Journal of Purchasing & Supply Management 28 (2022) 100770

Contents lists available at ScienceDirect

Journal of Purchasing and Supply Management

journal homepage: www.elsevier.com/locate/pursup

Editorial

Editorial: From judge to jury: the potential for crowd reviewing

A B S T R A C T

In this editorial, we introduce the four best papers selected from the IPSERA 2021 conference for this special issue of the Journal. We discuss how a double-blind crowd review (CR) process was used as an alternative to the traditional "two reviewer" double-blind peer review system. After a brief review of the literature and the existing debate around academic peer review, we introduce the main characteristics of CR. Next, at the core of our discussion, we report on the results of a pilot CR project that we conducted to review JPSM articles associated with the 2021 IPSERA Conference. We describe in detail the review process, and we illustrate the feedback received from reviewers and authors on the CR process, as well as our perspective as guest editors. Finally, we draw some conclusions and present recommendations for CR in the Purchasing and Supply Management (PSM) field. The CR pilot contributes to the wider debate around peer-reviewing by offering insights into the experience of different stakeholders and by highlighting the benefits and pitfalls of CR.

1. Introduction to the IPSERA 2021 conference special issue

The year 2021 was quite exceptional with Covid-19 shattering our ways of living, working, and traveling. Being successfully held amidst the uncertainty caused by the pandemic, the IPSERA 2021 conference was also exceptional as it was held online for the first time in its 29 years of history. Papers presented online embraced the conference theme, "Purchasing innovation and crisis management", and highlighted the effects of the pandemic on our field, and the diversity of research in our community. The four best papers finally selected for this special issue also center on innovative practices in supply management: supplier scoring methods (Schotanus et al., 2021); value creation through public procurement (Malacina et al., 2022); performance-based contracts (Nikulina and Wynstra, 2022); and supplier diversity and economic inclusion (Sordi et al., 2022). Public procurement, which gained considerable global attention during the Covid crisis, continues to be an important context for research (see Schotanus et al., 2021; Malacina et al., 2022). Papers in this special issue also address their topics in a variety of ways, including systematic literature reviews (Malacina et al., 2022), secondary data analyses (Schotanus et al., 2021), case studies (Nikulina and Wynstra, 2022), and a Notes and Debates paper in an emerging area where the aim is to steer a new research agenda (Sordi et al., 2022).

- Schotanus et al. (2021) explore a technical aspect of public procurement tenders, namely rank reversal. Based on an analysis of public tenders in the Netherlands, the authors show that rank reversal is more of a problem than one could think and make specific recommendations about the desired characteristics of supplier scoring methods.
- Malacina et al. (2022) pose a fundamental question: how does public procurement create value? Through a systematic literature review, the authors isolate the different components of value, the potentially relevant purchasing practices, and the connection between the two.
- Nikulina and Wynstra (2022) explore supplier motivation in taking part in multi-party performance-based contracts. The authors use the lens of expectancy theory to unveil contract design attributes whose presence or absence can foster or prevent supplier motivation and make recommendations for organizations adopting such contracts.
- Sordi et al. (2022) introduce an important topic for sustainable supply chain management literature, the diversity and economic inclusion of suppliers. The authors define the building blocks of economic inclusion across the supply base and propose avenues for future research. A call for papers for a JPSM SI, "Improving Supplier Diversity and Inclusion in Supply Chains" is now open on this topic (JPSM, 2022).

For the papers shortlisted for the IPSERA 2021 conference special issue, we piloted a novel 'crowd review' (CR) process. CR has been fully implemented by some journals, but this is the first time it has been used in our field. In CR, multiple reviewers simultaneously, and in interaction, evaluate a manuscript. The interaction is still blinded but provides opportunities for dialogue about a paper's strengths and weaknesses and it can increase the speed of the reviewing process. In recognizing that reviewing can be onerous, intensive, and slow, we piloted CR for JPSM's IPSERA 2021 Special Issue, to explore whether having multiple reviewers conducting their reviews interactively with each other could add value and increase the efficiency of parts of the review process. In this Editorial, we reflect on the pilot to assess whether our community could benefit from CR. We outline the assumptions in a traditional review process in JPSM, discuss how CR works, elaborate on the CR pilot for the IPSERA 2021 conference special issue, discuss reviewer, author, and guest editor feedback on the CR process, and make some suggestions for future use.

https://doi.org/10.1016/j.pursup.2022.100770
Received 7 April 2022; Accepted 10 April 2022
Available online 30 April 2022

2. Examining the traditional review process

Traditionally, prospective manuscripts that pass initial desk review in JPSM go through a process of double-blind peer review, usually to two or three reviewers, and subsequent revision rounds, before any are accepted and published. The traditional review process, in an ideal situation, has the following characteristics:

- Double blinding ensures that reviewers do not know who the authors are, and vice versa, to reduce potential bias, nor do reviewers know who the other reviewers are and how they are evaluating the manuscript;
- Peer review provides an assessment of research quality, contributes to improving the quality of published manuscripts, and can potentially detect errors and prevent fraud (Smith, 2006);
- All reviewers, as subject and methods experts, are expected to provide constructive and in-depth reviews assessing study positioning, quality and completeness of the literature review/conceptual background, theoretical soundness, methodological rigor, and ultimately a sufficient novel and relevant contribution; and
- Based on the reviewers' recommendations, an Associate/Handling Editor provides a summary judgment in the categories of reject, revise and resubmit, or accept. For revisions, the author(s) is given a clear indication of the main areas for development. An Editor-in-Chief checks the review team's work providing additional expertise if required and makes a final decision. This cycle may be repeated one or more times until the paper reached the publishable standard, or is rejected.
- Rejection rates are typically very high. At JPSM, for example, in 2021 75% of submissions were desk rejected (i.e., without being sent out to reviewers) and 15% were rejected after review – only one in ten papers submitted is published. While considered the 'gold standard' by most journals, the peer review process is not without pitfalls (Smith, 2006), as it is still ultimately a human system (Rennie, 2016).
- As most reviewers are not paid for their work, a recent estimate finds that the combined total cost of providing reviews is a billion-dollar donation, equating to roughly 100 million productive hours in a single year (Azcel et al., 2021);
- Reviews from multiple reviewers of the same manuscript are also often inconsistent, as reviewers have different perspectives, and expectations, and they have gaps in their knowledge, which also differ between reviewers (Peters and Ceci, 1982; Smith, 2006). This creates challenges for authors and editors to reconcile opposing views, for example where reviewer 1 recommends acceptance or minor revisions and reviewer 2 (it is always reviewer 2!) suggests rejection. An associate editor or editor will then have to provide an independent assessment to resolve the conflicting opinions, which often involves going out to new reviewers, adding further time to the process;
- Reviewers should be selected based on their expertise *and* willingness to review, and free from major selection bias of the editors (for example due to their direct connections). Studies from other fields however report issues of geographic bias (Kovanis et al., 2016) and pressures on early career researchers (McDowell et al., 2019);
- Although the peer review itself is double-blind, the full process is not. Editors typically know the authors' names and may be more likely to accept manuscripts from known authors/institutions, a form of ingroup-outgroup bias (Seeber and Bacchelli, 2017)
- Peer review is typically conservative, leaving less chance for innovative manuscripts in terms of methods or theory to be published (Brezis & Birukou, 2020). Similarly, reviewers and editors tend to reject negative or null findings (Smith, 2006) and replication studies hardly ever appear in the same (level) journal as the original investigation first proposing a particular intervention or concept (Gattiker et al., 2021).

3. *What is crowd review?*

While it is still very rare for the vast majority of academic journals, CR is being taken up in some fields, notably in the natural sciences, where single-blind or open (i.e., non-blind) reviewing is more common, and so CR may fit more readily (Ross-Hellauer, 2017). Several journals have adopted CR to varying degrees: Atmospheric Chemistry and Physics (ACP), Frontiers, Synlett, and SynOpen (van Gemmeren and List, 2021) have implemented a form of CR in their process. ACP and Frontiers have implemented a sequence of traditional peer review first and collaborative interactive discussion second (ACP, n.d., online; Frontiers, n.d., online).

In essence, CR uses a shared review platform to speed up the reviewing process and to decentralize decisions away from two or three reviewers to a (small) group of experts that evaluate the manuscript in interaction with each other. The crowd review process has the following characteristics:

- Reviewing becomes participatory and interactive; reviewers see each other's evaluations and can respond to each other, although are still blinded. A shared platform allows for collaboration and transforms reviewing from an individual, independent activity to a community effort (Bobak et al., 2022);
- Crowd reviews could be more efficient if reviewers build upon each other's comments (upvoting or seconding an existing comment) or if reviewers can focus on specific parts of the paper that they are more knowledgeable about, such as the theory or methods;
- Reviews are time-bound to compress the review time while increasing the number of reviewers. In the natural sciences, the allocated review time may be as short as 72 h;
- Journals typically maintain editorial control over the process, selecting reviewers to take part in a crowd (Heinemann et al., 2021; Stoye, 2017). In addition, an (associate) editor makes the final decision;
- Anonymity can still be guaranteed, with platforms automatically pseudonymizing reviewer's real identities, which are visible to editors but not to authors or other reviewers; and
- Shared platforms can include regular office tools such as Google Docs or Office 365, but dedicated software platforms also exist. Platforms on the market include Filestage (www.filestage.io), Frontiers Collaborative Review (https://www.frontiersin.org/about/review-system), or Pubpeer for post-publication review (www.pubpeer.com).

4. The CR process for IPSERA 2021 special issue

For the IPSERA 2021 SI, the IPSERA Executive Committee, JPSM Editors, and IPSERA 2021 SI guest editors agreed to pilot a CR process using a CR platform, to explore opportunities to improve the review process (notably its speed and quality). As speed and reviewer fatigue issues increase as JPSM submissions rise (Tate et al., 2022), CR has the potential to reduce reviewers' effort as they were not expected to review all parts of the manuscript and the imposed timeframe could reduce the

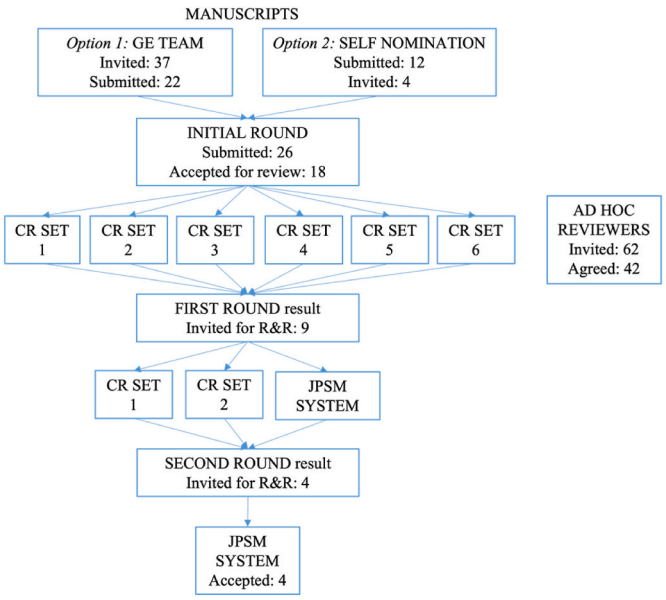

Fig. 1. CR process for the JPSM IPSERA SI pilot.

overall time for review. CR was considered to offer interesting potential for collaborative dialogue to support developmental reviews and the crowd could reduce the risks from an 'outlier' review through 'more eyes' on the paper. Additionally, following training from the software provider, CR was seen to offer a (more) enjoyable reviewing experience.

Fig. 1 illustrates the steps and outcomes in the pilot CR process. Manuscripts submitted via guest editor invitations or self-nominations underwent an initial guest editors' review. Out of 26 articles submitted, 8 were not taken forward for review owing to limitations regarding contribution or the execution of a robust research strategy. Eighteen were selected for the CR pilot. For the first CR review stage, 62 reviewers were invited, with 42 participating. Reviewers were anonymous to each other and authors, but visible to editors and guest editors. Reviewers were given advance notice (about 10–20 days before the actual review period) that they were expected to perform their review in a time window of 10–15 days. The 18 submitted manuscripts were clustered into six sets based on topic similarity and the sets were handled separately by three guest editors. The senior guest editor reviewed the decisions on all the manuscripts and handled manuscripts with contrasting reviews. The number of manuscripts in each set ranged from 2 to 4, and the number of reviewers for each set ranged from 4 to 9.

Reviewers could indicate their feedback on any part of the manuscript by adding text or graphics (like a common text editor) or type comments, to which other reviewers could also comment – all

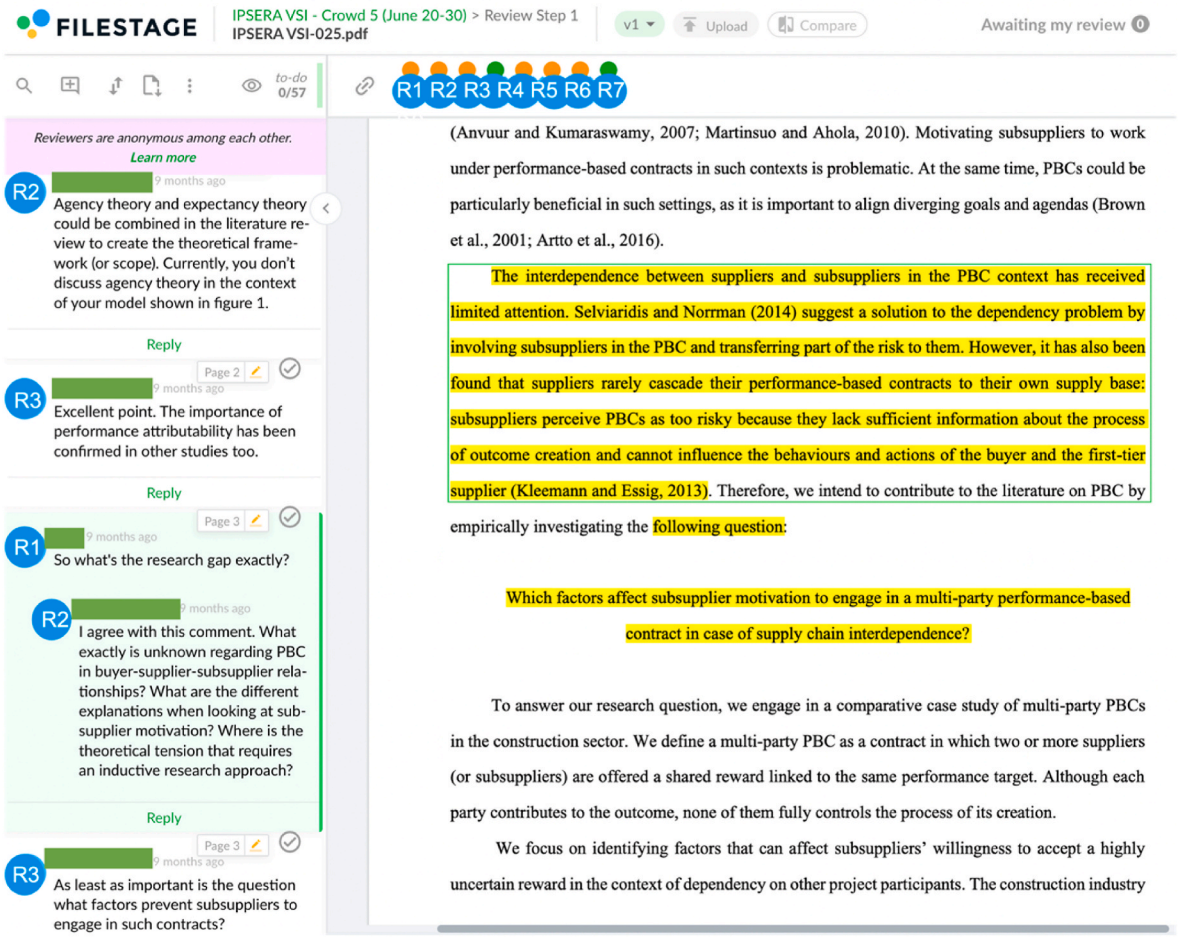

Fig. 2. Screenshot of CR platform.

anonymously. Fig. 2 illustrates a screenshot from the CR platform (with the authors' permission).

In most cases, reviewers provided comments for all the manuscripts in their set whereas in some cases one or two reviewers focused their efforts on certain manuscripts and indicated only a few comments for the others. For some articles, expert methodology reviewers were sought to comment specifically on the method. The number of comments for each manuscript was between 50 and 90, some of which were rather brief asking for clarification or indicating agreement with another reviewer. Some reviewers preferred to add one major remark at the start of the manuscript, summarizing their key comments. A limitation of the platform used was that there was no option for reviewers to indicate their final recommendation and so guest editors emailed reviewers a short survey to confirm their final decisions.

The authors received a detailed guest editor letter highlighting the main comments and prioritizing the required interventions. To avoid authors' fatigue from responding to the high number of reviewer comments on their manuscripts, guest editors allowed authors to address reviewers' comments either by grouping them or responding point-to-point, as authors saw fit. After the first CR round, articles invited for revision were processed either via CR again (e.g., if there was a need for further development, or if there were only a few minor issues), or via the regular JPSM submission system with only two reviewers (mostly relying on the same reviewers involved in the CR process). As a result of the second-round review, four manuscripts were accepted for publication in this SI.

5. Feedback from our community: how was the pilot CR process perceived?

Online feedback polls were distributed to reviewers and authors (including authors of papers not finally selected) to assess their experience with CR and to assess whether it could add value and increase the efficiency of the review process. A series of attitudinal questions were asked predominantly against 5-point Likert-type scales.

5.1. Reviewers' perspective

Reviewers and the SI guest editors were the most involved with the CR platform as they could experiment with the online CR environment and process, whereas the authors received the review report but did not interact directly with the system. Responses from 27 reviewers were received (64% response rate) and are reported in Appendix A, along with a thematic summary of open comments. Reviewers rated the CR as efficient (3.96/5) but had a lower score for its effectiveness (2.92/5), with the overall satisfaction of 3.7 out of 5. They perceived both advantages and disadvantages.

One of the most cited benefits of CR for reviewers was the possibility to see and build on other reviewers' comments. As well as being a motivating factor for reviewers, collaborative CR was seen to add value for authors by focusing effort on the most critical aspects of the manuscript and by cultivating a supportive, collegial, and developmental environment, and limiting excessively positive or negative comments. Whilst still retaining anonymity, collaborative reviewing enabled rich, developmental reviews, and allowed a focus on one's core expertise within a relatively shorter timeframe. Interestingly, seeing other comments were simultaneously reported as a concern as it could influence a reviewer's ability to express an independent, impartial judgment. On a practical level, the proliferation of comments from other reviewers was not always easy to track, increasing the effort it would take to respond to comments in the CR.

While some reviewers felt pressured to complete the review in time, and despite the frustration with the number of automatic notifications received whenever another reviewer completed the task, the crowd activity did serve as an incentive to complete the reviews on time. Reviewers generally liked the system interface although issues were raised regarding functionality and transparency, which could be improved.

The complexity and excessive effort required for CR were identified as problematic for some reviewers, and even if planned, the request to process 3–4 manuscripts in a relatively short amount of time (10–15 days) overwhelmed several reviewers. Potential risks were raised around review quality, as the combination of needing to process large quantities of information and pressures on time could lead reviewers to sacrifice accuracy and take shortcuts. The amount of effort required triggers other concerns, like the increased chance of free riding and the reviewers' frustration with a lack of formal recognition.

In sum, reviewers strongly recommended the CR process for the first rounds of the conference SI rather than for regular submissions to JPSM. Several reviewers suggested running the process in two steps to avoid issues of bias: "*I believe that the first round of review should be conducted in a conventional, i.e., non-collaborative, manner. This is so that the reviewer can make an independent assessment. Instead, platforms such as this can be useful in the subsequent review rounds, because they allow reviewers to work together on a synthesis that can bring together the various requests.*" [Anonymous reviewer].

5.2. Authors' perspective

A total of 20 responses were received from all authors (51% response rate) with some representing multiple manuscripts. Thirteen authors shared their names. Of these, four people represent accepted manuscripts, nine rejected ones. The authors' responses are reported and organized into themes in Appendix B.

Authors were more critical of CR than reviewers, with 80% rating the CR process as poor or very poor, and across the various performance dimensions, efficiency (in terms of turnaround time) was the only one receiving a positive evaluation. All items related to the quality of the review (i.e., effectiveness, clarity, and easiness to respond to reviewers' comments) scored poorly/very poorly.

Despite an acknowledgment of the benefits of speed and multiple constructive comments from CR, there were more comments related to disadvantages. While authors recognize the editors' help in summarizing comments and prioritizing interventions, some still felt excessive pressure to answer all the comments, even though they had the option to group questions/responses. The risk of bias was not mentioned frequently compared to reviewers' comments. In line with reviewers, authors considered CR as a valuable alternative for early-stage manuscripts, but not for regular journal submissions.

5.3. Guest editors' perspective

As an editorial team, we worked closely together and had regular discussions throughout the CR pilot. The pilot demanded an increased workload, but we were able to sustain our enthusiasm to experiment with CR for this Special Issue – although this may not transfer if using CR for regular submissions. We share reviewers' concerns regarding the efforts needed to manage the platform but recognize that this is partly due to the multi-purpose cloud solution that is not specific to academia. Whilst providing functionality benefits, it did require additional manual steps. For example, the lack of a pre-selected pool of reviewers in the CR system meant that we had to manually handle all the workflow that normally is automated in the review system, including scouting

reviewers, inviting them, sending reminders, monitoring progress, and sending confirmation emails. Moreover, we had to send a separate form to reviewers to collect the final recommendations (major revision-minor revision-reject) because this function was not available in the system. This extra step was important because the overall assessment was not always self-evident from the comments. For example, some manuscripts had 90% reject recommendations even though the tone of comments on the manuscript was fairly positive and developmental. As editors, we also missed reviewers' confidential comments to the editor, which is a standard feature of conventional academic review systems.

Coordinating reviews was particularly challenging. For a typical sample crowd of 3 manuscripts and 8 reviewers, we had to confirm that reviewers were available in the same time window to leverage the interaction of crowd reviewing. We also had to negate the practical risks of separating reviewers (42) and authors (39), who were involved simultaneously and with some conducting both roles. We were mindful of potential ethical issues related to the tone of comments stemming from interactivity or potential breaches of anonymity.

Following each CR, the GEs generated a review report (average 2–3 pages) for the authors. The reports synthesized the reviews for rejected papers, and for papers invited for revisions, they also prioritized issues to be addressed for revise and resubmit - a substantial task given that many submissions had 100+ comments with replies and re-replies.

While authors in the survey reported feeling overwhelmed by the volume of comments from the CR, some authors confirmed through unsolicited feedback that they appreciated the quality of the feedback, including from authors with submitted articles that were eventually rejected.

In analyzing the efficiency and effectiveness of the CR for this pilot, we observed variable approaches across the crowds. The crowds that iteratively provided comments over the full 10-day period provided better results that drew on the power of dialogue within the crowd. In these sets, we observed much more interaction between reviewers, and it allowed reviewers to pay more attention to individual sections of the manuscript if they saw other sections had been sufficiently attended to. The crowds that provided inputs only once, and all at the end of the period, had little opportunity for interaction.

6. Discussion

In agreeing to the JPSM editors' call to pilot CR, we were keen to explore how our community could use this alternative to augment traditional processes. Our expectations were not that this would provide a flawless replacement for conventional peer-review, but traditional reviews are also imperfect. In this concluding section, we summarize the feedback from the CR pilot and provide recommendations for IPSERA and JPSM communities, as well as for the larger academic audience interested in CR. We provide guidance in terms of innovating current review processes to improve the trade-offs between scientific rigor, manuscript development, researchers' learning, and efficient use of resources.

An important premise is that the conventional double-blind review is not flawless, as elaborated earlier in this article, and different reviewers can dramatically change the outcome of a review. The whole team of guest editors has recent experiences, no doubt familiar to many, of manuscripts that were rejected from one journal and accepted into another (without fundamental revision). Or worse, changes requested by one journal, before a subsequent rejection, had to be completely reversed to make the manuscript acceptable to the next journal. This is a symptom of the reliance on only two reviewers in conventional reviews. The main reason for limiting the reviewers' number to two is a legacy issue, as historically, editors had to print and mail manuscript copies to reviewers, and reviews were similarly printed and mailed back to authors. It is interesting to note that the review process did not change since then, despite being fully digitized. CR may counteract the risk inherent in reviewer selection by enlarging the pool of reviewers, by delegating the responsibility for identifying strengths and weaknesses to the collective, in which each reviewer adds their expertise, and hopefully then being able to acknowledge and recognize their own knowledge limitations too, deferring specific aspects to others where appropriate. Diverse expertise in a crowd allows for heterogeneity in opinions, leading potentially to more comprehensive and valuable feedback, whilst minimizing the skewed impact from outliers.

As guest editors, we know that finding appropriate reviewers, ensuring good quality reviews, providing clear editorial directions, and doing so in a reasonable amount of time represent a real challenge. Restructuring the way reviewers' pools are created and maintained and consideration of the use of CR might represent a viable solution. As outlined in the various feedback sections above, the additional time invested is probably considered excessive for a regular review process, although the steep learning curve and lack of an integrated repository of reviewers in the CR platform are likely to resolve over time.

The basic tenet of CR leverages the power of the crowd to reconcile the paradox expressed by reviewers: on the one hand, the majority welcome with great enthusiasm the possibility to interact with others; on the other hand, about 25% of reviewers lamented the amount of effort required to review more than one manuscript at once. We believe that the lack of experience with CR should be factored in. Whilst we are mindful of reviewers' time, breaking from traditional review approaches appeared difficult at times, and the different ethos of CR was not necessarily fully communicated or embraced in the pilot. In examining how CR has matured in the natural sciences, we see that individual contributions are smaller than conventional reviews, but the final aggregate result is larger. Akin to crowdsourcing platforms, where great amounts can be raised with a multitude of small donors, or with apps like Duolingo where the emerging translation of many learners is considered better than single professional translators, CR aims to democratize decisions through a more inclusive and transparent process spreading the review activity amongst a greater number of people. Reviewing is not simply a matter of quantity, but also of quality. In this regard, we agree with one reviewer's comment: "*effectiveness per individual reviewer is lower, but with more reviewers and the usual high standard of editor input, I reckon the value for authors will be as good (especially when speed of decision is factored in)*". Once accustomed, reviewers should be able to flexibly adapt to the manuscript and decide where and how much to focus their effort. We observed that this happened naturally in some crowds of our pilot, in line with the initial guidance provided and without the need to closely monitor reviewers.

The reviewer base should be large enough to enable the collaborative benefits of CR, reduce reviewer overload, and allow younger scholars to participate. In absolute terms, reviewers' characteristics are not necessarily different than traditional reviews, even though CR would probably require a broader pool of reviewers to start with, to avoid the repeated involvement of the same reviewers. Reviewers should also be diverse in experience, methods, topics, geography, and so on, to cover the required mix of expertise and perspective for each manuscript. Reviewers from outside our field would be particularly important to assess multidisciplinary and innovative research designs/subjects. The unique feature of CR is the involvement of such a mix of expertise at the same time, instead of doing so across multiple stages of review and journals. In our pilot, reviewer effort was increased because they were each reviewing three or four manuscripts at once, rather than just one as in

Table 1
Challenges and solutions with crowd review (the most important recommendations are indicated in bold).

Challenge	Possible solution
Editors' time required to manage reviewers and manuscript workflow	• **Prepare a pool of selected reviewers upfront** • **Integrate workflow management into the review platform**
Reviewer's effort	• **Ensure a sufficiently large pool of reviewers** • Provide clear review guidelines (e.g. not necessary to review the entire manuscript, interact with other reviewers) • Balance the number of articles and reviewers per crowd • Accumulate experience and develop best practices
Reviewers' bias	• **Ensure diversity of reviewers (different backgrounds, expertise, seniority, culture)** • Periodically reshuffle the reviewer crowds • Editorial assessment and guidance
Authors' confusion	• Ensure that editors combine all comments into a series of recommendations • Provide clear guidelines to authors on how to handle reviewers' comments • Allow authors to respond to comments directly on the platform
User interface and technical aspects	• Integration with publisher system • Reviewer database and management • Workflow management • Distinguish type of users • Enable function for final decisions and editors' comments • Make the process transparent

conventional reviews. However, having reviewers work on multiple manuscripts of a similar topic might be an efficient option in specific situations, such as conferences or special issues, or when a methodological check is needed on multiple manuscripts.

More doubts on the value of CR arise for regular journal submissions, rather than focused and timebound special issues, but we have offered our recommendations on how to anticipate possible issues, see Table 1. While CR has been adopted successfully in other fields, we suggest that our community is not yet sufficiently mature to adopt CR in the initial stages of review of a full manuscript and our systems make it hard to accommodate. Indeed, following the examples of other journals, we would need a large, flexible, and diverse set of reviewers with the capacity to review in short timeframes, adapting their approach and expert focus to specific parts of a manuscript to fully exploit the power of the crowd. Editorial teams, in turn, would play a more prominent role in synthesizing reviewers' comments and providing authors with clear follow-up guidelines.

The risk of bias in CR was emphasized by many reviewers, yet we found little, if any, evidence of this in the pilot, and as with conventional reviews, the authors and reviewers remained double-blinded throughout. Our impression is that reviewers were overzealous in indicating the risk of bias. Heterogenous review comments were observed in many reviews, suggesting that while there is a *possibility* of generating a bias through seeing others' comments, this did not translate into an actual bias. Further, conventional review processes are also not bias-free, linked with reviewers' homophily (Brezis and Birukou, 2020), both in their selection and their ontological positions and research design preferences. Carefully crafted review guidelines, CR training, editorial oversight, and experience with CRs should be sufficient to exploit the benefits of reviewers' confrontation (as they report) and minimize the risks. Another alternative might be to split the review process into separate steps and allow the reviewers to interact when it is most beneficial.

A series of minor (but important) technical aspects would greatly ease a CR process, see Table 1. The main obstacle was the lack of integration with the publisher platform which prevented automation of workflow. Relatedly, reviewers were free to choose when to review, but the pilot results suggest that CR works best by reserving reviewers' time in advance (for example through a booking system) and making sure that all reviewers will provide their feedback in a specified time frame (e.g., one or two weeks). Critical functions should be considered for any CR platform adoption, specifically, the provision for final reviewers' decisions (i.e., reject, revise, accept) and the ability to give confidential comments to the editor. The different types of users should be distinguished and ideally, a dedicated menu for summarizing editorial comments should be available.

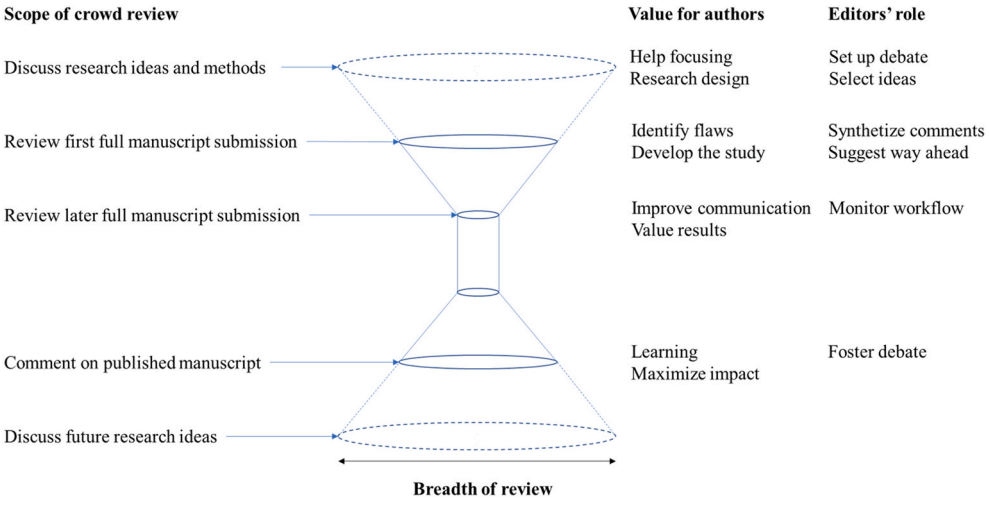

Fig. 3. Crowd review at different research stages.

7. Conclusions

An important question remains from the pilot of CR for the IPSERA 2021 SI: would we recommend our community to adopt CR? Our pilot experienced inherent challenges and exposed concerns and frustrations, but we believe that CR can be a viable alternative to augment conventional reviews in a variety of situations. At the heart of this view is an increasing recognition that conventional peer review is under strain and is suboptimal (Brezis and Birukou, 2020), with claims of bias and parochialism leading to calls for the rigor of peer review to be reconsidered (Casnici et al., 2017). As far back as 1999, the *British Medical Journal* was debating issues of reviewer bias and lack of transparency. In the words of Smith (1999, p. 5) "*Peer review will become increasingly a scientific discourse rather than a summary judgment.*"

In our view, CR is a tool that can serve different objectives, and that can be tailored accordingly for our community. Fig. 3 provides an overview of the possible uses of CR at different research stages: the value for authors as well as the role for editors within CR increase from bottom to top. Although outside of the scope of the pilot, in the survey all groups highlighted the potential of CR for paper development workshops, such as the IPSERA Doctoral Workshop, or conference submissions and book chapters. In other fields CR is used for reviews of research protocols or data (Fitzpatrick, 2010), and even open theorizing, in which multiple authors/reviewers collectively analyze research material such as data or script (Leone et al., 2021). In a developmental mode, CR can be a powerful tool to discuss alternative research ideas and help authors find the best alignment between research questions and available methods, *before* conducting a study. This kind of discussion may not happen within the boundaries of a journal but could take place on other meet-up occasions promoted by academic associations like IPSERA. Recent examples show that some journals are already using new forms of gated processes to select potential manuscripts based on extended abstracts or research protocols, even though CR is not yet adopted (JOM, 2022; PMR, 2022).

In contrast to using CR to develop new research, it could be beneficial at the latter stage of manuscript development, once the main concerns have been addressed. Here the crowd can help authors fine-tune the argumentation thread to improve the readability, and hopefully the impact, of the research. In our pilot, there was some success using CR in this way and several reviewers suggested holding the process in two stages, the first one being a conventional review and only the second being subject to CR.

Reviewing does not necessarily have to stop with the manuscript publication, as extant open review examples show (e.g., Ross-Hellauer, 2017). Journals like the *Academy of Management Journal* regularly publish commentaries to published articles and responses from authors to create debate, but this is still done traditionally and statically. CR can build diverse discourse with our community – scholars, organizations, policymakers – to engage in debates about published studies and encourage reflection and idea generation for future research. Platforms like Pubpeer.com have been built for this purpose of post-publication review.

In conclusion, we are keen to stress that despite the effort required, the complaints received (and the appreciations) as a guest editor team, we did not lose our initial enthusiasm for CR as a tool to augment the development of research in our field, and to support collaborative communities. Challenging the status quo of conventional review processes has opened rich conversations and ideas around how we can ensure the integrity of robust processes, while simultaneously enabling and supporting innovative and diverse research. There is a legitimate risk that CR drains resources, but – if properly managed – it can improve the reviewing process, in terms of both quality and speed. If we can assure quality, the ability to speed up the review process is likely to be increasingly important. Returning to the context of the IPSERA 2021 conference, the pandemic affected every aspect of our lives. Our ability to respond to the changing world demands a delicate balance of quality and speed – and of course, collaboration.

Acknowledgments

We would like to acknowledge the contributions of all the authors and reviewers who took part in the crowd reviewing pilot. Thanks are extended to Louise Knight and Erik van Raaij for initiating the pilot, arranging training, and conducting the feedback survey. The pilot used Filestage as the crowd review platform, and we thank them for allowing us complimentary access.

Appendix A. Feedback from reviewers involved in the crowd review (CR) pilot

As illustrated in Figure A1, 60% of reviewers viewed themselves as 'experienced/highly experienced', with the rest being more junior scholars who were invited to ensure a balance in the reviewers' crowd. The respondents were motivated reviewers, as all of them value reviewing as a way to learn about research, with an average of 3.9 on a scale from 1 = "no personal value" to 5 = "highly valuable" (and no one selecting 1 or 2).

The overall evaluation of the CR platform was 3.7 on a scale from 1 = "very poor" to 5 = "excellent" (with only three people selecting 1 or 2). This can be considered a positive outcome, even though the system is likely perceived as improvable, given that 26% of respondents assigned a neutral evaluation. The editorial team was conscious of the pressures on reviewers' time and the lack of flexibility in the process. For example, reviewers were typically required to assess two or three manuscripts at the same time, consider the comments from other reviewers, and provide a recommendation in a two-week timeframe. Looking at the aggregated responses, respondents perceive CR as more efficient than conventional reviewing (average 3.96/5) but less effective (with an average of 2.92). The main expected advantage, as perceived by reviewers, was added value for authors who have been able to receive a higher number of comments than usual in a shorter time. Yet, this only averaged 3.2, and reviewers felt that while reviewing was somewhat easier/faster, it was probably less in-depth.

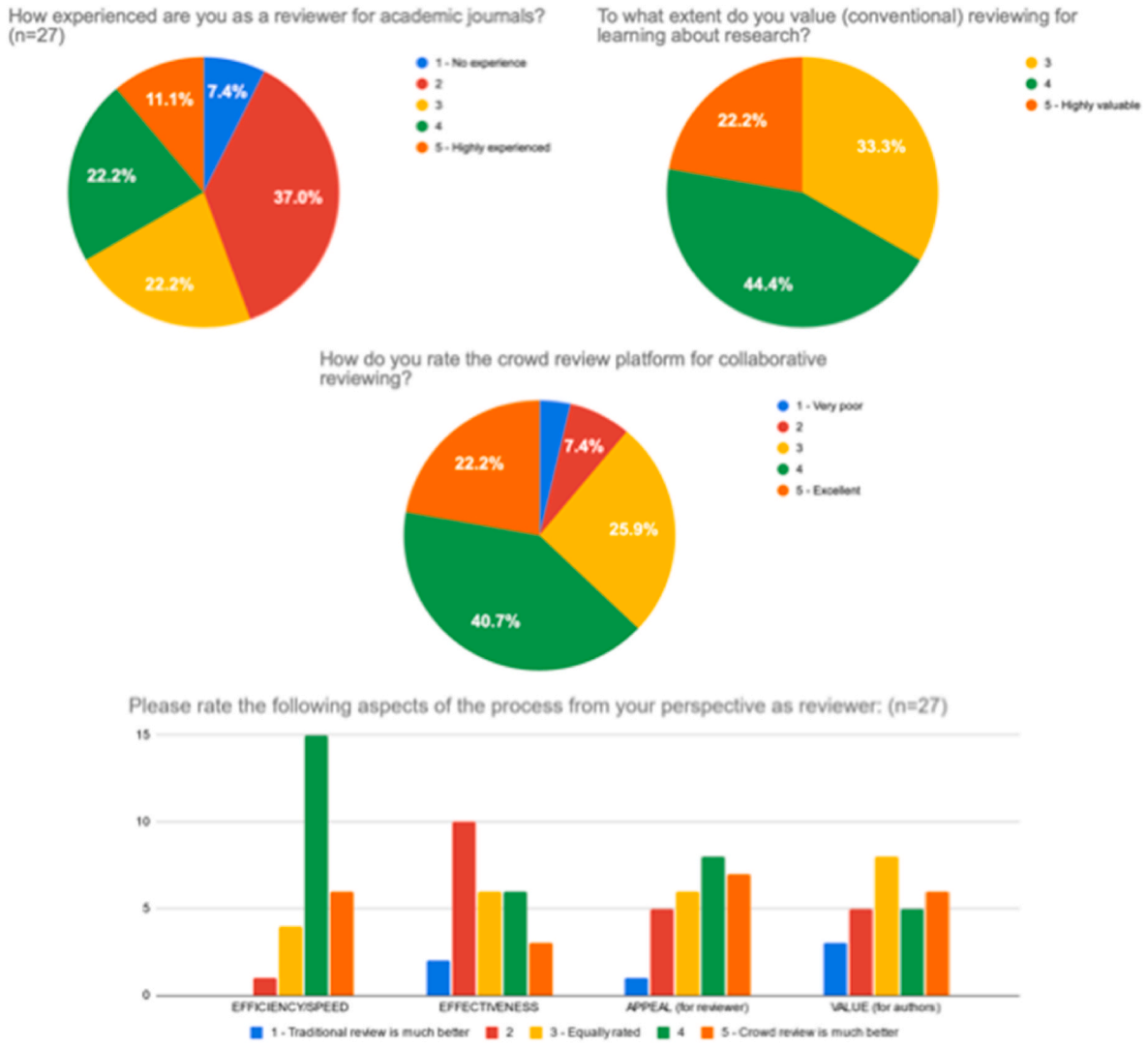

Fig. A1. Summary of reviewers' survey

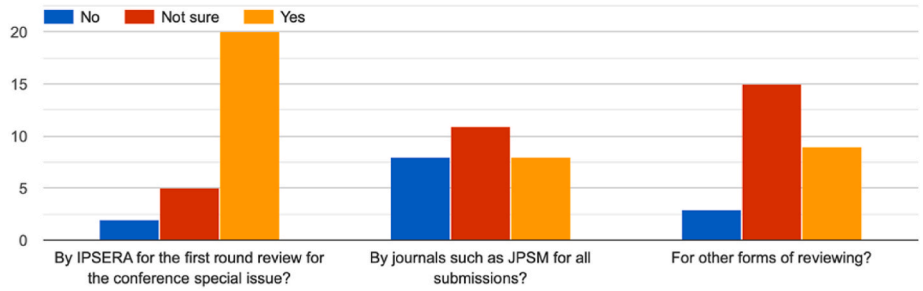

Fig. A2. Reviewers' recommendations

Table A1
Main pros for reviewers

Pros	Quotations
Teamwork	• *"I got to see other reviewers' comments and reiterate them where appropriate."* • *"What I enjoyed is that you can see other reviewers' feedback - this is an author benefit also"* • *"Felt less pressure to be expert on all aspects, but it does rely on [a] group of reviewers."* • *"… you can also see the progress of other reviewers, which additionally motivates you to complete your own review"* • *"… it was good to already be able to react to comments by other reviewers."* • *"I like the opportunity to respond to another reviewer's comments!..Would have liked to have more chances for interaction between reviewers."*

(*continued on next page*)

Table A1 (*continued*)

Pros	Quotations
	• *"Also it will help with a review of reviewers, if a critique appears overly critical or insensitive there is some scope for another reviewer to soften it. I loved this process."*
Developmental review	• *"The platform is very interesting and helpful to provide more precise and constructive comments directly on the paper."*
	• *"… sometimes reviewers go in different directions and AE's don't always provide clear direction on which direction to go making it harder for authors to respond."*
	• *"Worked better as iterative process where I went back into the review at least twice to see and respond to developing comments which requires reviewer resource and commitment."*
	• *"Inspiration from other comments and indication of points that you may not have recognised yourself or that you had a different attitude"*
	• *"Valuable feedback for the authors through many very helpful comments."*
	• *"I think the open feature of this format is valuable, both in terms of learning but not least to keep a good and sober tone in the comments (constructive, forward thinking, but also honest if the paper was not good enough)"*
Focus on core competence	• *"I could comment on pieces of the paper I knew most about/was more interested in reviewing."*
	• *"… it allows you to comment more specifically on your area of expertise and not to feel you need to comment on an area where you (a) are not an expert (b) can see others have contributed sufficiently with their expertise."*
	• *"I like not needing to comment on all aspects of a paper - for example if I am not an expert on a certain method that I can refrain from commenting [on] that part".*
Overall experience and user interface	• *"The platform is very good, and very intuitive for the reviewer. It was easy, and I could also see the other reviewers' comments, comment on them"*
	• *"Really enjoyed the process."*
	• *"The platform was great and also the experience."*
	• *"It was a very neat approach … it was fun, and definitely a new and intriguing approach!"*
	• *"Thanks for a great effort and the initiative in taking the scholarly interaction into the digital era!"*
	• *"I thought this was brilliant, can I also say [guest editor] was a joy to deal with."*
	• *"The platform is quite user friendly and nice to use."*
Speed of review	• *"Urgency to provide feedback in due time"*
	• *"On the positive note I think getting notifications about reviews done by other reviewers encourages you to do the same, so timeliness increases."*
	• *"I reckon the value for authors will be as good (especially when speed of decision is factored in)"*
	• *I like … the speed*

Table A2
Main cons for reviewers

Cons	Quotations
Bias	• *"… sometimes you can feel a bit influenced by the other reviewer, maybe if we could have the option to do our own review first and after interact with the other reviewer this can make the process much more productive"*
	• *"Potential for reviewers to "go along" with other comments as a risk."*
	• *"I think it would be good to organize the reviewing process in two rounds, one in which it's not possible to see other reviewers' comments and the second in which each reviewer can comment others."*
	• *"Seeing other reviewer's comments introduces biases and makes it very difficult for reviewers to conduct a fully independent and objective review."*
	• *"… a reviewer' neutral view may be biased by an overly positive or negative review already visible, with a reviewer maybe tending to follow suit with the recommendations and themes expressed, and not developing their own independent judgment."*
	• *"I am not sure if seeing comments of other reviewers while doing your own work is a good or a bad thing. It could be a bit distracting."*
	• *"Even if I am not an experienced reviewer, the approach does not seem very "unbiased" to me, as you can see in real time the contributions."*
	• *"Seeing comments of others is a dubious aspect: do reviewers remain objective?"*
Complexity/Effort	• *"It was much harder to concentrate on reviewing itself. Some papers got over 89 comments! As an author and reviewer I think it's better to have max 4 reviewers working on the same manuscript to be able to provide well written, holistic, fully written feedback, not bunch of tiny comments in different parts of the paper."*
	• *"Being asked to review four papers in one sitting was too much."*
	• *"I fear the massive amount of work for the Editor/AE to synthesize such unstructured comments - that not everyone is willing to take."*
	• *"… it was hard to see where the own comments were and distinguish them from other reviewers. also getting back to the paper several times is a nice idea, but also more time consuming … for the reviewer"*
	• *"The time frame was too limited also considering that this is an intense period with exams and conferences"*
	• *"I think to review 3 papers within 10 days is a bit difficult, and the quality of reviewing is not the same."*
Quality of review	• *"I fear the quality of feedback that will be provided to the authors could be deteriorated, and the confusion in the review steps could increase."*
	• *"In comparison to regular review, need to ensure you still get the overall review and key points because this could be missed as they are higher level summaries rather than in-text comments."*
	• *"… the organized and formal comment of the reviewer at the end of the review should be required to facilitate the editors work"*
	• *"This approach might be quicker, but sacrificing accuracy for speed is not appropriate for a formal review process."*
System integration and Transparency	• *"The platform should include the possibility to provide a final recommendation (instead this was asked through a separate form)"*
	• *"There was no 'reject' option for papers."*
	• *"Not clear to me how the editors make a decision based on the multiple and not always consistent reviews."*
	• *"At least in our group, there has been no decision as to whether and which paper will go forward, so it would be nice to be informed promptly."*
	• *"Would like to give an overall assessment of each manuscript with a limited set of rating scales and an overall "reject, major revision, minor revision, accept" recommendation to the editor."*
Free riding	• *There is a high risk of free riding, especially with a limited timeframe and multiple papers to be reviewed."*
	• *"… many of the reviewers seemed to have waited till the last day"*
Authors' frustration	• *"I feel that it might be more difficult to accept a reject decision (i.e., authors might get more frustrated), due to the number of comments received."*
Spam	• *"Receiving an email alert every time a review took place was unnecessary."*
	• *"I started to dislike all the messages the system kept pushing in my email. But apparently, these can be managed by settings."*

Appendix B. Feedback from authors involved in the crowd review (CR) pilot

Respondents were all experienced (74% with more than 4 articles journal publications). The perceived need to change to CR was generally low, with 65% generally satisfied with conventional reviews and a further 18% neutral. From the open comments, authors recognize that the quality of conventional reviews is highly dependent on the editorial team and the quality/commitment of reviewers, but they perceive it as more structured, clearer, and easier to handle than CR.

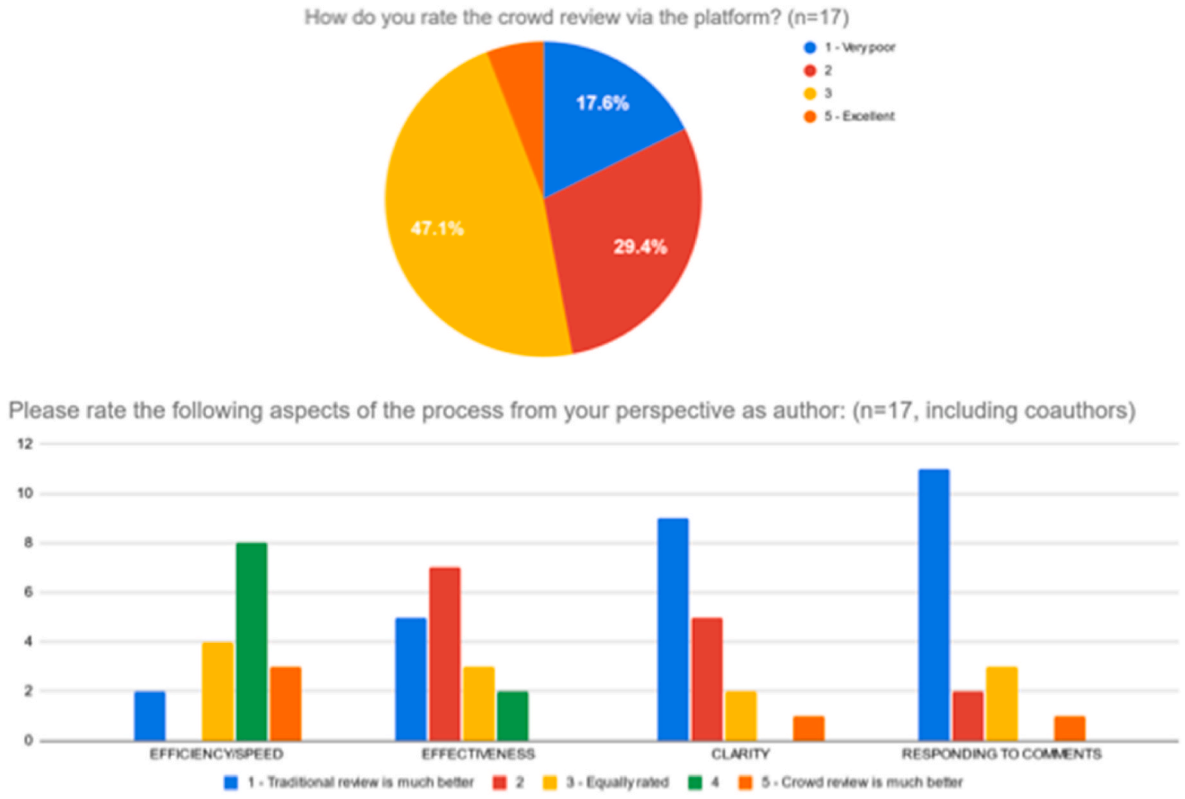

Fig. B1. Authors' perspective

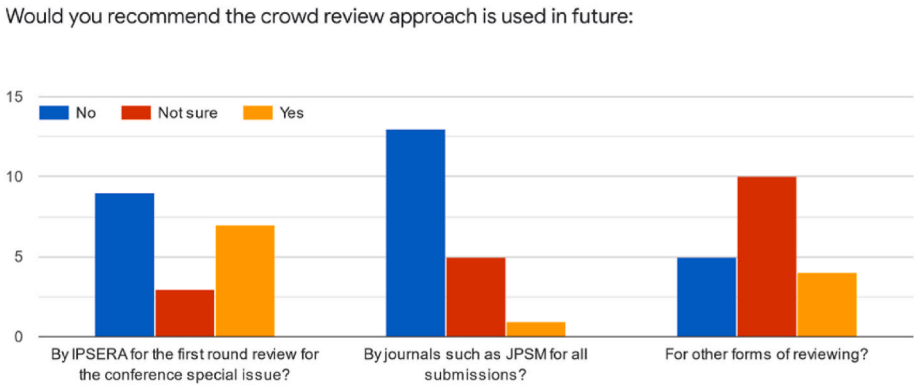

Fig. B2. Authors' recommendations

Table A3
Main pros for authors

Pros	Quotations
Speed of review	• *"The biggest 'pro' is the speed"* • *"The process is quick."* • *"Fast turnaround"* • *"Fast"*
Quality of comments	• *"Most of the comments were very good and useful. They really improved the paper"* • *"Multiple perspectives; reviewers can see and add to other reviewers' comments - consensus forming."* • *"Multiple comments, reviewers can have a fruitful dialogue"*

Table B2
Main cons for authors

Cons	Quotations
Confusion from multiple reviewers	• *"… the problem is that reviewers act the same time"* • *"My impression was that there were too many contradictory comments. Luckily for us, the handling editor did an extremely good work in interpreting the importance of the comments to us authors. However, that left an impression that the view of the handling editor became dominant and maybe also required a lot of work from the editor."* • *"… most comments were scattered across the document, mainly about minor issues, lacking a comprehensive view, often conflicting and contradictory."* • *"I really did not like to receive dozens of different comments, not connected [to] each other and not easy to follow. It looks like a sort of work done among authors of the paper, but this was a review and so it was difficult to understand how to work to actually improve the paper"* • *"I had 10 reviewers (!!). Getting everyone to agree in the review is complex, since they often have different opinions on the research (10 different people give very different advice)."* • *"Comments/criticisms too diverse"* • *"Confusing comments since dialogs might be lengthy and reviewers sometimes disagree what an author should do. The comments were not always full sentences but a couple of words and sometimes it was unclear what the point was"* • *"One major drawback, at least in how the reviews were formatted/shared, was the fragmentation: we had to sort of puzzle together which comments where from which reviewer (this was possible via the IDs but still time consuming)."* • *"What reviewers need a response? What reviewers were going to look at the next version? There needs to be a single point of contact – versus multiple co-editors with differing opinions."* • *"Last year's review round was more like price shooting. Everybody that was involved could "shoot" on "something". As if the patient has ten different doctors trying to cure. The review process may have succeeded but the patient is dead."*
Effort	• *"When we [revised] the paper, we prepared a detailed response to both the editor and all the single comments, which took ages to compile and resulted in a very long document, I am not sure it was really a value added …"* • *"… the downside is it is a lot of work to answer to all the comments; hard to organize them in a nice response without repetition; and some of them really do not add value - yet still require time and thinking about the response."*
Bias	• *"Tending to biased reviews"*
Complexity	• *For me, the crowd review result was just a cumbersome process and a lengthening of time.*
User interface	• *"You should be able to reply to the comments with the same platform to make it really effective."* • *"The PDF with comments was not user friendly … the output of the platform is not convenient"*
Transparency	• *"… no clear outcome from process that I could see i.e., major revisions etc."*
Review thoroughness	• *"Some reviewers look at sections of the paper in isolation i.e., only seem to read and comment on method section"*

References

ACP (not-dated); Peer review. Interactive Review Process. Interactive Public Discussion, Peer Review, And Publication In ACP. Available online from: https://www.atmospheric-chemistry-and-physics.net/peer_review/interactive_review_process.html.

Bobak, C.A., Muse, M., Giffin, K.A., Williamson, D.A., Greene, C.S., Moore, J.H., Wall, D. P., 2022. Human Intrigue: meta-analysis approaches for big questions with big data while shaking up the peer review process. Biocomputing 156–162, 2022. November 2021.

Casnici, N., Grimaldo, F., Gilbert, N., Squazzoni, F., 2017. Attitudes of referees in a multidisciplinary journal: an empirical analysis. J. Assoc. Informat. Sci. Technol. 68 (7), 1763–1771.

Gattiker, T.F., Hartmann, J., Wynstra, F., Pagell, M., Cantor, D., Yan, T., Tate, W., 2021. Testing the shoulders of giants—replication research using registered reports. J. Supply Chain Manag. 1–6. https://doi.org/10.1111/jscm.12276.

Heinemann, M.K., Gottardi, R., Henning, P.T., 2021. Select crowd review": a new, innovative review modality for the thoracic and cardiovascular surgeon. Thorac. Cardiovasc. Surg. 69 (5), 387–388.

JOM, 2022. Call For Papers: Pre-Approved Research Designs For Field Experiments. Available at: https://onlinelibrary.wiley.com/pb-assets/assets/18731317/JOM%20SI%20CFP%20-%20Pre-Approved%20Research%20Designs%20for%20Field%20Experiments-1632748345897.pdf.

JPSM, 2022. A Special Issue for Improving Supplier Diversity and Inclusion in Supply Chains. Available at: https://www.journals.elsevier.com/journal-of-purchasing-and-supply-management/call-for-papers/a-special-issue-for-improving-supplier-diversity-and-inclusion-in-supply-chains.

Kovanis, M., Porcher, R., Ravaud, P., Trinquart, L., 2016. The global burden of journal peer review in the biomedical literature: strong imbalance in the collective enterprise. PLoS One 11 (11), e0166387.

Malacina, I., Karttunen, E., Jääskeläinen, A., Lintukangas, K., Heikkilä, J., Kähkönen, A. K., 2022. Capturing the value creation in public procurement: a practice-based view. Journal of Purchasing and Supply Management 28, 100745. https://doi.org/10.1016/j.pursup.2021.100745. In this issue.

McDowell, G.S., Knutsen, J., Graham, J.M., Oelker, S.K., Lijek, R.S., 2019. Co-reviewing and ghostwriting by early-career researchers in the peer review of manuscripts. Elife. https://doi.org/10.7554/elife.48425.

Nikulina, A., Wynstra, F., 2022. Understanding supplier motivation to engage in multi-party performance-based contracts: the lens of Expectancy theory. Journal of Purchasing and Supply Management. 58, 100746 https://doi.org/10.1016/j.pursup.2022.100746. In this issue.

PMR, 2022. Call For Papers: Strategic Public Procurement: A Public Management Perspective. Available at: https://think.taylorandfrancis.com/special_issues/strategic-public-procurement-public-management-perspective/?utm_source=TFO&utm_medium=cms&utm_campaign=JPG15743.

Peters, D., Ceci, S., 1982. Peer-review practices of psychological journals: the fate of published articles, submitted again. Behav. Brain Sci. 5 (2), 187–195. https://doi.org/10.1017/S0140525X00011183.

Rennie, D., 2016. Let's make peer review scientific. Nature 535, 31–33.

Ross-Hellauer, T., 2017. What is open peer review? Systematic Review. F1000Research 6.

Schotanus, F., van den Engh, G., Nijenhuis, Y., Telgen, J., 2021. Supplier selection with rank reversal in public tenders. Journal of Purchasing and Supply Management 28, 100744. https://doi.org/10.1016/j.pursup.2021.100744. In this issue.

Seeber, M., Bacchelli, A., 2017. Does single blind peer review hinder newcomers? Scientometrics 113 (1), 567–585. https://doi.org/10.1007/s11192-017-2264-7.

Smith, R., 2006. Peer review: a flawed process at the heart of science and journals. Journal of the Royal Society of Medicine 99, 178–182. https://doi.org/10.1258/jrsm.99.4.178.

Sordi, A., Tate, W., Huang, F., 2022. Going beyond supplier diversity to economic inclusion: where are we now and where do we go from here? Journal of Purchasing and Supply Management, 100751. https://doi.org/10.1016/j.pursup.2022.100751. In this issue.

Stoye, E., 2017. Chemistry journal introduces 'intelligent crowd' peer review. In: Chemistry World. Royal Society of Chemistry. Jun. 8.

Tate, W., Di Mauro, C., Carnovale, S., Knight, L., 2022. Transitions, opportunities and challenges – change and continuity at JPSM. Journal of Purchasing and Supply Management 28 (1), 100755. https://doi.org/10.1016/j.pursup.2022.100755.

van Gemmeren, M., List, B., 2021. How and why crowd reviewing works. Synlett 32 (9), 885–891.

Melek Akın Ateş
Sabanci Business School, Sabanci University, Turkey

Davide Luzzini
EADA Business School, Spain

Joanne Meehan
University of Liverpool Management School, University of Liverpool, United Kingdom

Robert Suurmond*
School of Business and Economics, Maastricht University, The Netherlands

* Corresponding author.
E-mail address: r.suurmond@maastrichtuniversity.nl (R. Suurmond).