

# Are three heads better than two? How the number of reviewers and editor behavior affect the rejection rate

David M. Schultz

Received: 5 June 2009 / Published online: 13 August 2009  
© Akadémiai Kiadó, Budapest, Hungary 2009

**Abstract** Editors of peer-reviewed journals obtain recommendations from peer reviewers as guidance in deciding upon the suitability of a submitted manuscript for publication. To investigate whether the number of reviewers used by an editor affects the rate at which manuscripts are rejected, 500 manuscripts submitted to *Monthly Weather Review* during 15.5 months in 2007–2008 were examined. Two and three reviewers were used for 306 and 155 manuscripts, respectively (92.2% of all manuscripts). Rejection rates for initial decisions and final decisions were not significantly different whether two or three reviewers were used. Manuscripts with more reviewers did not spend more rounds in review or have different rejection rates at each round. The results varied by editor, however, with some editors rejecting more two-reviewer manuscripts and others rejecting more three-reviewer manuscripts. Editors described using their scientific expertise in the decision-making process, either in determining the number of reviews to be sought or in making decisions once the reviews were received, approaches that differ from that of relying purely upon reviewer agreement as reported previously in the literature. A simple model is constructed for three decision-making strategies for editors: rejection when all reviewers recommend rejection, rejection when any reviewer recommends rejection, and rejection when a majority of reviewers recommend rejection. By plotting the probability of reviewer rejection against the probability of editor rejection, the decision-making process can be graphically illustrated, demonstrating that, for this dataset, editors are likely to reject a manuscript when any reviewer recommends rejection.

**Keywords** Journal · Editor · Monthly Weather Review · Peer review · Rounds of reviews · Reviewer agreement

**PACS** 01.20

---

D. M. Schultz (✉)  
Division of Atmospheric Sciences and Geophysics,  
Department of Physics, University of Helsinki, and Finnish Meteorological Institute,  
Erik Palménin Aukio 1, P.O. Box 503, 00101 Helsinki, Finland  
e-mail: David.Schultz@fmi.fi

**Mathematical Subject Classification (2000)** 00A99

**JEL Codes** Z00

## Introduction

Debate rages about whether the peer-review process for scientific journals is fair. Can the peer reviewers of a scientific manuscript make an adequate assessment of its quality? Should authors alter their vision of the manuscript beholden to the whims of reviewers who may have only taken a cursory look at the manuscript or are less familiar with the material than the author? (e.g., Frey 2003) Can editors be trusted to make unbiased decisions?

As I am on the editorial boards of three scientific journals and the chief editor of another one (*Monthly Weather Review*), my concern about peer review is a practical one. How can I promote a fair system within the constraints of the peer-review system imposed by the publishers? The origin of this article came from a discussion with another chief editor who felt that it was harder to get papers published in his journal because his editors commonly used three reviewers, whereas editors at my journal were more likely to use two reviewers. He argued that more reviewers would be more likely to uncover fatal flaws with the manuscript, thus providing more guidance to the editor to reject the manuscript. A literature search turned up little on how the number of reviewers affects the rejection rate, so the topic seemed ripe for investigation.

The number of reviewers used in the review process is not to be considered lightly. One drawback of using a larger number of reviewers is finding enough available, qualified volunteers to perform reviews and requesting a larger number of reviews repeatedly from them. Such a continued demand on the best reviewers can burn out even the most eager ones. Another drawback to using a larger number of reviewers is the possible increase in the length of time a manuscript spends in the review process. Specifically, editors wait longer for more reviewers to submit their reviews, authors spend more time addressing a larger number of concerns, and manuscripts may undergo more rounds in review to satisfy the larger number of reviewers. On the other hand, perhaps the most important advantage to having more reviewers is the increased input that they provide to the authors to improve their manuscripts, especially if the reviewers come from different subdisciplines, different approaches (e.g., theoretical, modeling, observational), or different schools of thought. If improving the quality of scientific manuscripts is the *raison d'être* of peer review, then a larger number of reviewers should be paramount.

But, what does this increased breadth mean for the editor's decision-making process? Many scholars would argue that, to have validity, peer review—like any good scientific endeavor—should exhibit “reproducibility,” “reliability,” or “precision” (e.g., Lindsey 1988; Strayhorn et al. 1993; Ernst et al. 1993). Specifically, more reliable decisions are made when there is more agreement between reviewers, and more reviewers are likely to increase the likelihood of reviewer agreement (e.g., Lindsey 1988; O'Brien 1991). Others have argued that more reviewer agreement leads to lower rejection rates at journals (e.g., Zuckerman and Merton 1971; Hargens 1988; Eberley and Warner 1990; Hargens and Herting 1990).

Unfortunately, most studies of peer review have shown that the agreement between reviewers is often poor because of the large variability in the knowledge, experiences, and values of reviewers (e.g., as reviewed in chapter 6 of Weller 2001). For example, Ernst et al. (1993) received reviews on a single manuscript from 31 expert editors. These reviews

were quite disparate, with nearly all measures of manuscript quality receiving scores from the reviewers for the very best through the very worst rankings. Given such discrepancy, more reviewers providing comments on a manuscript give the editor greater confidence in getting a broader perspective from the scientific community with the hope of obtaining more agreement about the manuscript from the different reviewers. Indeed, O'Brien (1991) found that more reviewers resulted in more reliability, which could mean that editors should use three or more reviewers per manuscript, as advocated by Glenn (1976) and Lindsey (1988). Bornmann and Daniel (2009) found that including the recommendation of a late-arriving third review, after an editorial decision had already been made with two reviews, would have changed the editorial decision in 37 (23%) of the 162 manuscripts with a late third reviewer, or 2% of all 1,899 manuscripts in their dataset. Of those 37 manuscripts, 23 (62%) would have led to rejection of an otherwise accepted manuscript. On the other hand, in discussing Ernst et al. (1993), Weller (2001, pp. 182–183) said, “the range of [the reviewers’] recommendations indicates that submitting a manuscript for more review might not solve an editor’s dilemma.” Thus, whether using more reviewers leads to more reliable editorial decisions is unclear.

Of course, these arguments hinge upon whether editors actually do make decisions based on agreement between reviewers. Although sometimes editors may decide the fate of manuscripts based on reviewer agreement (democratically, by majority vote), oftentimes editors are swayed by the most critical reviews (e.g., those that identify the most severe flaws). Sometimes editors may follow most closely the recommendations of influential, respected, or knowledgeable reviewers, or they may disregard reviews that are biased in one way or another. Other times editors may seek out opposing viewpoints on a manuscript to weigh the relative arguments of each in making their decision. Even the same editor may treat different manuscripts differently, employing one decision-making process for some manuscripts, but using a different process for other manuscripts. Finally, two editors, when presented with the same sets of reviews on a manuscript, may make different decisions regarding its suitability for publication—one editor’s “major revisions” is another editor’s “reject.” Thus, arguments about the “reliability” of the peer-review process may not be relevant to the practical business of how most editors make decisions.

In a practical sense, knowing whether the number of reviewers affects rejection rate may provide guidance for managing a peer-review system for a journal. The purpose of this article is to use real data from the peer-review process of a scientific journal to address the question, “do editorial decisions differ whether they are made with the guidance from two or three reviewers?” As a byproduct of answering this question, the relative importance of reviewer agreement in editorial decision-making can also be assessed.

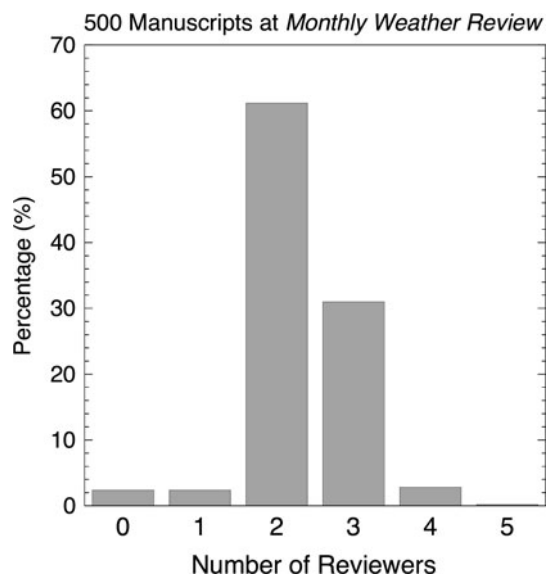
## Monthly Weather Review

To test whether the number of reviewers affects the rejection rate, we construct a dataset of manuscripts from the monthly meteorology journal *Monthly Weather Review* (*MWR*). *MWR* is the oldest continuously operating meteorology journal in the world. Initially published in 1873 by the United States government (first the Office of the Chief Signal Officer in the War Department, then the Weather Bureau, the Environmental Science Services Administration, and the National Oceanic and Atmospheric Administration), *MWR* was transferred to the nonprofit American Meteorological Society in 1974, where *MWR* eventually became one of the nine scientific journals currently published by the society. *MWR* was ranked 13th out of 51 journals in the category “Meteorology & Atmospheric Sciences” by the Institute of

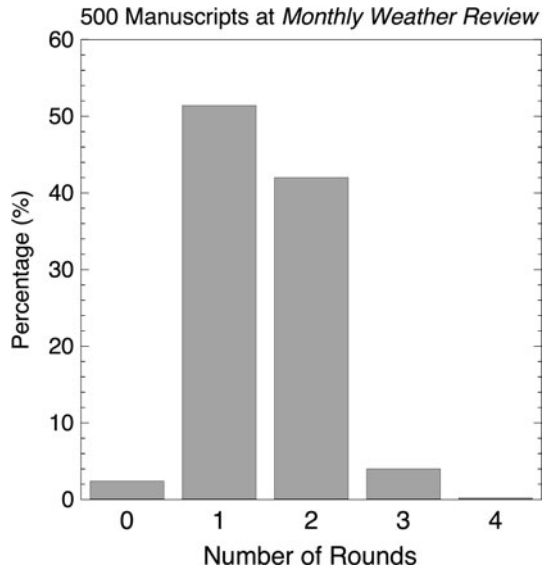
Scientific Information Web of Knowledge, Thomson Reuters, Inc. (<http://www.isiwebofknowledge.com>) in 2007 with an impact factor of 2.267, a 5-year impact factor of 3.061, an immediacy index of 0.558, a cited half-life of >10 years, and 242 articles published. *MWR* receives about 400 submissions a year, of which 34% were rejected in 2006, which is nearly the 37% average for journals that publish articles on atmospheric science (Schultz 2009). Authors are assessed page charges to publish in *MWR*, although more than 5 years after being published, the articles become open access and are available for free online (Seitter 2002). Jorgensen et al. (2007) discussed the publishing process at American Meteorological Society journals, although a brief description of the peer-review process at *MWR* follows.

To submit an article to *MWR*, the author uploads the manuscript file to the Manuscript Tracking System. After an initial evaluation process at American Meteorological Society Headquarters to check for proper formatting, the manuscript is sent to the *MWR* chief editor to decide on the one editor (out of eight or nine at any given time at *MWR*) who will handle the manuscript. The editor is then responsible for finding peer reviewers and making a decision on the manuscript. At *MWR*, the editors have some flexibility in making decisions, with most editors seeking the advice of two or three reviewers (61.2 and 31.0% of all manuscripts, respectively), although one, four, and five reviewers have been used in the past (Fig. 1). Editors have even made decisions on manuscripts without sending them out for peer review (zero reviewers, 12 times or 2.4% of the manuscripts) (Fig. 1). The authors of manuscripts are known to the reviewers, but, in general, the reviewers remain anonymous to the authors. After reviews are sent to the editor, the editor makes an *initial decision* (accept as is, minor revisions, major revisions, reject, transfer), and the reviews and the editor's decision are sent to the author. If the initial decision called for minor revisions and the author resubmits a revised manuscript, the editor usually decides whether the author addressed the reviewers' concerns, and the manuscript tends to be accepted. If the initial decision called for major revisions and the author resubmits a revised manuscript, the editor usually sends the manuscript back to reviewers who recommended major revisions or rejection for additional rounds of reviews. After two or more rounds of reviewer comments, the manuscript may be accepted or rejected, the editor's *final decision*. At any point, the author may also withdraw

**Fig. 1** Histogram of the number of reviewers per manuscript



**Fig. 2** Histogram of the number of rounds of reviews before final decision



the manuscript from consideration. Rejected manuscripts are sometimes resubmitted and treated as new submissions, eventually being accepted or again rejected. Most manuscripts go through one or two rounds (51.4 and 42.0% of all manuscripts, respectively), although as many as four rounds have happened (Fig. 2).

## Data and methods

The dataset for this article consists of 500 manuscripts submitted to *MWR*, with the dates that decisions arrived at the publisher from 13 January 2007 to 2 May 2008, or a period of 15.5 months. Data for each manuscript included the editor's name (of which there were 12 during this time, including four new editors with less than six manuscripts each and one outgoing editor with less than 30 manuscripts), the number of rounds of reviews, the number of reviewers in the first round, the initial decisions by the reviewers in the first round, and the final decision by the editor. Reviewers that agreed to provide a review, but never delivered it, were not counted. The dataset was obtained as XML output from the Manuscript Tracking System, which was then converted into tabular data to import into a spreadsheet application.

The full dataset consisted of 500 manuscripts. For some analyses in this article, the focus is only on those manuscripts that had two reviewers (306 manuscripts) or three reviewers (155 manuscripts). This subset amounts to 461 (92.2%) of the 500 manuscripts in the dataset. For other analyses, manuscripts with a final decision of “transfer” or “withdraw” are excluded from the 500 (regardless of the number of reviewers), resulting in a reduced dataset of 482 manuscripts.

## Results

Table 1 compares the initial decisions for manuscripts with different numbers of reviewers. Because of the small number of manuscripts with zero, one, four, and five reviewers, the

**Table 1** Initial decisions for all 500 manuscripts partitioned by the number of reviewers for each manuscript: percentage of decisions for each initial decision

	Zero reviewers	One reviewer	Two reviewers	Three reviewers	Four reviewers	Five reviewers	Average
<i>N</i>	12	12	<b>306</b>	<b>155</b>	14	1	<i>500</i>
Accept (%)	25.0	8.3	<b>0.3</b>				<i>1.0</i>
Major revisions (%)	8.3	8.3	<b>46.7</b>	<b>49.7</b>	50.0		<i>45.8</i>
Minor revisions (%)		50.0	<b>25.8</b>	<b>23.9</b>	28.6		<i>25.2</i>
Reject (%)	41.7	25.0	<b>26.5</b>	<b>25.8</b>	21.4	100	<i>26.6</i>
Transfer (%)	8.3	8.3		<b>0.6</b>			<i>0.6</i>
Withdraw (%)	16.7		<b>0.7</b>				<i>0.8</i>

*N* is the number of manuscripts in each category. Values in *bold* represent an emphasis on these two columns as discussed in the text, and values in *italics* represent the average for the whole dataset

focus is on manuscripts with two and three reviewers (values in *bold*). The percentages of manuscripts in each initial-decision category are consistent regardless of whether two or three reviewers were used, with a slight tendency for larger percentages for major revisions and smaller percentages for minor revisions and rejections for manuscripts with three reviewers (Table 1). In particular, the rejection rates for manuscripts with two reviewers and three reviewers at initial decision were 26.5 and 25.8%, respectively. Even if this result were statistically significant, this 0.7% difference between rejection rates amounts to only 1–2 manuscripts, a result that would be difficult to identify in practice given the vagaries of the review process.

The results are similar when the final decisions of the manuscripts are considered: consistent percentages in each category regardless of whether two or three reviewers are used (Table 2). Indeed, the percentage of rejections is only 0.3% less for manuscripts with two reviewers than those with three reviewers (29.4 vs. 29.7%). These results indicate that the rejection rate for initial or final decisions is not substantially related to the number of number of reviewers.

Although the decisions at the journal as a whole may not depend on whether two or three reviewers are used, perhaps differences between individual editors offset each other. To test for editor-dependent behavior, the final decisions on two- and three-reviewer manuscripts were segregated by editor, and the rejection rates presented in Table 3. (Final decisions, rather than initial decisions, were considered to ensure the largest number of rejections, although the results were similar.) Data are presented for individual editors if

**Table 2** Final decisions for all 500 manuscripts partitioned by the number of reviewers for each manuscript: percentage of decisions for each final decision

	Zero reviewers	One reviewer	Two reviewers	Three reviewers	Four reviewers	Five reviewers	Average
<i>N</i>	12	12	<b>306</b>	<b>155</b>	14	1	<i>500</i>
Accept (%)	25.0	66.7	<b>68.0</b>	<b>67.1</b>	78.6		<i>66.8</i>
Reject (%)	41.7	25.0	<b>29.4</b>	<b>29.7</b>	21.4	100	<i>29.6</i>
Transfer (%)	8.3	8.3		<b>1.3</b>			<i>0.8</i>
Withdraw (%)	25.0		<b>2.6</b>	<b>1.9</b>			<i>2.8</i>

*N* is the number of manuscripts in each category. Values in *bold* represent an emphasis on these two columns as discussed in the text, and values in *italics* represent the average for the whole dataset

**Table 3** Rejection rate (%) and number of manuscripts (*N*) by editor (including transfer and withdrawal) for the 461 manuscripts with two or three reviewers

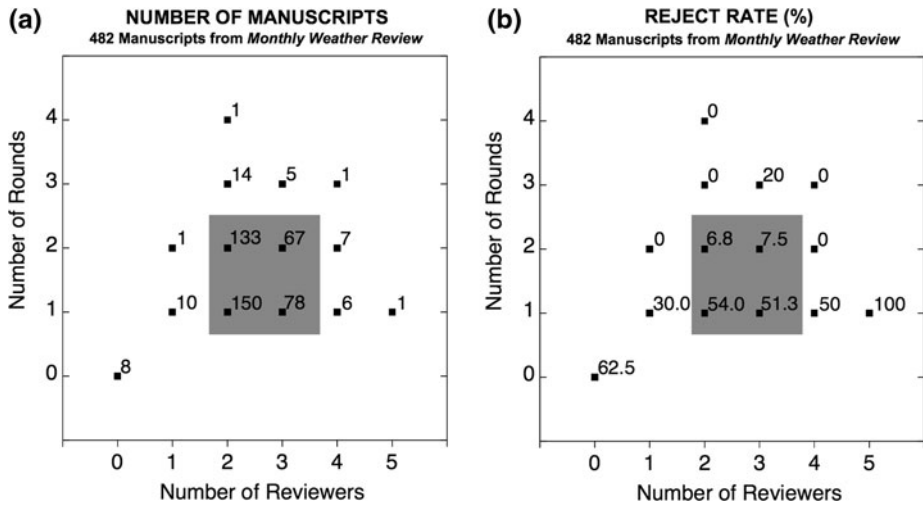
	Two reviewers		Three reviewers	
	Reject rate (%)	<i>N</i>	Reject rate (%)	<i>N</i>
<i>Editor A</i>	25.9	27	15.8	19
<b>Editor B</b>	<b>18.7</b>	<b>32</b>	<b>25.0</b>	<b>44</b>
<i>Editor C</i>	40.9	22	24.1	29
<b>Editor D</b>	<b>42.4</b>	<b>33</b>	<b>46.3</b>	<b>41</b>
Editors E, F, G	27.5	160	28.5	7
Other Editors (5)	31.3	32	26.7	15
Total	29.4	306	29.7	155

Values in *bold* represent editors for whom the reject rate for manuscripts with two reviewers is less than the reject rate for manuscripts with three reviewers. Values in *italics* represent editors for whom the reject rate for manuscripts with two reviewers is greater than the reject rate for manuscripts with three reviewers. The row labeled “Other Editors (5)” represents the sum of manuscripts from five editors who made less than 30 decisions each. The row labeled “Editors E, F, G” represents the sum of manuscripts from three editors who strongly prefer obtaining two reviewers

the editors had made decisions on more than 30 manuscripts (Editors A, B, C, and D). Five editors had not, so their statistics were combined into the row “Other Editors (5).” Three editors had a very strong preference for two reviewers (amassing 160 two-reviewer manuscripts but only 7 three-reviewer manuscripts), so their statistics are combined into the row “Editors E, F, G.” Rows in Table 3 are *bold* if two-reviewer manuscripts were rejected less than three-reviewer manuscripts (Editors B and D) and *italics* if two-reviewer manuscripts were rejected more than three-reviewer manuscripts (Editors A and C).

When interviewed about why their decisions might depend on the number of reviewers, Editors B and D admitted to similar approaches. They surmised that when they received a manuscript, they gave it an initial evaluation. They usually tried for three reviewers, but, if the manuscript seemed uncontroversial and well written (a “slam dunk”), then two experienced reviewers would suffice. Thus, two-reviewer manuscripts were more likely to get accepted than three-reviewer manuscripts because they had an initially favorable opinion by the editor. If Editors B and D thought that the manuscript might have a hard time being accepted, one reviewer was inexperienced, or a reviewer might not or did not submit a review, they tried harder to get three (or more) reviewers to ensure broader, more, and timely guidance. Thus, editors used their knowledge of whether a manuscript is likely to be successful in the review process to aid in determining the number of reviewers.

In contrast, Editors A and C had higher rejection rates for two-reviewer manuscripts than for three-reviewer manuscripts. Editor C was more likely to read a manuscript more carefully for which only two reviewers could be found to better inform a decision. Because some reviewers are reluctant to reject a manuscript that contains little new scientific insight but is well written, Editor C was probably more vigilant in rejecting such two-reviewer manuscripts. On the other hand, Editor A did not make a conscious connection between the number of reviewers and rejection rates. Put differently, if one of Editor A’s two-reviewer manuscripts that was rejected was converted to an accept decision and one three-reviewer manuscript that was accepted was converted to a reject decision, then Editor A would have had comparable reject rates regardless of the number of reviewers (22.2% for two-reviewer manuscripts vs. 21.1% for three-reviewer manuscripts). Thus, changing only two decisions



**Fig. 3** Number of reviewers versus number of rounds to final decision (excluding manuscripts with transfer and withdraw decisions): **a** Number of manuscripts in each category; **b** percentage of rejections in each category. *Gray-shaded area* emphasizes data discussed in the text

out of 46 would be enough to achieve parity in rejection rates for Editor A, hardly a robust result.

Does using more reviewers mean that manuscripts spend more rounds in peer review? Figure 3a shows the number of manuscripts (excluding transfers and withdrawals) plotted on a graph of the number of reviewers versus the number of rounds to final decision. The most common categories (gray shaded in Fig. 3a) were two and three reviewers with one and two rounds, amounting to 88.8% of the 482 manuscripts. For manuscripts with two reviewers, 50.3% went through one round of review and 44.6% went through two rounds. In comparison, for manuscripts with three reviewers, then 52.0% went through one round of review and 44.7% went through two rounds. Thus, more reviewers did not prolong the review process by increasing the number of rounds of reviews.

When rejection rates for these different categories are compared, no large differences between two- or three-reviewer manuscripts are apparent either (Fig. 3b). Specifically, of manuscripts that went through one round of review, 54.0% of two-reviewer manuscripts were rejected compared to 51.3% of three-reviewer manuscripts. Likewise, of manuscripts that went through two rounds of reviews, 6.8% of two-reviewer manuscripts were rejected compared to 7.5% of three-reviewer manuscripts. These differences are not of practical significance.

### A simple model of editor decision-making

To gain some insight into how the number of reviewers and the rejection rate varies for different decision-making processes by editors, consider the following simple model. For the purposes of this model, we consider that reviewers and editors make one of two decisions “reject” or “accept,” and only one round of reviews occurs. [“Accept” could be changed to “not reject” (accept or revisions required) and the number of rounds could be increased with no loss of generality.] Assume that the probability of rejection by any given

reviewer is  $r$ , the probability of acceptance is  $1 - r$ , and manuscripts likely to be rejected are evenly distributed among the editors.

Consider three different editorial strategies—simplified, but reasonable, approaches to decision-making at a journal. Editor Omnium (*omni* is Latin for “all”) only rejects manuscripts when all reviewers recommend rejection (i.e., accepts manuscripts if any reviewer recommends acceptance). This strategy favors authors, who are allowed to revise their manuscripts rather than having them be rejected. Editor Quisius (*quis* is Latin for “anyone”) always rejects manuscripts if any of the reviewers recommends rejection, similar to the “go with the low” philosophy described by Cicchetti (1991). This strategy shifts the advantage away from the authors, who must avoid any reviewer recommendations of rejection to get their manuscript published. Editor Populus (Latin for “people”) makes a democratic decision (majority rules), but if two reviewers are used and split their recommendations (i.e., one accept and one reject), then Editor Populus requests a third review to break the tie, similar to the “go with the mode” philosophy described by Cicchetti (1991) that favors reviewer agreement.

To calculate the probability of editor rejection based on these three different strategies, the probability of all possible reviewer outcomes is determined. For example, for two reviewers, the two reviews could both be reject, occurring with probability  $r^2$ ; both be accept, occurring with probability  $(1 - r)^2$ ; or be one reject and one accept, occurring with probability  $2r(1 - r)$ . The probabilities for three reviewers are similarly calculated. The probabilities of editor rejection for each of the different strategies are then computed for two and three reviewers as a function of  $r$  (Table 4).

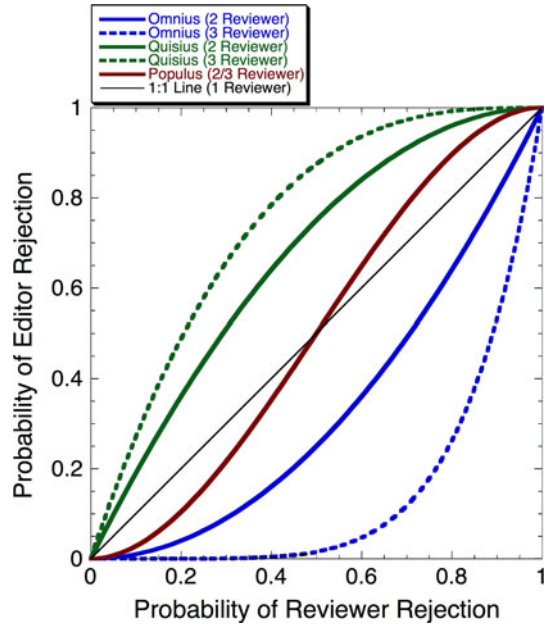
Plotting these algebraic expressions from Table 4 illustrates how the probability of the editor rejecting a manuscript varies as a function of the probability of reviewer rejection (Fig. 4). The different curves represent the three different editorial strategies for two- and three-reviewer manuscripts. Because Editor Omnium requires all reviewers to recommend rejection before rejecting the manuscript, the curves for Editor Omnium are concave up implying that the probability of editor rejection is less than the reviewer rejection (except at  $r = 0$  and  $r = 1$  where they are equal). The probability of Editor Omnium rejecting a manuscript is uniformly less for three-reviewer manuscripts than for two-reviewer manuscripts (Fig. 4), indicating the added difficulty of requiring agreement between three reviewers as opposed to two.

**Table 4** Probability of editor rejection according to three simple strategies

Editor	Strategy	Probability of editor rejection for two-reviewer manuscripts	Probability of editor rejection for three-reviewer manuscripts
Editor Omnium	Rejects when all reviewer recommend rejection	$r^2$	$r^3$
Editor Quisius	Rejects when any reviewer recommends rejection	$1 - (1 - r)^2$ $= 2r - r^2$	$1 - (1 - r)^3$ $= 3r - 3r^2 + r^3$
Editor Populus	Rejects when a majority recommends rejection (seeks out third reviewer if two reviewers are split)	$r^2 + 2r^2(1 - r)$ $= 3r^2 - 2r^3$	$3r^2(1 - r) + r^3$ $= 3r^2 - 2r^3$

The probability of any individual reviewer recommending rejection is  $r$

**Fig. 4** Probability of reviewer rejection versus probability of editor rejection for different editor decision strategies (Editors Omnium, Quisius, and Populus, described in Table 4) and the 1:1 diagonal line



The opposite is true for Editor Quisius who requires just one reviewer rejection to reject the manuscript. In this case, the curves are concave down implying that the probability of editor rejection is greater than the probability of reviewer rejection, and the curve for the three-reviewer manuscripts is uniformly higher than that for two-reviewer manuscripts (Fig. 4), indicating the higher likelihood that a third reviewer will recommend rejection.

Finally, the two curves for Editor Populus are equivalent and lie in between those of Editors Omnium and Quisius (Fig. 4). Editor Populus's curve is concave up for  $r < 0.5$ , where it is more similar to Editor Omnium's curves, and concave down for  $r > 0.5$ , where it is more similar to Editor Quisius's curves.

A 1:1 diagonal line in Fig. 4 represents the probability of editor rejection for a one-reviewer manuscript if the editor took the reviewer's recommendation verbatim [Editor Unius?, similar to the "go with the flow" philosophy described by Cicchetti (1991)]. Consequently, the increase in the number of reviewers from one to two has more importance to Editor Quisius's decision-making than the increase in the number of reviewers from two to three, as demonstrated by the larger difference between the 1:1 line and Editor Quisius's two-reviewer curve than between Editor Quisius's two-reviewer and three-reviewer curves. [That the addition of the third reviewer is less important to the editor's decision-making than the addition of the second reviewer is reminiscent of the same effect in reviewer reliability noted by O'Brien (1991).] Such a result is not duplicated for Editor Omnium (Fig. 4). In contrast, the number of reviewers does not affect Editor Populus's decision making. Thus, editors following Editor Omnium's or Editor Quisius's approaches will necessarily have a difference in rejection rates depending on the number of reviewers. A corollary is that an editor with parity between rejection rates for manuscripts with different numbers of reviewers must be deviating from a fixed strategy or following majority rules.

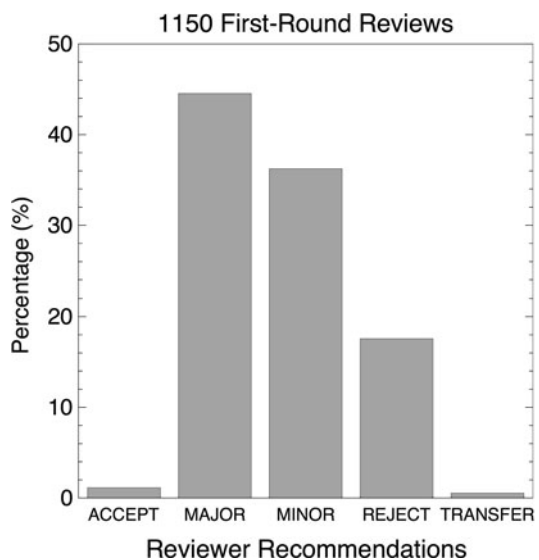
## Applying the model to real data from *MWR*

To apply this graph to real data, the probability of reviewer rejection  $r$  is needed. For this dataset, of the 1,150 reviewer recommendations in the first round of reviews for the dataset of 500 manuscripts, most of the reviewers recommended major (44.5%) or minor (36.3%) revisions (Fig. 5). Only 17.6% of reviewers recommended rejection, and 0.5% recommended transfer (Fig. 5). For the purposes of the model, rejection and transfer are combined for *MWR*, yielding an 18% probability of reviewer rejection ( $r = 0.18$ ). Individual editors may deviate from this 18%, depending on the quality of the papers they handle and the rigor to which their selected reviewers hold manuscripts. Indeed, in this study, 14.0% of Editor A's first-round reviews on two-reviewer manuscripts were recommendations for rejection, whereas 27.3% of Editor C's and Editor D's first-round reviews on two-reviewer manuscripts were recommendations for rejection.

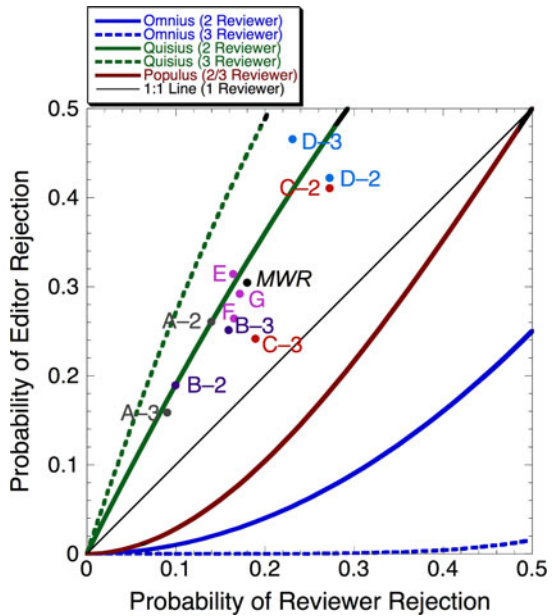
To show detail better, the lower-left quadrant of Fig. 4 is enlarged to create Fig. 6. A point, labeled *MWR*, is located at the 18% reviewer rejection rate and the 30.4% editor rejection rate for manuscripts from this dataset. This value of 30.4% is the sum of the average reject and transfer rates at final decision in Table 2. (The smaller rejection rate for this dataset compared to the 34% average in 2006 is because four editors were winding down their editorship during this period so they were tending to make more final "accept" decisions on manuscripts that they had been handling for a long time, which results in a reduced rejection rate compared to the long-term average.) The average rejection rate for this dataset lies near the two-reviewer curve for Editor Quisius, suggesting that the editors as a whole tended to reject manuscripts rather readily when reviewers recommended rejection.

Had the point for *MWR* occurred closer to Editor Populus's curve (about 8% editor rejection rate), then reviewer agreement would have been more important to the editors' decision-making processes. Had the point for *MWR* occurred closer to Editor Omnius's curves (0–3% editor rejection rate), then authors would have received the benefit of the doubt, avoiding rejection and being allowed to revise their manuscripts. Points generally

**Fig. 5** Histogram of reviewer recommendations during the first round of reviews



**Fig. 6** An enlargement of the lower left-hand side of Fig. 4: probability of reviewer rejection versus probability of editor rejection for different editor decision strategies (Editors Omnius, Quisius, and Populus, described in Table 4) and the 1:1 diagonal line. The point labeled *MWR* represents the average values for this dataset, and the other points labeled *X-n* represent the values for *n*-reviewer manuscripts by Editor X (Table 3). Points labeled E, F, and G represent values of two-reviewer manuscripts by editors who made decisions largely based on two reviewers (Table 3)



should not lie below Editor Omnius’s curves as that would indicate editors accepting manuscripts for which all reviewers recommended rejection. That the 30.4% rejection rate is so high relative to these two strategies of editor behavior indicates that the decisions by *MWR* editors as a whole tend to be influenced more by the most critical reviews than by agreement between the reviewers.

Had the point for *MWR* been much higher (about 35% and higher editor rejection rate), editors would have been rejecting manuscripts for which no reviewer had recommended rejection, which happens from time to time because the sum of all reviewers’ comments may require such extensive modifications to the manuscript that rejection may be appropriate. Such decisions, however, should not be the norm. (Such rejection rates are also possible in the scenario considered here because the 18% rejection rate is calculated from the initial reviews offered by reviewers and the probability of editor rejection is based on the final decision. A manuscript could undergo multiple rounds of revisions before being rejected, as shown in Fig. 3b where about 7% of manuscripts that undergo a second round of reviews are rejected.)

The rejection rate for any individual editor can be plotted on Fig. 6. For example, for manuscripts with two reviewers, the points for Editors A and B (A-2 and B-2) lie along Editor Quisius’s two-reviewer curve, indicating that these editors adopted a rather strict approach to decision-making (a single recommendation of rejection was usually sufficient to reject the manuscript). Indeed, each of Editor B’s rejections for two-reviewer manuscripts were guided by one reject or transfer recommendation and one major revision, and each of the eventual acceptances did not have a single reject recommendation, following Editor Quisius’s approach precisely (Table 5). Editor A was even more severe, not only following Editor Quisius’s approach closely, but rejecting two manuscripts that had no reviewer recommendations for rejection (Table 5). In addition, points E, F, and G represent the values for editors who made decisions largely based on two reviewers (“Editors E, F,

**Table 5** Fraction of editor final decisions on two-reviewer manuscripts adhering to strategies for Editors Quisius and Populus, segregated by decisions on accepted and rejected manuscripts

Editor	Two-reviewer manuscripts			
	Fractional adoption of the Editor Quisius strategy		Fractional adoption of the Editor Populus strategy	
	Accepted manuscripts	Rejected manuscripts	Accepted manuscripts	Rejected manuscripts
Editor A	<b>17/18</b>	<b>5/7</b>	17/18	1/7
Editor B	<b>24/24</b>	<b>6/6</b>	24/24	0/6
Editor C	12/13	8/9	<b>12/13</b>	<b>3/9</b>
Editor D	16/19	<b>12/14</b>	<b>16/19</b>	<b>3/14</b>

Values in *bold* represent an emphasis on these cells as discussed in the text

G” in Table 3). Interestingly, those points all clustered together near Editor Quisius’s two-reviewer curve and the average for *MWR*.

The points for Editors C and D (C–2 and D–2) are relatively close to Editor Quisius’s two-reviewer curve, although somewhat displaced downward. Editor C showed a slightly greater propensity to seek or receive reviewer agreement, with 3 of the 9 reject decisions being a result of both reviewers recommending rejection (although Editor C did reject one manuscript with two recommendations for revisions), and only 1 of the 13 accept decisions having one reviewer recommendation for rejection (Table 5). Similarly, Editor D found reviewer agreement in 3 of the 14 rejections, and all but three of the accepted manuscripts. In addition, two manuscripts were rejected after two rounds of reviews despite not having received a single reviewer rejection in the first round.

For manuscripts with three reviewers, a greater contrast among the editors can be seen (Fig. 6). In this case, the point for Editor A (A–3) is relatively close to Editor Quisius’s three-reviewer curve, indicating a relatively strict decision-making process. For example, Editor A employed Editor Quisius’s approach in 2 of the 3 rejected three-reviewer manuscripts and in 15 of the 16 accepted three-reviewer manuscripts (Table 6).

In contrast, the points for Editors B, C, and D (B–3, C–3, and D–3) are farther away from Editor Quisius’s three-reviewer curve, indicating a higher likelihood that these editors either sought or received reviewer agreement in their decisions. For example, the three editors all had multiple instances where Editor Quisius’s approach was not applied for

**Table 6** Fraction of editor final decisions on three-reviewer manuscripts adhering to strategies for Editors Quisius and Populus, segregated by decisions on accepted and rejected manuscripts

Editor	Three-reviewer manuscripts			
	Fractional adoption of the Editor Quisius strategy		Fractional adoption of the Editor Populus strategy	
	Accepted manuscripts	Rejected manuscripts	Accepted manuscripts	Rejected manuscripts
Editor A	<b>15/16</b>	<b>2/3</b>	<b>16/16</b>	2/3
Editor B	<b>25/31</b>	10/11	<b>31/31</b>	3/11
Editor C	<b>16/22</b>	7/7	<b>22/22</b>	3/7
Editor D	<b>18/21</b>	17/19	<b>21/21</b>	7/19

Values in *bold* represent an emphasis on these cells as discussed in the text

accepted manuscripts (i.e., one reviewer recommended rejection, but the manuscript was not rejected). For all four editors, all accepted three-reviewer manuscripts followed Editor Populus's strategy of majority rules (Table 6).

In summary, this graph can be used to compare the decision-making behaviors of editors and determine how similar they are to simple editor strategies (Editors Omnium, Quisius, and Populus). Although perfect correspondence with these curves is not expected because editors may deviate from these simple strategies, manuscripts may go through multiple rounds of review, or editors may incorporate multiple approaches depending on the manuscript, that the graphical method performs so well in illustrating real-life editor behavior is satisfying. That many of the points from this dataset lie close to Editor Quisius's two-reviewer curve provides strong support for the argument that, when considered in bulk, decisions made by editors at *MWR* are better described by the Editor Quisius strategy than the Editor Populus strategy.

## Summary and conclusion

To test whether the number of peer reviewers providing guidance to the editors at a scientific journal affects the rate of rejection, 500 manuscripts submitted to *Monthly Weather Review (MWR)* were examined. Because 92.2% of manuscripts used two or three reviewers, analysis was primarily limited to these manuscripts. The rejection rates were similar for two- or three-reviewer manuscripts, either at initial decision or final decision. Individual editors, however, may have had different rejection rates depending on their behavior in managing the peer-review process. Some editors rejected a larger percentage of two-reviewer manuscripts, whereas other editors rejected a larger percentage of three-reviewer manuscripts. Interviews with editors indicated that editors often employed their scientific expertise in the decision-making process, either in determining the number of reviewers or in reading the manuscript more thoroughly with a smaller number of reviewers, a result that questions whether obtaining reviewer agreement is necessarily important at all times. Also, the number of reviewers did not affect the number of rounds that a manuscript spent in review or the success of the manuscript at each round.

Finally, a simple model illustrates three editor behaviors: rejection when all reviewers recommend rejection (Editor Omnium), rejection when any reviewer recommends rejection (Editor Quisius), and rejection when a majority of reviewers recommend rejection (Editor Populus). By plotting the probability of reviewer rejection against the rate of editor rejection, an assessment can be made of how each editor and the journal as a whole compare to these simple strategies. For *MWR*, editors were more likely to follow the recommendation of the most critical reviewer to reject (Editor Quisius) rather than to look for agreement among the majority of reviewers in deciding to reject the manuscript (Editor Populus), further casting doubt on reviewer agreement as a reality in editor decision-making, at least for this one journal.

A practical side of this study is that whether two or more reviewers were used in evaluating a manuscript did not significantly affect the outcome of that manuscript in the review process. Caution should be used in overgeneralizing this result, however, because this study was only conducted for a 15.5-month period at one scientific journal. Nevertheless, at least one other study has shown that the effect of using two or three reviewers is relatively small (Bornmann and Daniel 2009). This result may give solace to editors who may feel less stress to get that third reviewer if time, resources, or reviewers are unavailable. To achieve parity in rejection rates regardless of the number of reviewers (if

that were even a goal of the peer-review process), however, editors must either adopt a majority-rules approach to decision-making (Editor Populus) or alter their decision strategy in some other way from that of Editor Omnius or Editor Quisius. Given the incredible amount of valuable time that volunteer peer reviewers willingly devote, editors should minimize impositions on reviewers' time, especially if a fair decision could be made by using fewer reviewers. More reviewers does not necessarily solve the editors' dilemma of what to do with a manuscript with conflicting reviews, as Weller (2001, p. 183) has argued, but the dilemma may not matter if the decisions are fair regardless of the number of reviewers used.

This result may also bring some solace to authors who wonder whether the editor obtaining that third review at *MWR* negatively affected the outcome of the manuscript (the answer being, statistically, no). Of course, authors must still face the burden of addressing that third reviewer's comments, but presumably these comments (which will likely be different from those of the two previous reviewers) generally improve the manuscript. From my own experience as editor, I almost always feel that the third reviewer is worthwhile by providing an additional assessment of the manuscript. Even in situations where I have used four or five reviewers (for review-type articles or broad interdisciplinary articles), I would not have reduced the number of reviewers. Thus, raw statistics about how editors do their jobs may not reveal all the underlying factors, and more interviews with individual editors is likely required to better understand how editors make decisions. Editors can use their knowledge about the science and manuscript quality before peer review starts in making their decisions about how to conduct the peer review. But, it is also true that the vagaries and individualized decisions that go into each manuscript sometimes prevent a single best answer to how to handle manuscripts in the peer review process. The number of reviewers is only one such factor.

**Acknowledgments** I thank William Gallus, Chief Editor of *Weather and Forecasting*, for the initial discussion that inspired this study, Corinne Kazarosian and Todd Davis of the American Meteorological Society for creating the dataset, and Veikko Nyfors of the Finnish Meteorological Institute for formatting the dataset. I thank the editors for their insights into their decision-making process: Nolan Atkins, Steve Cohn, James Doyle, David Jorgensen, and Elizabeth Ritchie. I thank the following individuals for their comments on earlier versions of this manuscript: Brian Colle, William Gallus, Lowell Hargens, David Jorgensen, Paul Roebber, Keith Seitter, William Skamarock, and two anonymous reviewers. Partial funding for Schultz comes from Vaisala Oyj.

## References

- Bornmann, L., & Daniel, H.-D. (2009). The luck of the referee draw: The effect of exchanging reviews. *Learned Publishing, 22*, 117–125.
- Cicchetti, D. V. (1991). Reflections from the peer review mirror. *Behavioral and Brain Sciences, 14*(1), 167–186.
- Eberley, S., & Warner, W. K. (1990). Fields or subfields of knowledge: Rejection rates and reviewer agreement in peer review. *The American Sociologist, 21*(3), 217–231.
- Ernst, E., Saradeth, T., & Resch, K. L. (1993). Drawbacks of peer review. *Nature, 363*, 296.
- Frey, B. S. (2003). Publishing as prostitution?—Choosing between one's own ideas and academic success. *Public Choice, 116*, 205–223.
- Glenn, N. D. (1976). The journal article review process: Some proposals for change. *The American Sociologist, 11*, 179–185.
- Hargens, L. L. (1988). Scholarly consensus and journal rejection rates. *American Sociological Review, 53*, 139–151.
- Hargens, L. L., & Herting, J. R. (1990). Neglected considerations in the analysis of agreement among journal referees. *Scientometrics, 19*(1–2), 91–106.

- Jorgensen, D. P., Rauber, R. M., Heideman, K. F., Fernau, M. E., Friedman, M. A., & Schein, A. L. (2007). The evolving publication process of the AMS. *Bulletin of the American Meteorological Society*, *88*, 1122–1134.
- Lindsey, D. (1988). Assessing precision in the manuscript review process: A little better than a dice roll. *Scientometrics*, *14*(1–2), 75–82.
- O'Brien, R. M. (1991). The reliability of composites of referee assessments of manuscripts. *Social Science Research*, *20*, 319–328.
- Schultz, D. M. (2009). Rejection rates for journals publishing atmospheric science. *Bulletin of the American Meteorological Society*, *90*, accepted with revisions.
- Seitter, K. L. (2002). Opening access to the AMS Journals Online. *Bulletin of the American Meteorological Society*, *83*, 1361.
- Strayhorn, J., Jr., McDermott, J. F., Jr., & Tanguay, P. (1993). An intervention to improve the reliability of manuscript reviews for the *Journal of the American Academy of Child and Adolescent Psychiatry*. *American Journal of Psychiatry*, *150*, 947–952.
- Weller, A. C. (2001). *Editorial peer review: Its strengths and weaknesses* (342 pp.). Medford, NJ: ASIST Monograph Series, Information Today, Inc.
- Zuckerman, H., & Merton, R. K. (1971). Patterns of evaluation in science: Institutionalisation, structure and functions of the referee system. *Minerva*, *9*, 66–100.