Journal for Research in Mathematics Education
2017, Vol. 48, No. 4, 369–396

Journal Quality in Mathematics Education

Steven R. Williams and Keith R. Leatham
Brigham Young University

We present the results of 2 studies, a citation-based study and an opinion-based study, that ranked the relative quality of 20 English-language journals that exclusively or extensively publish mathematics education research. We further disaggregate the opinion-based data to provide insights into variations in judgment of journal quality based on geographic location, journal affiliations and publishing records, and experience in the field. We also report factors that survey respondents indicated were important indicators of journal quality. Finally, we compare our results to previous related rankings and conclude by discussing how our results might inform authors, editors, and evaluators in their efforts to publish and recognize quality research in mathematics education.

Keywords: Journal quality; Mathematics education research journals; Publishing; Research venues

Copyright © 2017 by the National Council of Teachers of Mathematics, Inc., www.nctm.org. All rights reserved.
This material may not be copied or distributed electronically or in other formats without written permission from NCTM.

Rating the quality of research remains an indispensable part of making decisions related to faculty hiring, retention, promotion, and resource allocation in university settings as well as decisions on research funding from both private and government sources. Two important components of research quality are (a) how useful the research is to other scholars or practitioners and (b) the opinion of scholars in the same field as to the novelty, validity, and importance of the work (Lester, 2005; Schoenfeld, 2007). These two components are often measured, respectively, by how often the published reports of research are cited and by seeking the opinion of knowledgeable experts in the field as to the quality of those reports. Thus, in determining the quality of a candidate's work, a promotion committee may consider both the number of citations to the candidate's published work and the expert opinion of that work expressed in external letters of review.

A commonly used proxy for the quality of research reports is the quality of the journals in which the research is reported (Cooper, Blair, & Pao, 1993). If a study of publications reveals that one journal publishes articles that are on average cited more frequently or judged to be typically of higher quality than those of another, then the journal is considered to be of higher quality, and publishing in that journal is taken as de facto evidence that the research itself is of higher quality. Although there have been a number of proposed methods for measuring journal quality, they generally fall into the two categories suggested above: those based on citation frequency (*citation based*) and those based on expert opinion (*opinion based*).

The guest editor for this article was Edward A. Silver.

Such quality measures have real consequences for individual scholars. How often researchers' works are cited can have a marked effect on salary (Hamermesh, Johnson, & Weisbrod, 1982) as can the perceived quality of the journals in which they publish (Cooper et al., 1993; Mittal, Feick, & Murshed, 2008). Journal rankings are used by many governments as a measure of research output and quality (Auranen & Nieminen, 2010) and are factored into decisions about resource allocation (Oswald, 2007). Decisions about tenure and promotion are also based in part on the quality of the journals in which a candidate's research appears (Kokko & Sutherland, 1999; Starbuck, 2005). Some universities have even instituted incentive programs to encourage publication in better journals (Manning & Barrette, 2005). Undoubtedly, having a clear idea of the quality of journals is of importance to scholars at all levels and across disciplinary areas, including the field of mathematics education.

Journal Quality Rankings

Citation-Based Methods

Citation-based methods attempt to measure the impact of a journal on a field by measuring the citations to articles in that journal. Counting total citations in a volume is one such measure. Other measures attempt to take into account the relative sizes of journals. For example, the Impact Factor (IF) is computed by dividing the number of citations in a given year by the total number of items published in the journal for the prior 2 years. Although there are related measures (e.g., the *immediacy index* and the *cited half-life*), the IF is the most commonly used measure across academic disciplines.

A significant problem with the IF and its associated measures is that they are available only for journals indexed in the Thomson-Reuters Science Citation Index (SCI) or Social Science Citation Index (SSCI) databases. According to Kurmis (2003), SCI captures less than 4% of the total number of science journals worldwide, and although we know of no comparable data for SSCI, Fairbairn et al. (2008) found that the IF is available for fewer than 20% of education journals. Currently, only three mathematics education journals, *Educational Studies in Mathematics* (*ESM*), the *Journal for Research in Mathematics Education* (*JRME*), and *Mathematical Thinking and Learning* (*MTL*), are included in the SSCI.

Two recent developments have allowed citation-based measures to be applied to a greater range of journals. In 2004, Elsevier launched the Scopus database as an alternative to the ISI Web of Science (home of the aforementioned SCI and SSCI). Using this new resource, the SCImago research group, a joint venture of four Spanish universities, developed the SCImago Journal Rank (SJR) indicator as a citation-based alternative to the IF. Roughly contemporaneous with the development of the SJR indicator was the introduction of the *h-index*, a method suggested by Hirsch (2005) for quantifying the impact of a scholar's writings. The measure has since been generalized to apply to the quality of a journal, in which case it represents the largest number h such that h published articles from the journal have been cited at least *h* times each. This index has grown in popularity as an alternative to the IF and is provided for journals in the Scopus database. Google Scholar metrics include a slight modification that is called the *h5-index*, which is the *h*-index of the articles published in the past 5 years for a given journal.

Leydesdorff (2009) compared the new citation-based quality measures (SJR and the *h*-index) with the classic measures provided by the SCI and SSCI. A factor analysis suggested that the IF and SJR were both measures of impact (as measured by citations) and that the *h*-index combined the factors related to size (number of articles published per year) and impact. This finding might imply that the SJR indicator can act as a substitute for the IF for journals that are not indexed in the SCI or SSCI. However, it likely cannot escape some of the long-recognized shortcomings of citation-based measures, which we now discuss.

Shortcomings of the IF and other citation-based measures. Substantial literature exists that questions the use of the IF and similar indices. For example, although citation-based methods are often seen as "objective" in that they depend on seemingly straightforward counts of references, it is clear that they necessarily depend on the nature of the database from which citations are drawn and, in the case of Google Scholar, on the nature of the search methods that locate citations. Recent studies (Meho & Yang, 2007; Norris & Oppenheim, 2007; Roales-Nieto & O'Neill, 2012) found substantial differences among the SCI and SSCI, Scopus, and Google Scholar databases in both journal inclusion and ability to locate references. Roales-Nieto and O'Neill (2012) found that the three databases reported quite different numbers of citations for the same article—enough so that quality indicators could differ by a factor of three among different databases. Some studies suggest that Google Scholar overreports citations (Meho & Yang, 2007; Norris & Oppenheim, 2007; Roales-Nieto & O'Neill, 2012), and in a series of articles, Jacsó (2006, 2008a, 2012) described numerous problems with the accuracy of the Google Scholar database with respect to journal ranking metrics. Furthermore, although the Scopus database does appear to be more reliable than the others, its rankings are affected by the decision not to include publications prior to 1996 and to ignore citations to such articles even though they appear in articles published since 1996 (Jacsó, 2008b).

Beyond issues of database and search accuracy, a number of studies have suggested reasons for caution with the exclusive use of the IF. Problems include bias toward certain kinds of articles (Aase, 2008), the tendency to differ across disciplines and over time (Althouse, West, Bergstrom, & Bergstrom, 2009; Seglen, 1997), and a susceptibility to being affected by publication lag, article lengths, self-citations,[1] and a number of other factors not obviously related to research quality (Seglen, 1997). Although these shortcomings are not enough to discount citation-based quality indices altogether, they do remind us that their use requires careful attention to the nature of the data underlying them and the purposes for

[1] Self-citations are citations from an article to other articles from the same journal.

which they are used. 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.

Opinion-Based Studies

Although many scholars argue that citation-based methods are superior because they are objective, others strongly favor opinion-based methods, reasoning that most decisions in an academic environment (e.g., publication acceptance, faculty advancement) are made based on the opinions of other scholars in the field (Albrecht, Thompson, Hoopes, & Rodrigo, 2010). Perhaps for this reason, ranking journals based on peer assessment has been a common and accepted method.

Shortcomings of opinion-based measures. To be sure, there are potential problems with opinion-based rankings. Those asked to rank journals on “quality” may not agree on what quality means, and they may be biased towards journals for which they have served on an editorial board or towards journals in which they have most often published (Axarloglou & Theoharakis, 2003). Responses also likely differ according to the sample of scholars who are asked to provide ratings (Wellington & Torgerson, 2005). Sellers, Mathiesen, Perry, and Smith (2004) suggested that such factors as academic rank, publication history, and departmental culture may all affect experts’ rankings.

Another potential problem with opinion-based rankings is that, compared with citation-based measures, they are time-consuming to create. Because of this issue, it is reasonable to ask whether expert opinions add to the information already available from citation-based methods. The research addressing this question is equivocal. Some studies have found a great deal of agreement between citation- and opinion-based rankings (Buela-Casal & Zych, 2010; Lowry et al., 2013); others have found agreement but with notable subtleties and exceptions (Mingers & Harzing, 2007). Still others have concluded that the two methods cannot be used to replace one another but should instead be used to help improve the overall validity of journal rankings (Jarwal, Brion, & King, 2009; Serenko & Dohan, 2011). It is likely that, at a minimum, the degree to which citation-based rankings align with opinion-based rankings would differ among academic fields.

Journal Quality Rankings in Mathematics Education

Because only three mathematics education journals are currently included in the SSCI and thus have published IFs, those interested in citation-based rankings of mathematics education journals have found it necessary to look elsewhere. We briefly summarize resources that are currently available for our field.

SJR and the *h*-Index

As noted above, the SJR and the *h*-index provide potential for obtaining citation-based rankings of mathematics education journals. Nivens and Otten (2017) recently examined these sources and compared them with other opinion- and citation-based studies (reviewed below). They found that the databases underlying these measures missed some mathematics education journals and that the measures disagreed with one another on the relative ranks of several well-known mathematics education journals. Although it is clear that the SJR and the *h*-index show promise, they do not currently provide a definitive answer to the problem of ranking journals in mathematics education.

Other Studies

To our knowledge, there are six studies that provide journal rankings for a substantial number of mathematics education research journals; three are citation based, and three are opinion based (see Table 1).

We provide the actual rankings from these studies in the Conclusions section where we also compare those rankings with the results of our studies. However, we argue that these six studies are not by themselves sufficient to provide a full picture of the quality of mathematics education journals. Studies C and D were conducted in order to inform merit, rank, and tenure decisions in a particular department. Therefore, they began with a small, idiosyncratic list of journals. Although the purposes of Studies A and B were more general, the samples of journals from which the citations were taken were again small and somewhat arbitrary. Studies A–C relied mainly on gross counts of citations without taking into account the number of published articles. The samples of experts whose opinions were obtained in Studies D and E were both small, also idiosyncratic in the case of Study D, and self-selected in the case of Study E. Each also focused on a limited geographical area. Finally, none of the first five studies (Studies A–E) underwent peer review, and so none are published in research journals. Indeed, Studies A–D exist only as brief reports circulated electronically among interested scholars, and Study E is distributed as an electronic resource on a CD. Thus, although these studies can provide some insights and starting places for further investigation, they have limitations.

By contrast, Study F is methodologically stronger, exists in published form, and has undergone a level of peer review. Therefore, it provides a better model for producing a quality ranking. As we discuss in the next section, we hope to build upon its strengths in order to provide complementary information on journal quality in our field.

Summary and Implications for Mathematics Education

Although there are and have been various attempts to provide quality rankings for mathematics education, none are definitive. The IF is not available for most mathematics education journals, and there is some indication that newer citation-based indices share the biases and scope limitations of the IF as well as potentially differing quite widely on individual journals (Leydesdorff, 2009; Nivens & Otten, 2017). Earlier citation-based studies were based on limited numbers of journals

Table 1
Prior Studies of Mathematics Education Journal Quality

Study	Name	Sources of data	Method	Results
		Citation-based studies		
A	Dreyfus (2006a)	Articles from five math education journals published in 2004–2006	Counted total citations both with and without self-citations	Ranked list of 17 journals
B	Dreyfus (2006b)	Eight volumes of Proceedings from PME 28 and 29	Counted citations	Ranked list of eight journals
C	Williams (2008)	Substantive articles from 11 math education journals published in 2005–2007	Counted total citations both with and without self-citations	Ranked list of 16 journals
		Opinion-based studies		
D	Williams (2008)	46 math education faculty from North America representing 75% of solicited responses together with some unsolicited responses	Ranked list of 22 journals in three tiers	
E	PAJE (Holbrook et al., 2009)	79 respondents self-identifying as having expertise in math education from among 803 responding; 89% from Australia or New Zealand	Ranked list of 23 journals	
F	Toerner & Arzarello (2012)	75 math education experts from 32 countries representing 82% of those solicited	Ranked list of 17 journals in four tiers	

that were not systematically selected. Opinion-based rankings have been based on small and, in some cases, idiosyncratic samples of respondents.

In an effort to address some of these concerns, we designed two studies, one citation based (Study 1) and one opinion based (Study 2), with the following goals:

1. For both studies, draw on prior research to provide a defensible and systematic way to choose a sample of journals in our field.

2. For the opinion-based study, increase the number of respondents providing judgments of journal quality.
3. Compare the results of both studies to make reasonable judgments about the relative contributions of citation- and opinion-based methods and propose a ranking of mathematics education research journals.

Along the way, we made some decisions that made the work more manageable, but these decisions also affect how our results should be interpreted. First, we chose to focus only on English-language journals. Second, in focusing only on journals that publish exclusively or extensively in mathematics education research, we excluded a number of respected education and psychology journals that publish research in our field. Our primary interest, however, was in obtaining a reliable ranking of mathematics education journals. Third, we limited our sample of journals to simplify the response process for those whose opinions we solicited. Finally, we focused mainly on North America in soliciting opinions about the journals, much as Toerner and Arzarello (2012) focused on Europe and the PAJE study (Holbrook et al., 2009) focused on Australasia. However, as discussed below, we also included a substantial number of international respondents outside of North America and are thus able to draw some conclusions about the effects of geographic region on judgments of journal quality in our field.

Study 1

Methods

We began our work with Study 1 (our citation-based study) with the intention of using its results to help determine the subset of journals to include in Study 2 (our opinion-based study). Because previous studies (see Table 1) suggested that *JRME* and *ESM* were consistently rated as top journals in the field, they provided a natural starting point. We examined all the articles from these two journals for the 4 years from 2010 to 2013 and tallied all journal references, keeping track of self-citations. Next, we narrowed this list of journals to include only those that publish extensively or exclusively English-language reports of mathematics education research. For example, references to published proceedings, journals aimed mainly at practicing teachers, or research journals publishing mainly in statistics education were excluded. From this list (in which *JRME* and *ESM* had the two highest citation counts), we identified the two journals with the next highest number of citations (the *Journal of Mathematical Behavior* [*JMB*] and *Zentralblatt für Didaktik der Mathematik* [*ZDM*]). This process was repeated by tallying all journal references from the 2010–2013 issues of these two journals, rank ordering them, and identifying the two additional journals with the next highest number of citations. We carried out this process until we had tallied the references from what were then the 10 most-cited journals in the list (see Table 2 in the Results section). Given that these journals were the top 10 most-cited journals in each of the five intermediate lists, we felt very confident about that portion of our list. Moreover, the list remained the same whether or not self-citations were counted. Ultimately, the process resulted in a list of 55 journals that met our criteria and were cited at least one time in the sample of articles from our 10 most-cited journals (see the Appendix).

Having counted all of the citations from all of the research articles in these 10 journals, we were close to having the data necessary to compute citation-based quality scores for each. Taking the IF as our model, we wanted a measure that would account for both the number of citations and the number of articles to which the citations could refer. Our modified metric differs from the traditional IF in that it accounts for citations in articles from these 10 most-cited mathematics education journals and not for citations from articles included in the entire SSCI. Given our decision to focus exclusively on mathematics education literature, however, we see this Modified Impact Factor (MIF) as a meaningful metric. We collected the last remaining data by counting the number of articles published each year in these 10 journals and by differentiating the citations to each journal by year. These data allowed us to compute MIFs for 2013 based on this 10-journal set of articles and citations.

Our plan for Study 2 was to ask participants to rank order journals, and 55 journals seemed like too many for that purpose. Only 32 of the journals had tallied 10 or more citations, but this list still seemed too long. A fairly substantial jump in number of citations occurred between the 19th and the 20th journals on the list, so we capped our Study 2 list at an even 20. These 20 journals accounted for over 98% of the references that we tallied to English-language mathematics education research journals. We thus present the results for these same top 20 journals in Study 1.

Results

Table 2 contains the results of the citation-based study for the top 20 journals. Of note is the clear separation of *ESM* and *JRME* from the rest of the journals. These two journals each received more than three times the number of citations as the next highest cited journal in our list and together account for 52% of the 7,102 total tallied citations. The next five journals (*JMB*, *For the Learning of Mathematics* [*FLM*], *MTL*, the *Journal of Mathematics Teacher Education* [*JMTE*], and *ZDM*) also seem to be grouped together, accounting collectively for 32% of the total citations. There is a more moderate separation of this group from the remaining 13 journals. The most-cited journal in the bottom 13 has fewer than half of the citations of the next most highly cited journal. These 13 remaining journals account for approximately 15% of the total citations, making it difficult to argue for meaningful differences in the citation counts between journals within this group.

Table 3 presents MIFs for the top 10 journals. *JRME* and *ESM* are still clearly the top two journals; however, whereas in Table 2 their tallies were practically the same, here there is quite a large difference between the two MIFs, even larger than the difference between *ESM* and *JMB*. Nevertheless, we see again a large

Table 2
Citation Ranking of 20 Journals That Publish Research in Mathematics Education

Ranking	Journal	No. of citations (including self-citations)[a]	No. of citations (excluding self-citations)
1	*Educational Studies in Mathematics* (*ESM*)	2,729	1,872
2	*Journal for Research in Mathematics Education* (*JRME*)	2,188	1,854
3	*Journal of Mathematical Behavior* (*JMB*)	848	554
4	*For the Learning of Mathematics* (*FLM*)	625	507
5	*Mathematical Thinking and Learning* (*MTL*)	490	429
6	*Journal of Mathematics Teacher Education* (*JMTE*)	630	427
7	*Zentralblatt für Didaktik der Mathematik* (*ZDM; The International Journal on Mathematics Education*)	740	376
8	*Mathematics Education Research Journal* (*MERJ*)	263	175
9	*International Journal of Math Education in Science and Technology* (*IJMEST*)	526	166
10	*School Science and Mathematics* (*SSM*)	307	122
11	*International Journal of Science and Mathematics Education* (*IJSME*)		97
12	*Investigations in Mathematics Learning* (*IML; formerly FOCUS on Learning Problems in Mathematics*)		96
13	*Teaching Mathematics and Its Applications* (*TMA*)		75
14	*The Mathematics Educator* (*TME*)		71
15	*Research in Mathematics Education* (*RME*)		59
16	*International Journal for Technology in Mathematics Education* (*IJTME; formerly International Journal of Computer Algebra in Mathematics Education*)		55
17	*Journal of Computers in Mathematics and Science Teaching* (*JCMST*)		50
18	*Canadian Journal of Science, Mathematics and Technology Education* (*CJSMTE*)		46
19	*PRIMUS* (*Problems, Resources, Issues in Undergraduate Mathematics Studies*)		42
20	*The Montana Mathematics Enthusiast* (*TMME*)		29

[a] These entries only exist for those 10 journals from which we tallied the citations.

Table 3
5-Year MIFs for the Top 10 Journals With and Without Self-Citations

Journal	Articles published in 2008–2012	With self-citations		Without self-citations	
		2013 citations to 2008–2012 articles	2013 MIF	2013 citations to 2008–2012 articles	2013 MIF
JRME	90	109	1.21	83	0.92
ESM	278	206	0.74	128	0.46
JMB	122	65	0.53	26	0.21
MTL	79	40	0.51	34	0.43
JMTE	135	56	0.41	43	0.32
ZDM	349	122	0.35	79	0.23
FLM	138	25	0.18	17	0.12
IJMEST	500	79	0.16	14	0.03
MERJ	111	17	0.15	7	0.06
SSM	225	13	0.06	8	0.04

difference between *ESM* and the next group of journals. Moreover, the same five journals follow *JRME* and *ESM*, although in a slightly different order. It is also the case that *FLM* is quite a bit lower than the other four. The MIF takes into account the total number of articles published, and it is interesting to note the wide range of number of articles published in these journals over a 5-year period. We also computed the MIF both with and without self-citations (IFs typically include rather than exclude self-citations). The ordering of the MIFs are similar for the computations with and without self-citations, although *JMB* and *IJMEST* move downward when self-citations are not included.

Study 2

Methods

We designed a survey to solicit participants' opinions on the relative quality of the 20 most-cited journals from Study 1 (see Table 2). Participants were contacted by email and given a link to a web-based survey. They were asked to sort the 20 journals into six categories—five categories ranging from High Quality to Low Quality and a sixth category (Not Sufficiently Familiar) for journals with which individuals felt they were too unfamiliar to rank. Furthermore, we asked participants to rank journals within categories, resulting in a rank from 1 to n for the n journals each respondent chose to rank. The survey then prompted participants to consider the extent to which various journal characteristics (e.g., publishes on a wide range of topics, is peer reviewed, most articles are of high quality) influenced their ranking decisions. Finally, the survey asked for demographic information, including in which of the 20 journals participants had published in the previous 5 years, how long they had been active members of the mathematics education research community, and their current academic rank or status.

Although we hoped that the results of our work would be useful to the field as a whole, we made many of our decisions about scope based on what would be most useful for making decisions at institutions like our own. The North American Chapter of the Group for Psychology in Mathematics Education (PME-NA) seemed to be the perfect pool for our desired audience—this research conference publishes proceedings in English, is mathematics-education specific, and is well attended by mathematics educators from across North America. We compiled the contact information for all proceedings papers from PME-NA 2012 through 2014. The 2014 meeting was held jointly with PME, thus contributing a large international set of participants.

Although these proceedings provided a broad base of participants, we worried that this selection might exclude a number of experienced individuals who regularly publish their research in journals but not in these proceedings. To address this possible limitation, we collected the names of the editorial teams and boards for the 20 selected journals. Those whose names appeared on both lists were left on the Editorial list and removed from the PME-NA list. The final numbers of participants from these two groups, as well as their participation rates, are reported in Table 4.

Although we would have preferred higher response rates, they were typical for online surveys (Sheehan, 2001). Mathematics educators from 46 different countries were among our respondents. However, as would be expected from our sampling procedure, about two thirds (68%) of our respondents were from North America. Just over one sixth (18%) were from Europe, and the remaining 14% were scattered among 21 countries in Africa, Asia, Australasia, the Middle East, and South America. Throughout our results section, we compare survey results of respondents from North America (Region 1) with those from outside North America (Region 2).

Table 4
Survey Participation Rates

		Completed rankings		Completed rankings and demographics	
List	Emailed	Number	Response rate	Number	Response rate
PME-NA	1,510	495	33%	461	31%
Editorial	425	169	40%	161	38%
Total	1,935	664	34%	622	32%

Of the 622 respondents providing information about their academic positions (which we refer to as *ranks*), assistants (22%), associates (20%), and full professors (23%) together with graduate students (18%) accounted for the vast majority. The remaining 17% described their ranks as Other (clarified by some of these respondents as indicating lecturers, research associates, postdocs, or emeritus faculty). In terms of time in the profession, roughly 60% of respondents had 10 or fewer years' experience in the field, about 20% had between 10 and 20 years' experience, and the remaining 20% had over 20 years in the field.

Recall that we asked respondents to rank the journals in two stages. First, we asked them to place each journal in one of six categories: High Quality, Medium-High Quality, Medium Quality, Medium-Low Quality, Low Quality, or Not Sufficiently Familiar. Second, we asked them to order the journals within each category from highest to lowest quality. This two-stage ranking allowed us to distinguish journals in terms of (a) how well known they are, (b) the quality categories provided in the first sorting stage, and (c) an overall ranking of all 20 journals. The first two characteristics could be measured by simple aggregates of categorical measures. The ranked lists of journals moved beyond categorical data to ordinal data, however, and required some decisions about how to combine them.

Although there is considerable literature about combining ranks, most particularly in voting, a reasonably simple and effective method still widely used is the Borda count, proposed two centuries ago by Jean-Charles de Borda (1781). It involves asking voters to rank each of n candidates from a first to a last choice. Points are then awarded such that every candidate gets n points for each first-place ranking, $n - 1$ points for each second-place ranking, and so on down to 1 point for each last-place ranking. These points are then summed for each of the n candidates to give each a Borda score. Modifications have been made to the original Borda count procedure to accommodate partial votes (in which not everyone ranks all n candidates) and similar variations (e.g., Emerson, 2013).

We chose to use a modification of the Borda count to give a score to each journal in the following way: Journals placed in the High Quality category received scores that maintained the order of the journals within the category and were evenly distributed between 4 and 5. Thus, if the respondent ranked Journals A, B, C, and D all as High Quality and ordered them such that A was best, followed in turn by B, C, and D, Journal A would receive 5 points, and the remaining three would be equally spaced between 5 and 4 with B getting 4.75, C getting 4.5, and D getting 4.25 points. Journals in the Medium-High Quality category would receive points between 4 and 3 in an exactly analogous way and so on down to journals in the Low Quality category, which would receive scores between 1 and 0. The points given to a journal were then summed across respondents to produce a Borda score for each journal. Because there were 664 respondents, Borda scores for each journal theoretically ran between 0 (if no respondent rated it) to 3,320 (if it were every respondent's top choice).

Because these Borda scores were computed from ranked data, we treated them as ordinal data in our analyses. For example, when comparing the Borda scores of journals from two different regions, we used a Mann-Whitney test for which the null hypothesis is that the distribution of ranked scores was the same for both regions.

Results

We first discuss how familiar respondents were with the journals in our sample and then turn to the respondents' opinions on journal quality. We conclude by discussing how demographic and other variables affected respondents' judgments of quality.

Familiarity with journals. By asking respondents to mark those journals with which they were not familiar, we were able to determine respondents' overall familiarity with the 20 journals in the survey. On average, our respondents reported familiarity with just under 13 of the 20 journals ($\bar{x} = 12.93$, $s = 4.171$). No significant difference was found between the mean number of journals familiar to respondents in Region 1 ($\bar{x} = 12.79$, $s = 4.158$) and Region 2 ($\bar{x} = 13.24$, $s = 4.191$). Region did make a statistically significant difference in participants' familiarity with 12 of our journals but did not affect familiarity with the other eight (see Table 5).

Length of time in the field also had an effect on familiarity with journals (see Table 6). The difference among means is highly significant ($F_{4,617} = 6.923$, $p << .001$), but the effect size is small ($\eta^2 = .043$).[2] On average, participants with 20 or more years of experience were familiar with about two more journals than those with 5 or less years of experience in the field.

Respondents' length of time in the field had a greater effect on their familiarity with some journals than it did on others. Chi-square tests on the number of respondents in each category of longevity being familiar with a particular journal showed significant effects for seven of the 20 journals. *CJSMTE*, *IML*, *FLM*, *IJMEST*, *JMB*, *TMA*, and *TMME* were better known among respondents with more experience than those with less.

Respondents' academic rank showed a similar effect on their familiarity with journals. The difference among means was again highly significant ($F_{4,617} = 9.82$, $p << .001$), but the effect size was also small ($\eta^2 = .060$).[3] On average, graduate students knew fewer journals than associate or full professors did, and those who chose Other for their rank were familiar with fewer journals than were assistant, associate, or full professors. In addition, as with time in the field, chi-square tests on the number of respondents in each category of rank being familiar with a particular journal showed significant effects, although this time for 12 of the 20 journals. Graduate students and those classifying their rank as Other were less familiar with *IML*, *FLM*, *IJMEST*, *IJSME*, *JMB*, *JMTE*, *MTL*, *MERJ*, *PRIMUS*, *SSM*, *MME*, and *ZDM*.

[2] Because homogeneity of variance was likely violated, we also computed the more robust Brown-Forsythe statistic, obtaining a value of 9.242 $df_1 = 4$ and $df_2 = 372.974$, and again $p << .001$.

[3] The result was again corroborated by a Brown-Forsythe analysis with $p << .001$.

Table 5
Percentage of Respondents (by Region) Familiar With Each Journal

Journal	Total (n = 664)	Region 1 (n = 454)	Region 2 (n = 210)	χ^2 (df = 1)
JRME	98.0%	98.2%	97.6%	0.286[a]
ESM	95.3%	93.4%	99.5%	12.13***
FLM	87.8%	85.7%	92.4%	6.014*
JMB	87.8%	89.0%	85.2%	1.884
JMTE	87.7%	88.3%	86.2%	0.605
MTL	86.4%	87.9%	83.3%	2.539
ZDM	80.9%	75.8%	91.9%	24.162***
MERJ	72.0%	70.3%	75.7%	2.115
IJSME	65.1%	59.3%	77.6%	21.31***
RME	65.1%	60.6%	74.8%	12.717***
TME	61.4%	69.4%	44.3%	38.177***
SSM	60.8%	70.7%	39.5%	58.596***
IJMEST	57.2%	51.1%	70.5%	22.02***
TMME	56.6%	58.4%	52.9%	1.777
IJTME	43.8%	39.6%	52.9%	10.177**
IML	41.1%	43.0%	37.1%	2.001
PRIMUS	40.5%	46.5%	27.6%	21.185***
CJSMTE	40.4%	36.3%	49.0%	9.627**
JCMST	34.8%	33.0%	38.6%	1.937
TMA	30.4%	22.5%	47.6%	42.914***

[a] In the chi-square test for *JRME*, one cell had an expected count of less than 5 (4.11).

* $p < .05$
** $p < .01$
*** $p < .001$

Judgments of journal quality. We now report the results of the various judgments of journal quality, first by quality categories and then by individual journals.

Distribution of quality categories. The 664 respondents' rankings of the 20 journals resulted in 13,280 votes being distributed among six categories. Just over one third of the votes reflected journals the respondents were not familiar with—that is, votes that placed journals in the Not Sufficiently Familiar category. The remaining votes were distributed among five levels of quality as shown in Table 7. These data show that few journals were considered to be of Low Quality, and only

Table 6
Mean Number of Familiar Journals

Years of experience	Number of respondents	Number of familiar journals	
		Mean	*SD*
0 to 5	217	12.16	4.654
5 to 10	153	12.63	3.804
10 to 15	74	13.27	3.489
15 to 20	56	14.54	3.78
More than 20	122	14.06	3.429

Table 7
Distribution of Votes Among Quality Categories

Quality category	Percentage	Cumulative percentage
High Quality	32.7	32.7
Medium-High Quality	30.4	63.1
Medium Quality	23.2	86.3
Medium-Low Quality	10.3	96.6
Low Quality	3.4	100.0

10% were considered to be of Medium-Low Quality. The overwhelming majority of votes cast by our respondents were for journals that they considered to be of Medium or higher quality. Thus, if our respondents were familiar with a journal, they tended to consider it to be of at least Medium Quality.

Ratings of individual journals. Table 8 shows the 20 journals from our survey along with their average Borda scores and the percentage of votes that they received in each of the five quality categories. For each journal, the largest number of votes received in any category is bolded and the votes for the largest two categories are in a shaded box. The sums for the two largest categories range from a high of 97.5% (for *JRME*) to a low of 61.6% (for *TME*). The average Borda scores allow us to give quality ratings to each journal that are meaningful in terms of the quality categories. For example, journals with average scores greater than 4.5 also have over 80% of their ratings in the High Quality category. Similarly, those with scores between 3.5 and 4.5 have roughly 80% of their ratings split between the High and Medium-High Quality categories. This pattern continues throughout the remainder of Table 8.

The information in Table 8 strongly suggests a division into four categories of quality. It is clear that *JRME* and *ESM* are together in a class of their own; they are the only two journals with more than 50% of the vote in any single category

Table 8
Percentages of Total Votes for Each Quality Category, Ordered by Average Borda Score

Journal	Mean Borda score (*SD*)	Percentages of votes in each quality category				
		High	Medium-High	Medium	Medium-Low	Low
JRME	4.69 (0.51)	**90**	7.5	1.8	0.3	0.3
ESM	4.58 (0.57)	**84.2**	13.1	1.7	0.9	0.0
JMB	3.94 (0.68)	**47**	40.5	11.1	1.2	0.2
ZDM	3.91 (0.75)	**46.6**	38.5	12.5	2.2	0.2
JMTE	3.87 (0.74)	**46**	39.3	12.5	1.5	0.5
MTL	3.80 (0.76)	**42.9**	41.3	13.6	1.7	0.5
FLM	3.70 (0.85)	35	**43.6**	16.3	4.6	0.5
RME	3.37 (0.84)	22.2	**44.2**	26.9	5.6	1.2
MERJ	3.31 (0.80)	16.1	**47.7**	29.7	5.4	1
IJSME	3.30 (0.88)	19	**40.3**	32.4	7.4	0.9
IJMEST	3.06 (0.91)	12.1	**37.6**	35.8	12.1	2.4
CJSMTE	2.93 (0.92)	7.5	32.1	**44**	12.7	3.7
IJTME	2.75 (0.92)	6.9	29.9	**41.6**	18.2	3.4
SSM	2.71 (0.97)	7.4	24.5	**40.6**	21.5	5.9
JCMST	2.58 (0.93)	4.8	26.4	**42.4**	19.9	6.5
TME	2.48 (1.05)	5.9	20.6	**37.3**	24.3	12
IML	2.47 (0.92)	4.4	17.2	**42.9**	27.5	8.1
PRIMUS	2.36 (1.03)	5.2	16.4	**34.2**	32.7	11.5
TMA	2.20 (0.93)	1.5	15.8	**35.6**	34.7	12.4
TMME	2.11 (0.96)	2.7	10.4	32.4	**35.9**	18.6

(in their case, the High Quality category), and their votes in that category are 90% and 84.2%, respectively. Thus, there is extremely strong agreement that they are the two journals of highest quality and in a sense constitute a Very High Quality category. It is important to recognize, however, that the next five highest ranked journals still have a substantial number of votes in the High Quality category (four of those five received a majority of their votes in the High Quality category). Thus, many respondents consider the next five journals also to be of High Quality. Finally, with one exception, the journals in the lowest category still received over half of their votes in categories of Medium Quality or higher; even the lowest of the four categories still represents mostly Medium Quality journals in the opinion of our respondents. It is also interesting to note that standard deviations of Borda scores for the bottom category are between 1.5 and 2 times those for the top category. Thus, there seems to be more agreement about what constitutes a High Quality journal than a Medium or Low Quality journal.

Effects of region and experience in the field on individual journal rankings. Table 9 provides the average Borda scores for each journal as determined by respondents from Regions 1 and 2. We performed Kruskal-Wallis tests of the hypotheses that the distributions of rankings for each journal were the same in each region. For half of the journals, the distribution of ranks varied across regions. For cases in which the differences in rank distributions are significant, significant pairwise comparisons of mean ranks are also provided. Effect sizes were generally

Table 9
Comparison by Region of the Distribution of the Average Borda Scores for Each Journal

Journal	Region 1	Region 2	Z
JRME	4.76	4.55	-9.116****
ESM	4.49	4.76	-8.836****
JMB	3.99	3.83	-3.393***
ZDM	3.85	4.02	-2.23*
JMTE	3.95	3.69	-3.969****
MTL	3.84	3.7	-2.434*
FLM	3.68	3.75	-1.409
RME	3.39	3.33	-0.902
MERJ	3.33	3.26	-0.382
IJSME	3.21	3.44	-2.665**
IJMEST	2.96	3.2	-2.476*
CJSMTE	2.91	2.96	-0.274
IJTME	2.69	2.84	-1.236
SSM	2.79	2.4	-3.062**
JCMST	2.64	2.45	-1.715
TME	2.58	2.18	-3.083**
IML	2.48	2.44	-0.034
PRIMUS	2.44	2.07	-2.256*
TMA	2.14	2.27	-0.998
TMME	2.05	2.25	-1.841

* $p < .05$
** $p < .01$
*** $p < .001$
**** $p << .001$

small (.10 to .16), the exceptions being *ESM* and *JRME*, for which the effect sizes were medium (.35 and .36, respectively).

Table 10 shows the overall ordering of journals by Borda scores for the two regions. With few exceptions, journals generally remained in the same quality category across regions. Those that changed are bolded. In both regions, the Very High and High Quality categories remain the same, although journals were ranked in different orders within these categories. More variation occurs within and among the Medium-High and Medium categories, but there is still a consistent core of journals in each.

Ratings of individual journals also varied only slightly across respondents with different amounts of time in the field. Table 11 shows the overall ordering of

Table 10
Overall Ordering of Journals by Borda Scores for Each Region

Quality category	Region 1	Region 2
Very High	*JRME*	*ESM*
	ESM	*JRME*
High	*JMB*	*ZDM*
	JMTE	*JMB*
	ZDM	*FLM*
	MTL	*MTL*
	FLM	*JMTE*
Medium High	*RME*	*IJSME*
	MERJ	*RME*
	IJSME	*MERJ*
	IJMEST	*IJMEST*
	CJSMTE	*CJSMTE*
	SSM	*IJTME*
	IJTME	
	JCMST	
	TME	
Medium	*IML*	***JCMST***
	PRIMUS	*IML*
	TMA	***SSM***
	TMME	*TMA*
		TMME
		TME
		PRIMUS

Table 11
Overall Ordering of Journals by Borda Scores for Each Level of Experience

	Years of experience (*n* = 622)				
Category	0 to 5 (*n* = 217)	5 to 10 (*n* = 153)	10 to 15 (*n* = 74)	15 to 20 (*n* = 56)	Over 20 (*n* = 122)
Very High	*JRME*	*JRME*	*ESM*	*JRME*	*ESM*
	ESM	*ESM*	*JRME*	*ESM*	*JRME*
High	*ZDM*	*JMB*	*JMB*	*JMTE*	*JMB*
	JMB	*JMTE*	*FLM*	*MTL*	*ZDM*
	JMTE	*ZDM*	*MTL*	*JMB*	*MTL*
	MTL	*MTL*	*ZDM*	*ZDM*	*JMTE*
	FLM	*FLM*	*JMTE*	*FLM*	*FLM*
	RME				
Medium High	*MERJ*	***RME***	***RME***	*MERJ*	*MERJ*
	IJSME	*IJSME*	*IJSME*	*IJSME*	*IJSME*
	IJMEST	*MERJ*	*MERJ*	***RME***	***RME***
	CJSMTE	*IJMEST*	*CJSMTE*	*IJMEST*	*IJMEST*
	IJTME	*SSM*	*IJMEST*	*CJSMTE*	*CJSMTE*
	TME	*CJSMTE*	***IJTME***	*SSM*	***IJTME***
	SSM	***IJTME***	*SSM*	*JCMST*	*SSM*
	JCMST	***PRIMUS***	*JCMST*		*JCMST*
		IML	***TMA***		
		JCMST	***IML***		
		TME			
Medium	***PRIMUS***	*TMME*	***PRIMUS***	***IJTME***	***IML***
	IML	***TMA***	***TME***	***IML***	***TMA***
	TMME		*TMME*	***TMA***	*TMME*
	TMA			***TME***	***TME***
				PRIMUS	***PRIMUS***
				TMME	

journals by Borda scores for each of the five levels of experience. Again, most journals remained in the same category of quality across all levels of experience; exceptions are bolded. As with the distribution by region, the overall pattern of stability of journals in the higher categories of quality is maintained with one exception, whereas somewhat more instability exists in the lower categories.

Effect of journal affiliation and publication record on individual journal rankings. We tested the hypothesis that the distributions of ranks of Borda scores for specific journals were the same for respondents who were affiliated with the journal (by being members of either the editorial team or the editorial boards for those journals) and those who were not. For 10 of the journals (*ESM*, *FLM*, *IJMEST*, *JRME*, *JMB*, *JMTE*, *MTL*, *MERJ*, *PRIMUS*, and *TMA*), distributions of journal rankings were different (and resulted in higher mean ranks) among respondents who were affiliated with the journal. We note the presence in this list of both Very High Quality journals and four of the five High Quality journals. This phenomenon thus appears to be more common among higher ranked journals.

We also tested the hypothesis that the distributions of ranks of Borda scores for specific journals were the same for respondents who had published in those journals in the past 5 years and those who had not. Distributions of journal rankings were different and resulted in higher mean ranks among respondents who had published in those journals for a majority of the journals (16 out of 20). The only journals for which no difference was found were *IJMEST*, *IML*, *TME*, and *ZDM*.

Factors reported as affecting journal quality decisions. From the existing literature and the use of pilot questions, we developed a list of 11 possible factors that might affect judgments of journal quality and asked respondents to rank them as being either very influential, somewhat influential, of minor influence, or as not being considered in their judgments. By assigning numerical values of 1–4 to these categories, we gave each of the 11 factors a Borda ranking as to their influence in respondents' judgments of journal quality. The results are shown in Table 12. The highest percentage for each factor is bolded and percentages adding to over 50% for one factor are in a shaded box. The first three factors listed proved very influential for more than 75% of the respondents, and the next three factors were very influential for nearly half of the respondents.

We also invited respondents to list other factors that had been influential in their relative quality ranking. Among the more popular factors provided by respondents was the quality of the peer-review process. For example, respondents felt that journals were of higher quality if they perceived that reviewers are rigorous and constructive in their feedback. They also recognized the importance of editorial teams that work closely with authors to improve the articles, both with respect to shepherding authors through the revision process and through quality editing in preparing the final version for print. A number of other responses referred to more personal indicators of quality. For example, respondents indicated that they considered whether they found the articles to be readable, interesting, and relevant to their current work as indicators of high quality. Finally, some responses were less about the quality of journal articles and more about the quality of the journal itself (e.g., journal structure, online presence, reliable publication timetable).

Table 12
Percent of Respondents Ranking Each Factor in Categories of Importance

Factor	Mean Borda score	Percent of respondents ranking each factor in categories of importance (*n* = 623)			
		Very influential	Somewhat influential	Of minor influence	Not considered
Most articles of high quality	3.88	**90.9**	7.4	0.6	1.1
Peer reviewed	3.76	**85.1**	10.6	1.1	3.2
High reputation among colleagues and experts	3.70	**76.1**	19.4	2.7	1.8
Cutting-edge research	3.39	**57.9**	28.9	7.5	5.6
Broad distribution	3.31	**51.0**	35.3	7.2	6.4
Editorial board leaders in field	3.23	**49.0**	33.5	8.7	8.8
Three or more reviews per manuscript	2.97	**40.3**	33.1	9.6	17.0
Wide range of topics	2.77	21.2	**44.8**	23.6	10.4
Ranks well on objective factors	2.60	23.8	**34.5**	19.9	21.8
Low acceptance rate	2.52	17.5	**37.9**	23.8	20.9
Specialized range of topics	2.28	7.9	32.3	**39.5**	20.4

Summary and Conclusion

Summary and Comparison to Other Studies

Tables 2, 3, and 8 summarize the major results of our studies and suggest that *JRME* and *ESM* are the two most cited and respected journals in our field by a substantial margin. They can reasonably be called Very High Quality journals. Five other journals (*JMB*, *ZDM*, *JMTE*, *MTL*, and *FLM*) compose a strong core of High Quality journals. These results echo those of Study F and compare very favorably with results from Studies A, B, C, D, and E (see Table 13). Thus, our studies provide empirical support for what is likely the conventional wisdom in our field regarding the best journals. The other 13 journals fell naturally into two other categories: Medium-High Quality and Medium Quality. Study 2 suggests that the journals in each of these four categories remain fairly stable across variations in respondents' region, rank, experience in the field, and personal

Table 13
Ranks of Journals from Studies 1 and 2 Compared With Ranks From Other Studies

	Citation based				Opinion based			
Journal	Study 1 (total citations)	Study A (Dreyfus, 2006a)	Study B (Dreyfus, 2006b)	Study C (Williams, 2008)	Study 2 (Borda scores)	Study D (Williams, 2008)	Study E (Holbrook et al., 2009)	Study F (Toerner & Arzarello, 2012)
JRME	2	2	2	1	1	1	1	2
ESM	1	1	1	2	2	2	3	1
JMB	3	4	3	4	3	5	5	4
ZDM	6	8			4			7
JMTE	7	6	7	6	5	3	6	5
MTL	5	5	5	5	6	4	7	6
FLM	4	3	4	3	7	6	4	3
RME	15				8	9		11
MERJ	8	10	6	8	9	7	2	10
IJSME	11				10	10		9
IJMEST	9	9		10	11	8	8	8
CJSMTE	18				12			12
IJTME	16				13		9	
SSM	10			9	14	12		
JCMST	17				15			
TME	14				16	15	12	
IML	12	7		7	17	11	10	
PRIMUS	19				18	13		
TMA	13				19		11	
TMME	20				20	14		13

associations with individual journals. However, these factors can affect the relative rank of individual journals within the broad quality categories.

At the same time, our two studies provided some additional interesting insights.

1. Although we were able to identify 55 English-language journals publishing extensively or exclusively in mathematics education, we found that a small handful of these journals account for the vast majority of citations (see Table 3). Indeed, when self-citations were excluded, the top seven journals account for roughly 80% of the citations we tallied.
2. Study 1 examined the 10 most frequently cited journals, taking into account both citations and number of published manuscripts, and created an MIF (see Table 3). The MIF scores retained the grouping of the top seven journals but differentiated between the top two journals, with *JRME* receiving nearly twice as many citations per manuscript published as *ESM* or *MTL*. This is one example of how citation-based and opinion-based studies complement one another and help to provide a fuller picture of journal quality.
3. We were somewhat surprised to find that on average, respondents in our survey were familiar with only 13 of the 20 journals in our sample. Lack of familiarity with journals was also noted by Toerner and Arzarello (2012), who reported that two thirds of their respondents were unfamiliar with 11 journals from their initial sample of 28. Studies in other fields have also shown a lack of familiarity with journals among respondents (e.g., Sellers, Mathiesen, Perry, & Smith, 2004).
4. It is significant that all 20 journals were rated at least Medium Quality by our respondents. Whether this is the result of optimism bias or reflects the state of our field is not clear. There is some evidence that mathematics education journals are held in high esteem among education journals in general. The PAJE database of educational journals (Fairbairn et al., 2008, Holbrook et al., 2009) ranks six of our seven top journals in the top 10% of all educational journals.
5. Respondents' rankings of about half of the journals were affected by their region, although the effect sizes were generally small. Other studies of education journal quality have found similar and even stronger effects (e.g., Wellington & Torgerson, 2005). Our respondents also showed a tendency to rank journals higher if they had an association with the journal. Editorial team and board members gave "their" journals more favorable rankings for half of the journals in the sample. Those respondents who had published in the journal in the previous 5 years also gave more favorable rankings (for 16 of the 20 journals). Here again, other studies have found similar results (e.g., Axarloglou & Theoharakis, 2003, in economics). Finally, time in the field made a difference in the rankings for only five of the 20 journals.

Limitations and Future Research

Although we feel confident about the four quality categories suggested by our data, we do not feel that the data support much inference about relative quality of journals within the categories. Indeed, Study 1 was unable to distinguish among the 13 least frequently cited journals. Perhaps a larger sample of citations from across all 20 journals would have provided Study 1 sufficient power to yield results for the lower category.

In a similar vein, a variation of Study 2 that gives respondents a random subset of the 55 journals, or gives arbitrary pairings and asks for relative rankings of those pairs, could allow for survey data to give us opinion ratings of all 55 journals without requiring the unrealistic expectation of rank ordering all 55 journals. Periodic surveys of this sort might be needed to correlate with citation study data.

We also wish to emphasize that neither study provides information about the 35 journals that were not in our group of 20. This includes relatively new journals as well as those that specialize in particular issues within our field. There could easily have been relatively new journals that are of high quality but that had not yet become sufficiently established to have amassed many citations in our data set.

References

Aase, S. (2008). What weight does impact factor carry? *Journal of the American Dietetic Association*, *108*(10), 1604–1607. doi:10.1016/j.jada.2008.08.005

Albrecht, C., Thompson, J. A., Hoopes, J. L., & Rodrigo, P. (2010). Business ethics journal rankings as percieved by business ethics scholars. *Journal of Business Ethics*, *95*(2), 227–237. doi:10.1007/s10551-009-0356-x

Althouse, B. M., West, J. D., Bergstrom, C. T., & Bergstrom, C. T. (2009). Differences in impact factor across fields and over time. *Journal of the Association for Information Science and Technology*, *60*(1), 27–34. doi:10.1002/asi.20936

Auranen, O., & Nieminen, M. (2010). University research funding and publication performance—An international comparison. *Research Policy*, *39*(6), 822–834. doi:10.1016/j.respol.2010.03.003

Axarloglou, K., & Theoharakis, V. (2003). Diversity in economics: An analysis of journal quality perceptions. *Journal of the European Economic Association*, *1*(6), 1402–1423. doi:10.1162/154247603322752584

Buela-Casal, G., & Zych, I. (2010). Analysis of the relationship between the number of citations and the quality evaluated by experts in psychology journals. *Psicothema*, *22*(2), 270–275.

Cooper, R. B., Blair, D. C., & Pao, M. L. (1993). Communicating MIS research: A citation study of journal influence. *Information Processing & Management*, *29*(1), 113–127. doi:10.1016/0306-4573(93)90027-B

de Borda, J.-C. (1781). *Memoire sur les elections au scrutin* [A note on polling elections]. Paris, France: Histoire de l'Académie Royale des Sciences.

Dreyfus, T. (2006a). *Citations of mathematics education research journals 2004–2006*. Unpublished manuscript.

Dreyfus, T. (2006b). *Refereed mathematics education research journals cited in PME proceedings*. Unpublished manuscript.

Emerson, P. (2013). The original Borda count and partial voting. *Social Choice and Welfare*, *40*(2), 353–358. doi:10.1007/s00355-011-0603-9

Fairbairn, H., Holbrook, A., Bourke, S., Preston, G., Cantwell, R., & Scevak, J. (2008, November–December). *A profile of education journals*. Paper presented at the annual conference of the Australian Association for Research in Education, Brisbane, Australia.

Hamermesh, D. S., Johnson, G. E., & Weisbrod, B. A. (1982). Scholarship, citations and salaries: Economic rewards in economics. *Southern Economic Journal*, *49*(2), 472–481. doi:10.2307/1058497

Hirsch, J. E. (2005). An index to quantify an individual's scientific research output. *Proceedings of the National Academy of Sciences*, *USA*, *102*(46), 16569–16572. doi:10.1073/pnas.0507655102

Holbrook, A., Bourke, S., Fairbairn, H., Preston, G., Cantwell, R., & Scevak, J. (2009). *Publishing in academic journals in education* [CD]. Melbourne, Australia: Professional Resources Services.

Jarwal, S. D., Brion, A. M., & King, M. L. (2009). Measuring research quality using the journal impact factor, citations and "ranked journals": Blunt instruments or inspired metrics? *Journal of Higher Education Policy and Management*, *31*(4), 289–300. doi:10.1080/13600800903191930

Jacso, P. (2006). Deflated, inflated and phantom citation counts. *Online Information Review*, *30*(3), 297–309. doi:10.1108/14684520610675816

Jacsó, P. (2008a). The pros and cons of computing the h-index using Google Scholar. *Online Information Review*, *32*(3), 437–452. doi:10.1108/14684520810889718

Jacsó, P. (2008b). The pros and cons of computing the h-index using Scopus. *Online Information Review*, *32*(4), 524–535. doi:10.1108/14684520810897403

Jacsó, P. (2012). Using Google Scholar for journal impact factors and the h-index in nationwide publishing assessments in academia—Siren songs and air-raid sirens. *Online Information Review*, *36*(3), 462–478. doi:10.1108/14684521211241503

Kokko, H., & Sutherland, W. J. (1999). What do impact factors tell us? *Trends in Ecology & Evolution*, *14*(10), 382–384. doi:10.1016/S0169-5347(99)01711-5

Kurmis, A. P. (2003). Understanding the limitations of the journal impact factor. *The Journal of Bone & Joint Surgery*, *85*(12), 2449–2454. doi:10.2106/00004623-200312000-00028

Lester, F. K. (2005). On the theoretical, conceptual, and philosophical foundations for research in mathematics education. *ZDM*, *37*(6), 457–467. doi:10.1007/BF02655854

Leydesdorff, L. (2009). How are new citation-based journal indicators adding to the bibliometric toolbox? *Journal of the Association for Information Science and Technology*, *60*(7), 1327–1336. doi:10.1002/asi.21024

Lowry, P. B., Moody, G. D., Gaskin, J., Galletta, D. F., Humpherys, S. L., Barlow, J. B., & Wilson, D. W. (2013). Evaluating journal quality and the Association for Information Systems Senior Scholars' Journal Basket via bibliometric measures: Do expert journal assessments add value? *MIS Quarterly*, *37*(4), 993–1012.

Manning, L. M., & Barrette, J. (2005). Research performance management in academe. *Canadian Journal of Administrative Sciences*, *22*(4), 273–287. doi:10.1111/j.1936-4490.2005.tb00374.x

Meho, L. I., & Yang, K. (2007). Impact of data sources on citation counts and rankings of LIS faculty: Web of Science versus Scopus and Google Scholar. *Journal of the Association for Information Science and Technology*, *58*(13), 2105–2125. doi:10.1002/asi.20677

Mingers, J., & Harzing, A.-W. (2007). Ranking journals in business and management: A statistical analysis of the Harzing data set. *European Journal of Information Systems*, *16*(4), 303–316. doi:10.1057/palgrave.ejis.3000696

Mittal, V., Feick, L., & Murshed, F. (2008). Publish and prosper: The financial impact of publishing by marketing faculty. *Marketing Science*, *27*(3), 430–442. doi:10.1287/mksc.1080.0361

Nivens, R. A., & Otten, S. (2017). Assessing journal quality in mathematics education. *Journal for Research in Mathematics Education, 48*(4), 348–368.

Norris, M., & Oppenheim, C. (2007). Comparing alternatives to the Web of Science for coverage of the social sciences' literature. *Journal of Informetrics*, *1*(2), 161–169. doi:10.1016/j.joi.2006.12.001

Oswald, A. J. (2007). An examination of the reliability of prestigious scholarly journals: Evidence and implications for decision-makers. *Economica*, *74*(293), 21–31. doi:10.1111/j.1468-0335.2006.00575.x

Roales-Nieto, J. G., & O'Neill, B. (2012). A comparative study of journals quality based on Web of Science, Scopus and Google Scholar: A case study with IJP&PT. *International Journal of Psychology & Psychological Therapy*, *12*(3), 453–479.

Schoenfeld, A. H. (2007). Method. In F. K. Lester (Ed.), *Second handbook of research on mathematics teaching and learning* (Vol. 1, pp. 69–107). Charlotte, NC: Information Age.

Seglen, P. O. (1997). Why the impact factor of journals should not be used for evaluating research. *British Medical Journal*, *314*(7079), 498–502. doi:10.1136/bmj.314.7079.497

Sellers, S. L., Mathiesen, S. G., Perry, R., & Smith, T. (2004). Evaluation of social work journal quality: Citation versus reputation approaches. *Journal of Social Work Education*, *40*(1), 143–160.

Serenko, A., & Dohan, M. (2011). Comparing the expert survey and citation impact journal ranking methods: Example from the field of Artificial Intelligence. *Journal of Informetrics*, *5*(4), 629–648. doi:10.1016/j.joi.2011.06.002

Sheehan, K. B. (2001). Email survey response rates: A review. *Journal of Computer-Mediated Communication*, *6*(2). doi:10.1111/j.1083-6101.2001.tb00117.x

Starbuck, W. H. (2005). How much better are the most-prestigious journals? The statistics of academic publication. *Organization Science*, *16*(2), 180–200. doi:10.1287/orsc.1040.0107

Toerner, G., & Arzarello, F. (2012, December). Grading mathematics education research journals. *Newsletter of the European Mathematical Society*, *86*, 52–54.

Wellington, J., & Torgerson, C. J. (2005). Writing for publication: What counts as a "high status, eminent academic journal"? *Journal of Further and Higher Education*, *29*(1), 35–48. doi:10.1080/03098770500037739

Williams, S. R. (2008). *Report on venue study*. Unpublished manuscript, Department of Mathematics Education, Brigham Young University, Provo, UT.

Authors

Steven R. Williams, Department of Mathematics Education, Brigham Young University, 193B TMCB, Provo, UT 84602; williams@mathed.byu.edu

Keith R. Leatham, Department of Mathematics Education, Brigham Young University, 193A TMCB, Provo, UT 84602; kleatham@mathed.byu.edu

Submitted April 26, 2016

Accepted September 23, 2016

APPENDIX

Table A1

Alphabetical List of 55 Journals That Publish (in English) Either Exclusively or Extensively in Mathematics Education Research

Journal
Adults Learning Mathematics – An International Journal
African Journal of Educational Studies in Mathematics and Sciences
African Journal of Research in MST Education
Canadian Journal of Science, Mathematics and Technology Education
Contemporary Issues in Technology and Teacher Education
Educational Studies in Mathematics
Electronic Journal of Mathematics and Technology
Eurasia Journal of Mathematics, Science & Technology Education
Far East Journal of Mathematical Education
For the Learning of Mathematics
Hiroshima Journal of Mathematics Education
Indonesian Mathematical Society Journal on Mathematics Education
International Journal of Math Education in Science and Technology
International Electronic Journal of Mathematics Education
International Journal for Mathematics Teaching and Learning
International Journal for Studies in Mathematics Education
International Journal for Technology in Mathematics Education (formerly International Journal of Computer Algebra in Mathematics Education)
International Journal for the History of Mathematics Education
International Journal of Innovation in Science and Mathematics Education
International Journal of Mathematical Education: Policy and Practice
International Journal of Science and Mathematics Education
Investigations in Mathematics Learning (formerly FOCUS on Learning Problems in Mathematics)
Issues in the Undergraduate Mathematics Preparation of School Teachers: The Journal (IUMPST)
Journal for Research in Mathematics Education
Journal of Computers in Mathematics and Science Teaching
Journal of Mathematical Behavior
Journal of Mathematics and Culture
Journal of Mathematics Education
Journal of Mathematics Education Leadership
Journal of Mathematical Modelling and Application

Table A1 (continued)

Journal
Journal of Mathematics Teacher Education
Journal of Science and Mathematics Education in Southeast Asia
Journal of STEM (*Science, Technology, Engineering, and Mathematics*) *Education: Innovations and Research*
Journal of Urban Mathematics Education
MathAMATYC Educator
Mathematical Thinking and Learning
Mathematics and Computer Education
Mathematics Education Research Journal
Mathematics Education Review
Mathematics Teacher Education and Development
Mathematics Teacher Educator
Mathematics Teaching-Research Journal (online)
Mediterranean Journal for Research in Mathematics Education
MSOR Connections
Nordic Studies in Mathematics Education (*Nordisk Matematikk Didaktikk*)
Philosophy of Mathematics Education Journal
PRIMUS (*Problems, Resources, Issues in Undergraduate Mathematics Studies*)
Pythagoras (AMESA's research journal)
Research in Mathematics Education
School Science and Mathematics
Teaching Mathematics and Its Applications
The Mathematics Educator
The Montana Mathematics Enthusiast
Turkish Journal of Computer and Mathematics Education
Zentralblatt für Didaktik der Mathematik (*ZDM*) *The International Journal on Mathematics Education*