposal, as detailed below, is that mathematicians of high quality examine journals' contents to assess their quality. By this criterion, a top tier journal would be one in which the large majority of papers published in it are of very high quality, and which regularly publishes papers of great significance to mathematics or a major subfield of it.

Various other criteria that could be used to rate mathematical journals were discussed:

- Quantitative bibliometrics, such as citation counts.
- Reputation, as determined by surveys.
- Evaluation of journal editorial processes.

As has been indicated, bibliometrics provide a very poor proxy for journal quality [1]. Strong arguments support the conclusion that this is true not only of the Impact Factor, but of any number that can be easily extracted from citation or similar data (such as SJR, SNIP, and MCQ). As was detailed in the IMU–ICIAM–IMS report "the sole reliance on citation data provides at best an incomplete and often shallow understanding" [3]. The need we are proposing to address aims to go beyond the shallow understanding of journals obtained from bibliometrics.

A reputational rating of journals can be valuable, and many mathematicians use one informally. However, it does not seem feasible to create a formal or extensive reputational rating with any official status, and consequently this would not likely meet the needs listed above.

The possibility of rating a journal's editorial processes rather than its outcomes was discussed in detail by the working group. This included the possibility of producing an endorsed list of journals with approved processes. We concluded, for two reasons, that this is not a viable alternative. First of all, it does not address the need for assessing the *quality* of journals. We believe that there is much less interest in process information than quality assessment, and so such a rating would not address the need we set out to meet.

Moreover, it does not seem feasible to produce an accurate assessment of journal editorial processes. The natural source of information about the editorial processes is the journal itself. However, recent years have seen a burgeoning of predatory journals. Many of these gain some or all of their income through author charges, for example justified as open-access fees. Peer review procedures are often minimal or non-existent, but the journal claims that it carries out full peer review. In short, deceitful description of editorial processes is part of the operating procedure of a growing group of journals. Therefore, if IMU/ICIAM were to endorse journals, it would not be able to do so based on information provided by the journals. However, getting such information independently would be difficult or impossible. Even editorial board membership is problematical. There are examples among even among highly predatory journals of very distinguished editorial board members. In some cases, the editors have agreed to allow their names to be included, but have no real involvement in the journal and are unaware of its practices. In others, their names were used without permission. Once again, without an impractical level of effort, an IMU/ICIAM committee would be unable to determine either the accuracy of editorial board membership lists or the involvement of the board in the journal production.

6. A detailed proposal for IMU/ICIAM journal ranking

The proposed process would be overseen by a rating committee (RC), which would be constituted by 8–12 appointees from each of IMU and ICIAM, leading to a committee of 16–24. Committee members should be experienced and respected mathematicians who are willing to commit some of their time and energy to this project in the coming year. In the case of the IMU, RC appointees could be proposed by the Committee on Electronic Information Communication.

The RC would begin its work by creating a rough list of important subfields of mathematics by joining the list of sections for ICM[8] and the ICIAM thematic panels,[9] and then making adjustments to obtain a list of 25–35 subfields. For each subfield, the RC would appoint 5–10 members to a rating panel for journals in the area, and designate one panel member as chair. Some journals would be assigned to more than one

[8] http://www.mathunion.org/activities/icm/icm-2010-program-structure/

[9] http://www.iciam2011.com/index.php?option=com_content&view=article&id=40&Itemid=45

panel, and some panels would be directed to more generalist journals or to more interdisciplinary journals. The RC would then finalize the charge to the panels.

The first job of the panels would be to review a list of mathematics journals and select from it the subset which regularly publishes articles relevant to the subfield. The starting list might, for example, be the journals which are reviewed cover-to-cover by MathSciNet or Zentralblatt, but panels would be able to select in addition journals they felt significant to their field that were not on the starting list. These results would be compiled by the RC, which would check journals on the starting list which were not identified by any panel, and consider whether they needed to be rated (and so assigned to one or more panels).

After adjustment by the RC, each panel would have a list of journals to rate. Based on the number of journals assigned to the panel, the panel would have the opportunity to bring in additional members to ensure a reasonable workload (e.g., each panelist rating between 10 and 20 journals, with each journal receiving three independent ratings). The panel chair would then assign the journal to panel members to obtain three ratings from each (perhaps fewer for journals which are on multiple panels). Panelists would declare any conflicts of interest, for example if they were assigned a journal for which they were an editor or had a close relationship with the publisher, in which case the journal would be reassigned. Then panelists would be expected to rate their assigned journals using the established tiers such as those described in the next section, based on the content of the journal in the past few years, using both their knowledge of the journal, papers, and field, and by reviewing tables of content and looking at papers as needed to form a confident opinion.

Once the panelists had finished their individual ratings, the panel would discuss, by email or teleconference, all the ratings of journals assigned to it, in order to come up with a consensus panel rating. These ratings would be forwarded to the RC, which would reconcile journals that were rated by multiple panels.

The above process would result in a preliminary assignment of a rating tier to each of the reviewed journals. This preliminary rating would be posted to the IMU and ICIAM web sites and announced to the mathematical community, initiating a three month period of public comment. Comments collected in this period would be forwarded to the RC, which could choose to forward relevant comments to panels and/or act on them. This process might result in some adjustments to the rating, which would be finalized, posted to the web, and widely publicized. A description of the rating process and a frank discussion of its appropriate uses, limitations, and possible misuses would be included with the disseminated results.

7. Rating tiers

We propose that each rated journals be assigned to one of four tiers.

- Tier 1: A top journal in mathematics or a major subfield of it. Almost all papers published are of very high quality, and it regularly publishes papers that are of great significance. Peer-review is applied consistently and rigorously, and editorial work is carried out by leading mathematicians.
- Tier 2: Very strong journal with a carefully run and reliable peer-review process. Papers are generally of high quality, and regularly papers are published which are of significant importance in at least a subfield of mathematics.
- Tier 3: Solid journal that generally publishes reputable work and follows accepted practices of peer review, but are generally less selective than journals of Tier 2, and paper quality is more variable. Such journals may play an important role in specific communities, but are usually not considered highly important to mathematics or a subfield globally.
- Tier 4: Journals not found to meet the standards of the other tiers.

8. Timeline

From the initiation of the process, we envisage the following timeline.

- Appointment of the RC, two months.
- Definition of the panels, assignment of journals to panels, appointment of panel members and chair, four months.
- Adjustment of panel membership, if needed, one month.
- Individual panel ratings, three months.

- Panel discussion and reconciliation of ratings, one month.
- RC discussion and reconciliation of ratings, one month.
- Public comment period, three months.
- Final adjustments and posting, one month.

This sums to 16 months, but there will be some overlapping of these processes, and some variations in the timelines of the different panels.

9. Financial considerations

Monetary costs of the proposed effort will be small. Panelists and RC members will serve on a volunteer basis, and it is expected that all or almost all of the communication among them will be via the internet. Support from the IMU and ICIAM will be required in disseminating the output of the process. There may be cases where panelists will not have direct access to journals they need to examine, and will need to be sent tables of contents or articles.

10. Risks of a rating scheme

No rating scheme is perfect, and the working group discussed a variety of potential problems and drawbacks. Although we have tried to design the process to minimize these, some risks and drawbacks inevitably remain. We list the primary ones here.

- The use and abuse of quantitative research evaluation, and the ranking of the research output of institutions, individuals, and even different fields, is very problematic. One may worry that by endorsing a system of expert ranking of journals, the IMU and ICIAM may been seen as condoning or giving support for other sorts of ranking, or for the use of a single measure to judge complex phenomena. As explained earlier, such formulaic processes do not reflect the reality and richness of mathematical research and their inadvertent affirmation could disadvantage mathematics compared with other disciplines.
- While we judge the workload involved to be manageable by the community and justified by the need, it does require a significant commitment of effort.
- Above we have criticized bibliometric ranking of journals as providing an inaccurate measure of journal quality. Another negative aspect of such rankings is that they have been misused in an attempt to evaluate individual departments and researchers. It is of great importance to acknowledge that, while the quality of a journal depends on the quality of the papers that appear in it, the quality of any individual paper is not determined by the quality of the journal in which it appears. A risk of any journal rating scheme is that this point will not be understood by some people, and the rating, however careful or accurate, may be used inappropriately as part of some other evaluation process. The IMU–ICIAM rating scheme should be prefaced by a clear discussion of the inappropriateness of such misuses.
- The richness and diversity represented by the many mathematical journals is not captured by the small number of tiers. Quality, as defined here, is only one aspect of a journal.
- Although the panel-based expert assessment proposed here is far more difficult to manipulate than citation-based metrics, one may worry about "gaming" by groups or individuals. (This would be mitigated by keeping the panel membership confidential.)
- The rankings could prove contentious and open to challenge.
- The rankings could prove deleterious for journals that do not achieve the higher tiers, and may make it harder for new journals until they have become established and been rated.

Instead of carrying out its own rating process, it has been suggested that the IMU and ICIAM could endorse and publicize the existing ratings done by other countries or other entities as an alternative to bibliometric ranking, and encourage more rankings in the hope that this would inundate the ISI citation impact factor with other rankings and make it less important.

Appendix A. Background

IMU and ICIAM jointly formed the Working Group on Journal Ranking and Pricing on 19 October 2010 as a first step towards answering the question "What actions should follow the findings of the report *Citation Statistics*?". This step was stimulated by the paper "Nefarious Numbers" by Doug Arnold and Kristine Fowler [1] and the proposal "Thoughts about journals and the role of international mathematical organisations" by Stefan Müller. The terms of reference of the Working Group (or, in brief, WG-JRP) are listed in Appendix B and the members of this working group are listed in Appendix C.

The terms of reference on possible ranking of journals dominated the group's early discussions. A summary of our discussions was presented to the IMU's Executive Committee (EC) on 27 February 2011. This led to two modification of the tasks entrusted to the working group, first, a recommendation that a new committee be formed to consider terms of reference on pricing and accessibility of journals, and second, additional instructions to the working group on journal rating, which are included in Appendix D.

Appendix B. Terms of Reference

(1) The working group is charged with the task of proposing ideas to ICIAM and IMU of how to actively react to the situation described above.
(2) Each proposed idea should be accompanied by an estimate of the efforts involved in establishing and maintaining an implementation of the idea in the long run.
(3) The working group is asked to comment on the possible effects of each proposal, in particular, on changes in the behavior of researchers, universities, funding agencies, publishers etc. Are these acceptable? May there be legal implications?
(4) One possible proposal could be a ranking system for journals created by the mathematical community. In addition, does it make sense to include preprint servers in such a ranking? If so, how? Is the workload in establishing and maintaining a community based system acceptable? And will in practice sufficiently many mathematicians be prepared to be involved on a continuing basis? How can this process be organized?
(5) The working group is also asked to consider what other possible options there may be for protecting against the inappropriate use of impact factors and similar manipulable indices for evaluating research.

Appendix C. Members of the Working Group

- Nalini Joshi (Chair, The University of Sydney)
- Douglas N. Arnold (University of Minnesota)
- Carol Hutchins (Library, Courant Institute of Mathematical Sciences, New York University)
- John D. S. Jones (Warwick University)
- Malcolm MacCallum (University of Bristol)
- Peter Michor (Universität Wien)
- Stefan Müller (Universität Bonn)
- Tao Tang (Hong Kong Baptist University)

Appendix D. Result of discussion with the IMU EC on 27-28 February 2011 (Perth WA, Australia)

The IMU Executive Committee agreed that the current Working Group on Journal Ranking and Pricing should continue to consider the following issues:

a) Though there is merit in providing a resource to support evaluation of mathematical research for countries that do not have a large population of "peers" who can offer appropriate evaluation, a journal rankings scheme endorsed by IMU could lead to negative consequences, which include the following:
   - The IMU becoming an administrator of an approval process (rather than a supporter of quality research in mathematics).
   - Journal rankings being misused to rank individuals.

– Encouraging "gaming" by groups or individuals to optimise their rankings.

b) Instead of rankings that are "negative", IMU could produce an endorsed list of journals, leaving aside journals that do not meet appropriate criteria.

c) Journal rankings purely within mathematical sciences may not be adopted by national research ranking schemes that require ranking of all sciences (e.g. the proposed Research Excellence Framework 2014 in the UK).

d) The IMU can promote accessibility by producing a list of endorsed free electronic journals, to counter the situation where journals without a paper circulation are given little or no recognition in evaluation exercises.

e) The IMU could persuade AMS to adapt the MCQ rankings to make it more widely applicable, for example, by broadening its citations to include journals that lie outside its current base of journals.

f) A mini-workshop should be held one afternoon at ICIAM2011 in Vancouver, Canada with the aim of producing an action list to be recommended to IMU and ICIAM. Members of the Working Group who would normally not attend ICIAM may request travel support from IMU.

References

[1] Douglas N. Arnold and Kristine K. Fowler, Nefarious Numbers, *Notices of the American Mathematical Society* **58** (2011) 434–437 and *Gazette of the Australian Mathematical Society* **38** (1) (2011) 9–16.

[2] Douglas N. Arnold. Integrity under attack: the state of scholarly publishing. *SIAM News*, 42(10):2–3, 2009. Spanish translation in *Gac. R. Soc. Mat. Esp.*, 13(1):21–25, 2010; Chinese translation in *Mathematical Culture* 4, 2010.

[3] R. Adler, J. Ewing, and P. Taylor, Citation statistics. Statistical Sciences 24 (2009), 1–14.