

This is the pre-peer reviewed version of the following article: [Cengher, Mirela, and Linda A. LeBlanc. "Editors' Perspectives on the Selection of Reviewers and the Quality of Reviews." Journal of Applied Behavior Analysis. November 08, 2023. <https://doi.org/10.1002/jaba.1033>, which has been published in final form at <https://doi.org/10.1002/jaba.1033>. This article may be used for non-commercial purposes in accordance with Wiley Terms and Conditions for Use of Self-Archived Versions.

Access to this work was provided by the University of Maryland, Baltimore County (UMBC) ScholarWorks@UMBC digital repository on the Maryland Shared Open Access (MD-SOAR) platform.

Please provide feedback

Please support the ScholarWorks@UMBC repository by emailing scholarworks-group@umbc.edu and telling us what having access to this work means to you and why it's important to you. Thank you.

Editors' perspectives on the selection of reviewers and the quality of reviews

Mirela Cengher¹ and Linda A. LeBlanc²

¹Department of Psychology, University of Maryland, Baltimore County, Baltimore, MD, USA

²LeBlanc Behavioral Consulting, Golden, CO, USA

Correspondence:

Mirela Cengher, Department of Psychology, University of Maryland, Baltimore County, 1000 Hilltop Circle, Baltimore, MD, 21250, USA.

Email: cengher@umbc.edu.

Editor-in-Chief: John Borrero

Handling Editor: Anna Petursdottir

Abstract

This paper describes the outcomes of a survey of 93 Editors and Associate Editors of behavior-analytic journals. We sought information about variables that influence their judgment of the selection of reviewers, selection of review panels, and quality of reviews. When selecting reviewers, participants rated highly expertise on the topic, history of conducting good reviews, and history of writing constructive and respectful reviews. When selecting review panels, participants rated highly stratifying reviewers based on their expertise, avoiding conflicts of interest, and the matching based on the area of expertise between reviewers and authors. When evaluating the quality of a review, participants rated highly considerations related to research design, the science underlying the main idea, and accurate interpretations of the data. Participants did not rate copyediting as impactful. Overall, the extent to which reviewer selection was influenced by membership in underrepresented groups varied. These findings can inform the development of training programs for teaching peer review repertoires.

KEYWORDS

behavior-analytic journals, editorial process, peer review

The system of peer review is fundamental to the development and dissemination of scientific discovery (Houry et al., 2012). Peer-reviewed publications are the building blocks of research careers, hence the “publish or perish” ethos (Rawat & Meena, 2014). The editorial process facilitates quality control of science, legitimizes scientific research, and self-regulates scientific communities (Tennant, 2018). The importance of the peer review system is also illustrated by its scale, with an investment of \$2 trillion globally and the production of more than 3 million peer-reviewed articles annually (Tennant & Ross-Hellauer, 2020). The investment includes expenses related to supporting existing businesses (e.g., publishing companies), direct research expenses, and new business projects (National Science Board, National Science Foundation, 2020).

Generally, academic journals adhere to the peer-review system when evaluating manuscripts submitted for publication. Given the global scale of the peer-review system, one would expect it to be well-developed and scientifically grounded. Yet, research on peer review shows that it is a relatively underdeveloped system that varies in process and outcomes (Tennant & Ross-Hellauer, 2020). As an example of variation in process, Song et al. (2021) conducted a scoping evaluation on peer-review guidelines in biomedical journals. The authors identified 78 records (i.e., 65 publications and 14 websites), which varied with respect to guidelines for writing reviews as well as resources used to develop these guidelines. Across articles, Song et al. identified the following recommendations for conducting peer review: attitude of the review, suggestions for becoming a reviewer, the content of the review, access to resources necessary for review work, managing ethical challenges as a reviewer (e.g., avoiding conflicts of interest), managing ethical challenges in the manuscript (e.g., plagiarism), the structure of the review, time management, and other (e.g., signing reviews, recognition, and collaborating with others). Although the articles

varied widely with respect to which of the aforementioned recommendations for conducting peer review were highlighted, the most commonly referenced criteria were content (95.3%) and structure (75%). Of note, only three of the studies included in the review by Song et al. reported systematic means of developing criteria, such as the review of theoretical and data-based literature (Christenbery, 2011), the inclusion of reference standards from peer-reviewed articles (Duchesne & Jannin, 2008), and the development of survey among reviewers plus workshop (Schüttpelz-Brauns et al., 2010). Based on these results, Song et al. suggested that there is a need to reduce variability and develop more uniform guidelines for peer review based on objective criteria.

Different approaches to constructing guidelines can result in variable recommendations for peer review. Further, the fact that existing guidelines have been developed based on individuals' experiences and preferences, rather than quantitative criteria, could contribute to variable recommendations for peer review as well. One way to reduce these sources of variability is to design studies using set procedures and quantitative data that can inform the development of recommendations for peer review. To accomplish this, Falkenberg and Soranno (2018) analyzed complete reviews and the corresponding manuscripts, for which Editors provided a quantitative ranking of the quality of the reviews. The authors found that the disagreement between the reviewer's recommendation and the Editor's decision for publication did not negatively influence the evaluation of the reviewer's performance. Further, Editors valued reviews that were comprehensive and constructive, and that included comments related to the study's goals, methods, results, and contribution to the literature. These outcomes were later replicated by Pranić et al. (2020).

In summary, current research shows that guidelines for peer review vary in process and that Editors in other disciplines (e.g., biomedicine) value comprehensive and constructive reviews.

Given the influence of the editorial process on scholarship and scholarly careers, it is important to assess the preferences and expertise of Editors and Associate Editors (AEs) who select reviewers, judge the quality of reviews, and appoint editorial review boards. Invitations to serve on an editorial board signal recognition for one's scientific merit (Bedeian et al., 2009)¹ and can particularly influence the career of relatively new scholars. Information on the variables that Editors and AEs consider when nominating scholars for membership on editorial boards can guide the actions that scholars need to access such reinforcers.

To our knowledge, no prior research has systematically evaluated what Editors and AEs value in a review and their selection of reviewers, selection of review boards, and nominations for editorial board membership. Further, given how much scientific standards differ, it is unclear whether what Editors and AEs value is constant across disciplines. The purpose of this manuscript was to conduct a survey of Editors and AEs that could inform the development of empirically-based guidelines for writing reviews for behavior-analytic journals. We attempted to determine the current status in our field and to inform future training of behavior-analytic scholars.

METHOD

Participants

We sought responses from participants who had served as an AE, guest AE, or Editor at least once for a behavior-analytic journal. We selected this inclusion criterion because only an Editor or AE has the experience of selecting reviewers and review panels and evaluating the work of reviewers, which were the topics of the survey. To access the greatest number of potential

¹ However, scientific merit is not always the leading variable that results in editorial board membership; see Bedeian et al. (2009).

respondents, we invited a broader group of individuals (i.e., all current Editors, AEs, and review board members) than we expected to qualify (i.e., only Editors, AEs, guest AEs). We created a database with all current² Editors, AEs, and editorial board members for journals that were part of the Society for the Experimental Analysis of Behavior or the Association for Behavior Analysis International. These included the *Journal of Applied Behavior Analysis*, *Journal of the Experimental Analysis of Behavior*, *Behavior Analysis in Practice*, *The Analysis of Verbal Behavior*, *Perspectives on Behavior Science*, *Behavior and Social Issues*, *Education and Treatment of Children*, and *The Psychological Record*. Although other behavioral journals exist, focusing on the journals published by these two organizations allowed us to be systematic in our recruitment efforts while accessing a large number of scholars.

Our database included the name of each individual who held one of the aforementioned roles, as well as their email address. When an email address was not available online, we emailed the journal Editor and asked for this information. We retrieved the email addresses of all but one current Editor, AE, or editorial board member across all of the targeted journals. The database was completed on January 1, 2022, and included 335 potential participants, representing considerably more individuals than we expected to qualify for the study. We included editorial board members because some serve or previously served in the role of guest AEs, AE, or Editor even though they might not currently serve in that role. However, not all editorial board members served as guest AEs at the time of the survey; as such, the final number of potential participants who met our inclusion criterion was lower than the 335 individuals included in our email list. The email sent to potential participants about the study specified the roles that we intended to include and the assumption was that many receiving the email would not even begin the survey if they did not

² At the time of the study (i.e., January 2022).

have the appropriate experience. Further, if an individual who did not qualify started completing the survey, they would have been excluded based on their responses to questions pertaining the editorial roles they served (described below).

Survey development

We developed the survey around four themes: (a) identification of reviewers and selection of review panels, (b) selection of reviewers on the first and second rounds of revisions, (c) evaluation of the quality of a review, and (d) editorial board membership nominations. We shared the initial draft survey with a panel of experts who served as Editors and AEs throughout their careers and, as such, had extensive expertise in conducting editorial work. These experts provided feedback on the content of the survey and are mentioned in the acknowledgment section. Based on the feedback received from the panel of experts, we developed a 13-question survey consisting of multiple-choice questions, comparative rating scales, and open-ended response questions (see Supporting Information A for the complete survey).

The first two questions (1 and 2) focused on demographic information, and the next three questions (3, 4, and 5) focused on the extent and type of experience with editorial work (i.e., professional experience). Question 5 was, "Have you ever served as a (guest) AE for a behavior-analytic journal?" All participants who responded "no" were directed to the end of the survey. In other words, the participants with no prior experience serving in an AE role were excluded from the study if they had not already self-selected out of participation. Thereafter, all participants who responded "yes" answered questions regarding (a) the identification of reviewers and selection of review panels (Section 3; Questions 6 and 7), (b) the selection of reviewers on the first and second rounds of revisions (Sections 5 and 6; Questions 8, 9, and 10), (c) evaluation of the quality of a

review (Section 7; Question 11), and editorial board membership nominations (Section 8; Question 12).

The first five questions had a multiple-choice response option. Further, Questions 6 through 12 included eight to 10 statements in which participants were asked to provide their ratings on a 5-point Likert scale, ranging from 1 (*not at all useful*) to 5 (*very useful*). The anchors varied across questions depending on content. For example, for the identification of new reviewers, the anchors were not at all useful, not useful, neutral, useful, and extremely useful. For selection of new reviewers, the anchors were not at all influential, not influential, neutral, influential, and highly influential (see Supporting Information A). The last question (Section 9; Question 13) allowed participants to type additional comments about the selection and assignment of reviewers. Responding to Questions 1 through 12 was required, whereas responding to Question 13 was optional. Questions pertaining to the following topics included an open-ended response option (i.e., “Other [please indicate:]”): identification of new reviewers (Question 6), selection a review panel (Question 7), and evaluation of the quality of a review (Question 11).

Survey distribution

We distributed the survey to all 335 potential participants via the Gmass function on February 1, 2022, and one additional time on February 17, 2022. The Gmass function is a mass email tool used to send information to large groups of people. The survey was made available to the participants via an anonymous link to Qualtrics, a survey software program. Participants provided their consent before participating in the survey. Qualtrics estimated that participants would take 10 to 15 min to complete the survey. We designed the survey so participants could only

complete it once from any given browser. The survey link remained active from February 1, 2022, to March 1, 2022.

Qualitative and quantitative analyses

Quantitative and qualitative research methods refer to the type of data that are generated and analyzed (Garbarino & Holland, 2009). In short, quantitative data are quantifiable (e.g., Likert-scale responses to a survey). Qualitative data can come in the form of text (e.g., open-ended responses to a survey) and are typically evaluated based on thematic content. Quantitative data allow one to evaluate the relation between variables (e.g., history of conducting good review and invitation to be part of an editorial board), whereas qualitative data allow one to place these relations in context (e.g., qualitative description of what makes a review “good”) and, as such, allow a more nuanced analysis. For quantitative analyses, we used descriptive statistics of central tendency (i.e., mean), range, standard deviations, and percentages rated based on anchors such as “not at all useful” or “very useful” (see Supporting Information A for anchors used with each question). These descriptive statistics were computed automatically in Qualtrics. Of the measures of central tendency, we reported mean because (a) our distribution of data was normal, and, as a result, (b) the mean, mode, and median were almost identical.

For qualitative analyses, we used thematic analyses to synthesize and interpret the responses to our open-ended questions (see below for a brief description of the process; Lochmiller, 2021). We first generated codes (e.g., responds quickly, history of accepting invitations, submits the review on time), which thematically synthesized different responses participants made. Then, we grouped these codes into larger themes (e.g., timeliness) that allow the evaluation of generality across responses. These themes uniquely described a set of responses, and

the organized summary of these themes highlights similarities and differences in responses. The Supporting Information contains the themes as well as all of the unedited responses. Specifically, Supporting Information B, D, F, and H includes open-ended responses; themes are summarized in Supporting Information C, E, G, and I.

RESULTS

One hundred thirty-seven responders (out of 355; 39%) clicked on the link to the survey. Of these, 136 (38%) agreed to participate after reviewing the consent form. One hundred and three responders (29%) indicated that they had served as an AE, guest AE, or Editor for a behavior-analytic journal (i.e., they qualified for participation), and 93 (26%) of them answered all questions and completed the survey. Below, we report data for the 93 participants who completed the survey. The authors of this manuscript did not complete the survey. Note that the email list was constructed to include many more recipients than were expected to qualify for participation in hopes of garnering more responses at the expense of the response percentage.

For Likert-response questions that allowed quantitative analysis, we report percentage agreement and the n of participants who selected the *useful* or *very useful* options (or other anchors that denote agreement or strong agreement) as well as means and standard deviations. For open-ended responses, we provide the results of thematic analyses.

Demographic information and professional experience

The participants identified with the following races and ethnicities: Asian (4%, $n = 4$), Black or African American (2.1%, $n = 2$), Brazilian Mix (1%, $n = 1$), Hispanic or Latinx (12.9%, $n = 12$), and White (81.7%, $n = 76$; see Table 1). Three participants (3.2%) chose not to disclose

their race. None of the responders identified as American Indian, Alaskan Native, Native Hawaiian, or Other Pacific Islander. Some participants selected more than one response option. The participants identified with the following genders: cisgender woman (44.1%, $n = 41$), cisgender man (48.4%, $n = 45$), and gender non-binary (1%, $n = 1$). Six participants (6.5%) chose not to disclose their gender. None of the responders identified as transgender women or men. This demographic breakdown is consistent with data reported by the National Science Foundation on Ph.D. recipients in the United States of America (National Science Board, National Science Foundation, 2020). Our sample included relatively equal numbers of individuals who identified as cisgender women and men, which contrasts with prior research indicating that women were underrepresented in science (e.g., Tennant, 2018), or that women in behavior analysis are more represented in roles that are typical for early career stages (Nosik et al., 2019)

As per the inclusion criterion, all participants reported conducting editorial work. The duration of their experience conducting editorial work skewed toward longer durations, with the majority of respondents indicating 6–10 years (24.7%; $n = 23$) and over 11 years (61.3%; $n = 57$; see Table 1) of editorial experience. Participants could serve in more than one capacity (e.g., Editor for one journal, AE for another), therefore the total number of roles reported exceeds the number of participants. Specifically, participants reported a total of 262 reported roles. This number serves as the denominator for the percentages reported below. Participants conducted editorial work in one or more of the following capacities: editorial board member (34.6%; $n = 91$), guest AE (31.9%; $n = 84$), AE (22.8%; $n = 60$), Editor (8.4%; $n = 22$), and other (2.3%; $n = 6$). These data suggest that the sample is skewed to more senior and experienced scholars, which is

consistent with the roles of AE or Editor. The other roles indicated by participants were guest reviewer, guest Editor, and AE mentee³.

There is a relatively equal distribution of editorial roles between genders, but not necessarily by ethnicity. For example, 23.3% of men ($n = 10$ of a total of 43) and 19.5% of women ($n = 8$ of a total of 41) reported serving in the role of Editor, and 58.13% of men ($n = 25$) and 68.3% of women ($n = 28$) reported serving in the role of AE. However, only 17.6% of the participants who identified as Black, Indigenous, or People of Color (BIPOC; $n = 3$ of a total of 17) reported having served as an Editor, compared to 25% of the White respondents ($n = 19$ of a total of 76). For all other editorial roles, the breakdown across racial groups is relatively equal. For example, 24% of BIPOC responders ($n = 11$) and 23% of the White respondents ($n = 49$) reported serving as an AE. Even though these data suggest that there is not a large discrepancy in the percentages of participants serving in each role, the absolute numbers are smaller for BIPOC behavior analysts.

Identification of reviewers and selection of review panels

Quantitative findings

Respondents endorsed the strategies they used to identify new reviewers, defined as individuals who had not performed reviews for the AE before. Table 2 displays the mean, standard

³ The *Journal of Applied Behavior Analysis* has a mentorship program whereby junior scholars (mentees) are paired with a senior scholar (mentor) to jointly serve in the AE role for manuscripts. The mentor provides guidance to the mentee in all aspects of the editorial process related to the role of the AE. The mentor and mentee typically work together on at least five manuscript submissions and the mentee may continue to independently serve in the role of Guest AE after the mentorship period has ended.

deviation, percentage agreement, and the number of participants whose ratings were high (i.e., *agreed* or *strongly agreed*) for each item in Section 3. This section included ten 5-point Likert scale questions. Participants rated three statements with a score of 4 or higher: (a) identifying scholars who conduct thematically related research ($M = 4.5$; $SD = 0.5$), (b) searching the reference list of the manuscript to be reviewed ($M = 4.1$; $SD = 0.7$), and (c) selecting editorial board members ($M = 4.0$; $SD = 0.8$). They rated six questions with a score between 3 and 4 (see Table 2). Overall, respondents provided the lowest ratings for the following statements: (a) asking their students ($M = 2.5$; $SD = 1.0$), (b) identifying individuals who are members of underrepresented groups ($M = 3.2$; $SD = 0.9$), and (c) asking for recommendations from one's network ($M = 3.5$; $SD = 0.9$). In summary, participants find reviewers who conduct thematically related work by looking through publicly available channels (e.g., editorial board lists, reference sections). Participants are less inclined to choose reviewers based on recommendations from their professional network, or based on factors related to diversity, equity, and inclusion.

Participants also answered questions regarding the selection of review panels. Participants completed a five-point Likert scale for nine statements on this topic (see Table 2). Participants rated four of the statements with a score of 4 or higher (noting that the strategy was *useful* or *extremely useful*), three questions with a score between 3 and 4 (noting that the strategy was *neutral* or *slightly useful*), and two statements with a score below 3 (noting relative *disagreement about the usefulness* of the strategy). Participants provided the highest scores for the following statements: (a) stratify the review panel based on the expertise of the reviewers ($M = 4.8$; $SD = 0.5$), (b) avoid known conflicts of interest ($M = 4.6$; $SD = 0.6$), and (c) match based upon at least one reviewer and the author(s)'s area of expertise ($M = 4.5$; $SD = 0.6$). Participants provided the lowest scores to the following statements: (a) match between the rank of at least one reviewer and

that of the author(s) of the manuscript ($M = 2.8$; $SD = 1.1$), (b) include only senior reviewers ($M = 2.7$; $SD = 0.9$), and (c) match based on other variables, such as the author and the reviewer being from underrepresented groups ($M = 3.1$; $SD = 1.0$). Overall, participants valued considerations related to balancing reviewers' and authors' areas of expertise and avoiding known conflicts of interest. The participants did not appear to find considerations related to the career level of the reviewers and authors, or factors related to diversity, equity, and inclusion useful.

Qualitative findings

Thirteen participants further explained the strategies they used to choose new reviewers through open-ended responses (see the Supporting Information B for themes and Supporting Information C for the open-ended responses). The common themes identified in the analysis of these findings were: (a) skills valued ($n = 10$; e.g., area of expertise and expertise ["Expertise in the area measured by the number and quality of publications," #11], history of conducting good-quality reviews), and (b) other strategies to identify new reviewers ($n = 2$; e.g., individuals from the portal nomination form, recruiting others' students and postdoctoral fellows). Although skills valued pertain to the selection rather than the identification of reviewers (see below for variables involved in the selection of reviewers), the other strategies used to identify new reviewers complement the quantitative results by providing alternative ways to identify new reviewers.

Forty-six participants also provided open-ended descriptions of strategies they use when selecting a review panel (see Supporting Information D for open-ended responses and Supporting Information E for themes). The most common themes were: (a) area of expertise ($n = 18$, e.g., "reviewers with knowledge in the topics covered in the paper," #22), (b) history of accepting reviews and availability ($n = 17$, e.g., "experience with that reviewers [sic], the courtesy and

professionalism of previous reviews,” #36), (c) timeliness ($n = 14$, e.g., “record of completing assigned reviews within reasonable time frame,” #38), (d) history of conducting good reviews ($n = 12$; described as prescriptive, including constructive feedback, using a respectful and professional tone), and (e) diversity in expertise and perspectives ($n = 9$; e.g., “I like to ensure [assumed] gender representation,” #2). One reviewer commented, “because this is a volunteer activity, sometimes it is just about who agrees to review. Ridiculous that publishing companies bring in so much money off the backs of unpaid reviewers,” (#46). These open-ended responses highlight that participants value reviewers who conduct good and timely reviews, and whose area of expertise overlaps with their current editorial needs. Area of expertise and diversity in perspectives overlap with variables identified as valued by participants through the quantitative analyses. History of accepting reviews, history of conducting good reviews, and timeliness closely relate to skills valued when identifying new reviewers, as described above.

Selecting reviewers for the first and second rounds of revisions

Quantitative findings

Once Editors and AEs identify potential reviewers, they decide whom to invite to review. In Section 6, the participants rated several statements regarding the decision to ask someone to review a new manuscript (see Table 2). This section consisted of nine 5-point Likert scale statements. The participants rated three statements with a score above 4, five statements with a score between 3 and 4, and one statement with a score below 3. The participants provided the highest scores for the following statements: (a) the individual has expertise in a domain where the AE needed reviews ($M = 4.8$; $SD = 0.4$), (b) the individual writes reviews that are respectful and

constructive ($M = 4.7$; $SD = 0.6$), and (c) the individual writes high-quality reviews that help in the editorial process ($M = 4.6$, $SD = 0.5$). The participants provided the lowest scores for the following statements: (a) the reviewer carefully copyedited the manuscript ($M = 2.4$; $SD = 1.1$), (b) the individual is a member of an underrepresented group ($M = 3.1$; $SD = 0.9$), and (c) the individual publishes articles in the journal for which the AE serves ($M = 3.4$; $SD = 1.1$). We did not include an open-ended question that could result in qualitative findings for this section, considering that participants likely use the same considerations when identifying and selecting new reviewers.

In Section 4, participants indicated their preference for selecting a new panel of reviewers if the manuscript underwent a second round of review (i.e., the initial decision was to invite resubmission). Seventy participants (75%) reported that they keep some reviewers and add at least one new reviewer, whereas 23 participants (25%) reported that they keep the same reviewers from the initial panel. Participants rated several statements regarding reasons why they choose not to invite a reviewer for the second round of revisions. Participants provided the highest rating for the following statements: (a) the review was not helpful in the editorial process ($M = 4.4$; $SD = 0.8$), (b) the reviewer has two or more manuscripts they are currently reviewing ($M = 3.7$; $SD = 1.1$), and (c) the review was harsh ($M = 3.5$; $SD = 1.2$). Participants provided a relatively neutral rating for the poor timeliness of the review ($M = 3.2$; $SD = 1.0$). Participants provided the lowest rating to the statement that the reviewer recommended rejecting the manuscript, which was incongruent with the decision made by the AE ($M = 2.5$; $SD = 1.1$). Therefore, a reviewer may not be invited for the second round of submission if writing harsh, poor-quality reviews or when balancing a high load of manuscripts. However, the prior recommendation and its correspondence with the initial decision generally do not influence the decision to issue a repeat invitation. The results for this question are consistent with the results for the previous question, suggesting that participants use

the same considerations to decide whether to invite reviewers for the first or subsequent round of revisions. We did not include an open-ended response option for this question.

Evaluation of the quality of the review

Quantitative findings

Further, the participants rated variables that influence their evaluation of the quality of a review. This section, Section 8, consisted of ten 5-point Likert scale items, followed by an optional open-ended question where participants could describe other variables they take into account when evaluating the quality of a review. Table 3 displays the mean, standard deviation, and percentage agreement (i.e., percentage of responders who indicated agreement or strong agreement) for each item in this third section. Participants rated eight of the 10 statements with a score of 4 or higher, one question with a score close to 4, and one statement only with a score below 3. Participants provided the highest scores for the following statements: (a) the reviewer evaluated whether the experimental design was appropriate ($M = 4.6$; $SD = 0.5$), (b) the reviewer evaluated whether the science underlying the main idea is strong ($M = 4.6$; $SD = 0.6$), and (c) the reviewer noted whether the interpretation of the data is accurate ($M = 4.4$; $SD = 0.6$). Participants provided the lowest scores to the following statements: (a) the reviewer conducts line edits ($M = 2.6$; $SD = 1$), and (b) the reviewer evaluates whether the cited literature is current and related to the research topic ($M = 3.9$; $SD = 0.7$). In summary, participants favor reviews that focus on evaluating the value of the research questions, the method used to answer the research question, and the interpretation of the results. These three variables are essential for a comprehensive evaluation of a manuscript. Participants place less value on reviews that focus only on details and provide

copyedits or evaluate the extent to which the literature review is up to date. The results for this section are consistent with the results in the previous section, suggesting that participants use the same considerations to evaluate the quality of reviews and to decide whether to invite reviewers for another round of revisions.

Qualitative findings

Twenty-three participants further described the variables that influence their evaluation of the quality of a review through open-ended responses (see Supporting Information F for open-ended responses and Supporting Information G for the themes). The common themes identified in the analysis of these findings were: (a) prescriptive and constructive feedback ($n = 11$; e.g., “the review was [...] prescriptive [included helpful details and recommendations],” #2), (b) respectful and professional ($n = 10$; e.g., “Tone of the review. I do not appreciate overly harsh reviews,” #6), (c) focus on the big picture rather than detail ($n = 4$, e.g., “I don’t expect reviewers to do line edits. I consider that the AEs [sic] job,” #7), and (d) unbiased ($n = 3$; e.g., “An absence of in-group gatekeeping,” #5). The focus on the big picture rather than, or in addition to, details is a recurrent theme across the Likert scale and open-ended responses in this section. One reviewer commented, “Most of my actual answers to these questions would be more nuanced than a Likert-style survey permits. [...] There is no formula for producing a good review,” (#23). In summary, despite the limitations imposed by the Likert-scale questions, participants highly value reviews that are prescriptive, constructive, unbiased, and respectful.

Editorial Board membership nominations

Quantitative findings

Section 8 included statements that illustrate reasons why an AE would decide to nominate or vote for someone to become an editorial board member (see Table 3). This section included the same nine 5-point Likert scale statements as in the previous section. The participants rated three of these statements with a score above 4, five of the statements with a score between 3 and 4, and one of the statements with a score below 3. The participants provided the highest scores for the following statements: (a) the individual writes reviews that are respectful and constructive ($M = 4.6$; $SD = 0.6$), (b) the individual has expertise in a domain where the Editor needs reviews ($M = 4.5$; $SD = 0.7$), and (c) the individual wrote a high-quality review that helped in the editorial process ($M = 4.5$; $SD = 0.7$). The participants provided the lowest rating to the statement that the reviewer carefully copyedits the manuscript ($M = 2.5$; $SD = 1.1$). The results for this question are aligned with the results from the previous two sections, denoting that participants use the same considerations to nominate reviewers for editorial boards. We did not include an open-ended question that could result in qualitative findings for this section.

Final comments

The last open-ended question invited participants to share any other comments on the topic of the survey (open-ended responses to which can be found in Supporting Information H and themes can be found in Supporting Information I). Twenty participants shared additional thoughts, and we grouped these responses into two recurrent themes. One recurrent theme pertained to variables that influence the decision to invite reviewers ($n = 13$; e.g., “I try to select reviewers who were not affiliated with each other, went to different graduate programs, work in different places, etc.,” #2). Most of these comments overlapped with other themes that were stated above, such as

balancing reviewers based on expertise, tone of the review, or avoidance of conflicts of interest.

However, a new topic emerged—diversity, equity, and inclusion. Two participants (#1, #16) questioned the ethics of overburdening junior reviewers or members from underrepresented groups. For example:

I have mixed feelings about seeking out junior reviewers. Yes, they need the experience and the exposure, but I feel like we also tend to take advantage of their willingness to review/unwillingness to say no compared to more senior individuals in the field. Spending too much time on editorial work early on can negatively impact career trajectories—it did in my case.” (#16)

Several other participants commented on factors related to diversity, equity, and inclusion ($n = 5$). Participants reported that they appreciate the focus on this topic (#1), that they nominate members of underrepresented groups for editorial boards, if conducting good reviews (#1), that they try to have an even split in assumed gender identity across reviewers (#4), and that they try not to over select individuals from underrepresented groups (#1). In contrast, some participants reported that there are too few members from underrepresented groups, and as such, this does not influence their decisions (#11), and the survey made them aware that they need to think more about this topic (#14).

DISCUSSION

Despite the ubiquity of the scientific peer review process in the last 100 years, there remains considerable variability in the process and our understanding of what constitutes quality in a review is nascent (Song et al., 2021; Tennant & Ross-Hellauer, 2020). We developed a survey to investigate the extent to which Editors and AEs of behavior-analytic journals value different

variables when selecting reviewers, when selecting review panels, and when making recommendations for editorial board membership. Our survey generated responses from 93 participants who served in the role of (guest) AE at least once for a behavior-analytic journal. Most often, participants identify scholars who conduct thematically related research, search the reference list of the manuscript to be reviewed, or search the editorial board or the reference list of manuscripts handled to identify new reviewers. When selecting reviewers for a new manuscript, participants identify scholars who have expertise on the respective topic, have a history of writing high-quality reviews that help in the editorial process, and write reviews that are respectful and constructive. When selecting review panels, participants rated highly considerations of stratifying the reviewers based on their expertise, avoiding conflicts of interest, and matching based on the area of expertise between reviewers and between the reviewers and authors. For a second round of review, participants indicated that they are unlikely to invite a scholar when they have a history of writing harsh, unhelpful reviews, or have multiple manuscripts they are currently reviewing. When evaluating the quality of a review, participants rated highly considerations related to the evaluation of experimental control, science underlying the main idea, and interpretation of data (including using the right interpretative system). Participants used the same considerations when deciding to re-invite reviewers for the second round of revisions and to nominate reviewers for editorial board membership.

Few studies (Falkenberg & Soranno, 2018; Pranić et al., 2020) evaluated Editors' and AEs' perspectives on the quality of peer review. We extended this body of literature by evaluating what variables Editors and AEs consider not only when evaluating the quality of a review, but also when choosing reviewers, selecting review panels, and choosing editorial board members. Further, we extended the research on peer review from other fields to behavior analysis. Future research

should focus on developing guidelines for writing reviews for scientific journals, which could then be used to develop training programs (for a first attempt in behavior analysis, see Cengher & Kodak, 2022). Behavior analysis has a long history of developing successful interventions for a wide range of skill deficits (e.g., behavioral skills training; Slane et al., 2021); as such, we are perhaps uniquely poised to develop effective training programs to teach junior scholars to conduct reviews for scientific journals.

The qualitative data generated by our survey indicate that Editors and AEs nominate reviewers to editorial boards based on expertise, history of conducting good reviews, and professionalism. Two of the participants also provided open-ended responses related to this topic; in the section Anything Else That Influences Your Evaluation of the Quality of the Review, a participant noted “Note that guidelines for reviewers are available; many scientific journals clearly and transparently state their expectations, in the review process” (Supporting Information F, #17). On another note, in the section Do You Have Any Other Comments About How You Select Reviewers or Assign Papers to Them, a participant commented:

I would also note that many journals should follow a standard for which is the MINIMUM requirement necessary for a review board member to be added. Consider JABA's [*Journal of Applied Behavior Analysis*]: an individual must be first or serve as the corresponding author on at least two manuscripts in JABA that year (sic)⁴ in order to be considered. It's a minimum bar that many other journals should consider adopting. I would also state that those expectations/requirements should be transparent (i.e., everyone should know what

⁴ JABA's guidelines indicate that one must serve as a first or corresponding author for two manuscripts in the past 5 years, rather than 1 year, to be nominated for editorial board membership. This information is provided along with a ballot that editorial board members receive every year to vote for new editorial board members.

those requirements are). Finally, I understand that there is some liberty for EICs [Editors-in-Chief] and AEs to select and advocate for others to be on boards. Those *same* minimum requirements should apply to those individuals for whom EICs or AEs advocate to join a board (Supporting Information H, #20).

These comments suggest that at least some Editors and AEs support the development of transparent and objective standards for nominations to editorial boards across journals.

Participants generally provided lower scores for the influence of factors related to diversity, equity, and inclusion (e.g., considering factors related to these topics when choosing reviewers) than other variables that impact selection. However, in the open-ended question sections, many participants commented on this topic. For example, in the section Strategies Used when Selecting a Review Panel, a participant commented, “I (...) try to include BIPOC without overburdening anyone. I try to bring diverse perspectives” (Supporting Information D, #9), whereas another commented:

I try to find reviewers who can provide multiple, and sometimes divergent perspectives on the topic, to provide the author(s) with the opportunity to consider comments from both those who are 'supportive' of the work and those who may present otherwise not considered viewpoints (Supporting Information D, #15).

In the section Do You Have Any Other Comments about How You Select Reviewers or Assign Papers to Them, a participant commented:

I appreciate the questions about selecting reviewers who are from underrepresented groups—I do think about this and try to balance not overselecting people from underrepresented groups, risking asking them to do more than their "share" of editorial work. If someone from an underrepresented group does a good job reviewing a few times, I

will advocate for them to be invited to the editorial review board (as that will be a stronger CV-builder than repeatedly doing guest reviews) (Supporting Information H, #1)

Another participant commented:

In many questions, (sic) was asked about underrepresented groups. I confess I only think about it when the manuscript demands it; for example, I would prefer (and have preferred) female scientists to review manuscript (sic) about sexism or feminism; but my usual go to think [sic] about reviewers is the manuscript topic. Maybe this has to change (Supporting Information H, #10).

Yet another commented, “There are not enough underrepresented researchers in my area to make that a reasonable influence on my decisions” (Supporting Information H, #11). Collectively, these comments suggest that Editors and AEs value diversity of viewpoints and expertise in ways that may not have been captured by the survey questions.

One limitation of our study was that we did not define underrepresented groups and, as such, participants likely responded using the constructs developed based on their learning history. Possibly, open-ended responses were variable due to this omission. Future research should define underrepresented groups when asking questions about this topic, and consider a more comprehensive evaluation of these different ways of evaluating participants' perspectives on factors related to diversity, equity, and inclusion. In addition, future research could focus on the underrepresentation of BIPOC researchers in the scientific leadership roles of our behavioral peer-review process. Although equal representation may be desirable, members of underrepresented groups are asked to do more service (i.e., in this case, write reviews; see Carson et al., 2019), which does not necessarily translate to promotion (i.e., to editorial board membership, or AE role; see comments from participants on this topic above). Thus, Editors and AEs might consider

inviting reviewers from diverse backgrounds, particularly if they provide mentorship and additional supports to facilitate success and do not place undue burden on the same reviewers. In addition, reviewers who are international authors have diverse research and clinical experiences and other personal characteristics (e.g., cultural variables) that may bring a unique perspective when evaluating a manuscript. In general, Editors and AEs could consider balancing editorial boards in a way that ensures that individuals from various underrepresented groups contribute to the evaluation of scientific papers to increase diversity in thought, provided that there is a sufficient pool of qualified individuals who can perform the work.

Another limitation of our study lies in the recruitment process. We created a database with all current Editors, AEs, and editorial board members of several journals in behavior analysis and contacted them all to invite participation, hoping to capture former Editors and AEs who still serve on editorial boards. However, it is likely that at least some former Editors and AEs were not included in the invitation list. This invitation process also likely lowered our response percentage (i.e., the percentage of all who were emailed who completed the survey) because we knowingly emailed people who would likely not qualify. We also limited the invitations to those behavior-analytic journals published by two organizations, though the invitees certainly may have served as AEs for other behavior-analytic journals (e.g., *Behavioral Interventions*, *Behavior Modification*) in the past. Possibly, the AEs and Editors of other behavior-analytic journals would have responded differently than those we sampled.

Most questions in our survey had a Likert scale, which allowed us to quantitatively evaluate how our sample of participants valued different variables in the peer review process. The open-ended questions generated qualitative data that are reported in full in the Supporting Information. The qualitative data supplemented the quantitative data in two ways. First, these

different ways of analyzing responses allowed us to evaluate areas of agreement and disagreement. As an example of agreement, both qualitative and quantitative data indicate that Editors and AEs value certain characteristics (e.g., timeliness, tone, and history of conducting good quality reviews). As an example of disagreement, even though many participants indicated that they value timeliness in one open-ended section, close-ended responses did not indicate that it negatively impacted the AE's likelihood of issuing a repeat invitation. Therefore, timeliness seems to be valued, but its absence was not always perceived negatively or as a reason to avoid future invitations. A potential explanation for this counterintuitive finding is that our survey investigated the absolute value of these variables, which may not be aligned with their relative value. For example, all things being equal, timeliness is important. However, when the AE has to choose between timeliness and other variables, such as the quality of the review, they may choose the latter. That is, they may select a reviewer who takes considerable time but writes great reviews, rather than one whose reviews are not as useful despite meeting the deadline. In a perfect world, perhaps the AE would prefer a review that is both good and submitted on time.

The second way that qualitative analyses supplemented quantitative analyses was by providing context or supplementing the topics covered in the close-ended questions. For example, several participants commented on the fact that conducting reviews is unpaid labor (Strategies Used When Selecting a Review Panel, Supporting Information D, #46), that reviewers (especially from underrepresented groups) may be overburdened (Strategies Used When Selecting a Review Panel, Supporting Information D, #9, #46; Do You Have Any Other Comments About How You Select Reviewers or Assign Papers to Them, Supporting Information H, #9, #16), and that there is a need for a transparent process of nominating reviewers to editorial boards (Do You Have Any Other Comments About How You Select Reviewers or Assign Papers to Them, Supporting

Information H, #20). These comments suggest that Editors and AEs consider variables other than the ones we included in our survey in their editorial board. Further, these comments will hopefully lead to action that can improve the equity and transparency of the editorial practices in behavior analysis.

Participants commented on the fact that quantitative research may not capture the complexity of the peer review process. For example, in the section Strategies Used When Selecting a Review Panel, a participant wrote:

Each manuscript invites special considerations—type of research (sometimes it's helpful to have an expert in methods), review or theoretical paper, etc. Matching algorithms would have to take into account manuscript type, complexity of methods (expertise needed), and other considerations (Supporting Information D, #32)

In the section Do You Have Any Other Comments about How You Select Reviewers or Assign Papers to Them, a participant commented: “It's a nuanced process. Your initial results are likely to be interesting, but I wonder if a more qualitative approach would yield fruitful additional details” (Supporting Information H, #20). Despite this limitation, this study represents the first investigation of variables that influence the decisions Editors and AEs make and, therefore, provides preliminary information that can inform future research. Future research should employ qualitative measures, such as focus groups (for a recent example of this approach, see St. Peter et al., 2023). Further, future research should inform the development of tutorials on writing reviews for scientific journals, which could be empirically validated (see Cengher & LeBlanc, 2023, for one such example).

To conclude, the system of peer review is fundamental to the development and dissemination of scientific discovery (Houry et al., 2012). It is necessary to continue developing

methods to improve the peer review system. This study provided preliminary data illustrating variables Editors and AEs consider when choosing reviewers, deciding on the composition of review panels, and choosing editorial board members.

ACKNOWLEDGEMENTS

The authors thank John Borrero, Jim Carr, Caio Miguel, Tyra Sellers, and Jason Vladescu for their comments on early drafts of the manuscript and survey. We also want to thank Anne Brodsky for providing consultation on qualitative data analysis and interpretation, as well as Ann Jeanette Santos for creating the survey In Qualtrics. The authors have no conflict of interest to declare.

CONFLICTS OF INTEREST

The authors have no conflict of interest to declare.

DATA AVAILABILITY STATEMENT

Under Supporting Information, the reader can find the survey questions (Supporting Information A), open-ended answers to questions grouped in different sections (Supporting Information B, D, F, and H), and thematical analysis of open-ended questions fr each section (Supporting Information C, E, G, and I).

REFERENCES

- Bedeian, A. G., Van Fleet, D. D., & Hyman, H. H., III. (2009). Scientific achievement and editorial board membership. *Organizational Research Methods*, 12(2), 211–238.
<https://doi.org/10.1177/1094428107309312>
- Carson, T. L., Aguilera, A., Brown, S. D., Peña, J., Butler, A., Dulin, A., Jonassaint, C. R., Riley, I., Vanderbom, K., Molina, K. M., & Cené, C. W. (2019). A seat at the table: Strategic engagement in service activities for early career faculty from underrepresented groups in the academy. *Academic Medicine: Journal of the Association of American Medical Colleges*, 94(8), 1089–1093. <https://doi.org/10.1097/ACM.0000000000002603>
- Cengher, M., & Kodak, T. (2022). Development of a reviewer mentoring program in the analysis of verbal behavior. *The Analysis of Verbal Behavior*, 38(2), 85–92.
<https://doi.org/10.1007/s40616-022-00172-x>
- Cengher, M., & LeBlanc, L. A., (2023). Reviewing manuscripts for behavior analytic journals: A primer. *Journal of Applied Behavior Analysis*.
- Christenbery, T. L. (2011). Manuscript peer review: A guide for advanced practice nurses. *Journal of the American Academy of Nurse Practitioners*, 23(1), 15–22.
<https://doi.org/10.1111/j.1745-7599.2010.00572.x>
- Duchesne, S., & Jannin, P. (2008). Proposing a manuscript peer-review checklist. *Neuroimage*, 39(4), 1783–1787. <https://doi.org/10.1016/j.neuroimage.2007.09.065>
- Falkenberg, L. J., & Soranno, P. A. (2018). Reviewing reviews: An evaluation of peer reviews of journal article submissions. *Limnology and Oceanography Bulletin*, 27(1), 1–5.
<https://doi.org/10.1002/lob.10217>
- Garbarino, S., & Holland, J. (2009). *Quantitative and qualitative methods in impact evaluation and measuring results* [Discussion Paper, University of Birmingham].

- Houry, D., Green, S., & Callaham, M. (2012). Does mentoring new peer reviewers improve review quality? A randomized trial. *BMC Medical Education*, 12(1), 1–7.
<https://doi.org/10.1186/1472-6920-12-83>
- Lochmiller, C. R. (2021). Conducting thematic analysis with qualitative data. *The Qualitative Report*, 26(6), 2029–2044. <https://doi.org/10.46743/2160-3715/2021.5008>
- National Science Board, National Science Foundation. (2020). *Science and Engineering Indicators 2020: The State of U.S. Science and Engineering*. NSB-2020-1.
<https://nces.nsf.gov/pubs/nsb20201/>
- Nosik, M. R., Luke, M. M., & Carr, J. E. (2019). Representation of women in behavior analysis: An empirical analysis. *Behavior Analysis: Research and Practice*, 19(2), 213–221.
<https://doi.org/10.1037/bar0000118>
- Pranić, S. M., Malički, M., Marušić, S. L., Mehmani, B., & Marušić, A. (2020). Is the quality of reviews reflected in editors' and authors' satisfaction with peer review? A cross-sectional study in 12 journals across four research fields. *Learned Publishing*, 34(2), 187–197.
<https://doi.org/10.1002/leap.1344>
- Rawat, S., & Meena, S. (2014). Publish or perish: Where are we heading? *Journal of Research in Medical Sciences: The Official Journal of Isfahan University of Medical Sciences*, 19(2), 87–89.
- Song, E., Ang, L., Park, J. Y., Jun, E. Y., Kim, K. H., Jun, J., Kim, K. H., Jun, S., Park, S., & Lee, M. S. (2021). A scoping review on biomedical journal peer review guides for reviewers. *PLoS One*, 16(5), e0251440. <https://doi.org/10.1371/journal.pone.0251440>
- Schüttpelz-Brauns, K., Stosch, C., Matthes, J., Himmelbauer, M., Herrler, A., Bachmann, C., Huwendiek, S., Huenges, B., & Kiessling, C. (2010). Recommendations for reviewing a

- manuscript for the *GMS Zeitschrift für Medizinische Ausbildung*. *GMS Zeitschrift für Medizinische Ausbildung*, 27(5), 1–14. <https://doi.org/10.3205%2Fzma000712>
- Slane, M., & Lieberman-Betz, R. G. (2021). Using behavioral skills training to teach implementation of behavioral interventions to teachers and other professionals: A systematic review. *Behavioral Interventions*, 36(4), 984–1002.
<https://doi.org/10.1002/bin.1828>
- St. Peter, C. C., Brand, D., Jones, S. H., Wolgemuth, J. R., & Lipien, L. (2023). On a persisting curious double standard in behavior analysis: Behavioral scholars' perspectives on procedural fidelity. *Journal of Applied Behavior Analysis*, 56(2), 336–351.
<https://doi.org/10.1002/jaba.974>
- Tennant, J. P., & Ross-Hellauer, T. (2020). The limitations to our understanding of peer review. *Research Integrity and Peer Review*, 5(1), 1–14. <https://doi.org/10.1186/s41073-020-00092-1>
- Tennant, J. P. (2018). The state of the art in peer review. *FEMS Microbiology Letters*, 365(19), fny204. <https://doi.org/10.1093/femsle/fny204>

TABLE 1 Survey results for Sections 1 and 2

Question and Response Options	Response Option Count
With which race do you identify? (Select all that apply)	
American Indian or Alaskan Native	0
Asian	4
Black or African American	2
Hispanic or Latinx	12
Native Hawaiian or Other Pacific Islander	0
White	76
Prefer not to answer	3
Specify: _____	1
With which gender do you identify?	
Cisgender woman	41
Cisgender man	43
Transgender woman	0
Transgender man	0
Gender non-binary	1
Prefer not to answer	6
Specify: _____	2
For how many years have you been conducting editorial work (i.e., served as an editorial board member, [guest] associate editor, or editor) for scientific journals?	
0–1 years	2
2–5 years	11
6–10 years	23
11+ years	57
What roles do you currently fulfill for the academic journal(s) that you serve? Select all that apply.	
Editorial board member	91
Guest Associate Editor	84
Associate Editor	60
Editor	22
Other (please indicate):	6

Note: In the “specify” or “other” rows, participants were asked to indicate what specific

demographic information they identify with, or professional experience they have, respectively.

For the question about gender that participants identify with, two selected “Specify” and indicated

that they identify as male; as such, in the manuscript we reported that the n of cisgender men was 45.

TABLE 2 Survey results for Sections 3, 4, 5, and 6

Section and Item	<i>M</i> (SD)	Number and Percentage (out of 93) of Participants who Selected <i>Agree</i> or <i>Strongly Agree</i>
Section 3: Rate the extent to which the following strategies are useful in identifying new reviewers (i.e., individuals who have not performed reviews for you before).		
Identifying scholars who conduct thematically related research	4.5 (0.5)	93, 100%
Searching the reference list of the manuscript to be reviewed	4.1 (0.7)	81, 87.1%
Selecting editorial board members	4.0 (0.8)	77, 82.8%
Using search tools (e.g., use predetermined keywords to search for articles and authors on a topic)	3.9 (0.8)	71, 76.3%
Identifying early career researchers or junior colleagues	3.8 (0.6)	69, 74.2%
Asking for suggestions from reviewers who decline	3.6 (1.0)	62, 66.7%
Identifying scholars who publish in the respective journal	3.7 (0.8)	58, 62.4%
Asking for recommendations from your network (e.g., asking colleagues for recommendations of good reviewers for stimulus equivalence research)	3.5 (0.9)	57, 61.3%
Identifying individuals who are members of underrepresented groups	3.2 (0.9)	35, 37.6%
Asking your students (not based on any other strategy listed above or below)	2.5 (1.0)	15, 16.1%
Other (please indicate):		
Section 4: Rate the extent to which the following you find useful in selecting a review panel.		

Stratify based on the expertise of the reviewers	4.8 (0.5)	91, 97.9%
Avoid known conflicts of interests	4.6 (0.6)	89, 95.7%
Match based on at least one reviewer's and the author(s)'s area of expertise (e.g., both verbal behavior researchers)	4.5 (0.6)	88, 94.6%
Balance of reviewers' typical level of analysis (e.g., detail-oriented, big-picture focus)	4.1 (0.7)	75, 80.6%
Include at least one junior reviewer	3.5 (0.7)	47, 50.5%
Stratify reviewers based on rank (e.g., one student, one junior, one senior)	3.2 (0.9)	37, 39.8%
Match based on other variables (e.g., both the reviewer and the author are members of an underrepresented group, such as non-native English speakers)	3.1 (1.0)	31, 33.3%
Match between the rank of at least one reviewer and the author(s) of the manuscript (e.g., senior author->senior reviewer)	2.8 (1.1)	19, 20.4%
Include only senior reviewers	2.7 (0.9)	17, 18.3%
Other (please indicate):		
Section 5: Rate the extent to which the following contribute to your decision to ask someone to review new manuscripts that you handle.		
The individual has expertise in a domain where you need reviewers	4.8 (0.4)	93, 100%
The individual wrote a high-quality review that helped you in the editorial process	4.6 (0.5)	91, 97.9%
The individual writes reviews that are respectful and constructive	4.6 (0.6)	88, 94.6%
The individual submitted the review on time	3.8 (0.9)	63, 67.7%

The individual publishes articles in the journal where you serve as an Associate Editor or Editor	3.4 (1.1)	56, 60.2%
The individual is a student or an early career scientist who can benefit from writing reviews	3.5 (0.8)	54, 58.1%
The individual publishes frequently	3.5 (0.8)	48, 51.6%
The individual is a member of an underrepresented group	3.1 (0.9)	28, 30.1%
The individual carefully copyedits the manuscript	2.4 (1.1)	13, 14.0%
Section 6: If the manuscript will undergo a second round of revisions (i.e., an invitation to resubmit), how do you select the new panel of reviewers? Select the strategy that you employ most frequently.		
Keep some reviewers, and add at least one new	N/A	70, 75.3%
Only prior reviewers	N/A	23, 24.7%
Only new reviewers	N/A	0, 0.0%
You indicated that sometimes you don't ask the same reviewers to serve on the second panel of revisions. Rate the extent to which the following reasons lead to your decision of not to invite a reviewer to serve on the second panel of revisions. Note. The following question was asked of the 70 participants who responded "yes" to the question, "keep some reviewers and add at least one new;" thus, the number and percentage are reported out of 70 respondents.		
The review was not helpful in the editorial process	4.4 (0.8)	66, 94.3%
The reviewer has two or more manuscripts that they are currently reviewing	3.7 (1.1)	45, 64.3%
The review was harsh	3.5 (1.2)	39, 55.7%
The review was not submitted in time	3.2 (1.0)	30, 42.9%
The reviewer recommended rejecting the manuscript, and you decided to invite resubmission	2.5 (1.1)	14, 20.0%

Note: The mean (*M*) and standard deviations (*SD*) represent aggregate data from the 5-point Likert scale, ranging from not at all useful or not at all influential (1) to extremely useful or highly

influential (5) agreement. The denominator for percentages was the total number of participants (93).

TABLE 3 Survey results for Sections 7 and 8

Section and Item	<i>M</i> (SD)	Number and Percentage (out of 93) of Participants who Selected <i>Agree</i> or <i>Strongly Agree</i>
Section 7: Rate the extent to which the following influence your evaluation of the quality of a review.		
The reviewer evaluated whether the research design is appropriate	4.6 (0.5)	93, 100%
The reviewer evaluated whether the science underlying the main idea is strong (i.e., the aim or hypothesis is well-founded)	4.6 (0.6)	92, 98.9%
The review notes whether the interpretation of the data is accurate	4.4 (0.6)	90, 96.8%
The reviewer evaluated whether the authors used the correct interpretative system (e.g., statistical analyses, visual inspection, effect sizes) to analyze the data	4.4 (0.6)	89, 95.7%
The reviewer evaluated whether the article makes a meaningful contribution to the field	4.4 (0.6)	89, 95.7%
The reviewer evaluated whether the methods are explained well enough that the experiment can be replicated	4.4 (0.6)	89, 95.7%
The reviewer evaluated a study based on the rationale, methods, and results, rather than one or just a few of these variables	4.4 (0.7)	85, 91.4%
The reviewers interpreted the data in a conceptually systematic manner	4.2 (0.7)	85, 91.4%
The reviewer evaluated whether the cited literature is current and related to the research topic	3.9 (0.7)	76, 81.7%
The reviewer provided line edits	2.6 (1.0)	18, 19.4%
Other (please indicate):		
Section 8: Rate the extent to which the following contribute to your decision to nominate or vote for someone to become a member of the editorial board.		
The individual writes reviews that are respectful and constructive	4.6 (0.6)	89, 95.7%

The individual wrote a high-quality review that helped you in the editorial process	4.5 (0.7)	87, 93.5%
The individual has expertise in a field where you need reviewers	4.5 (0.7)	86, 92.5%
The individual publishes articles in the journal where you serve as an Associate Editor or Editor	4.0 (1.0)	69, 74.2%
The individual submitted the review on time	3.8 (0.9)	65, 69.9%
The individual publishes frequently	3.7 (0.8)	63, 67.7%
The individual is a member of an underrepresented group	3.5 (1.1)	56, 60.2%,
The individual is a student or an early career scientist who can benefit from writing reviews	3.3 (1.0)	43, 46.2%
The individual carefully copyedits the manuscript	2.5 (1.1)	16, 17.2%

Note: The mean (*M*) and standard deviations (SD) represent aggregate data from the 5-point Likert scale, ranging from not at all influential (1) to highly influential (5) agreement. The denominator for percentages was the total number of participants (93).