CSR Evaluation Support Services:

Editorial Board (Two-Stage) Review Survey Report

March 2010

Contents

Background

The Center for Scientific Review (CSR) conducted a pilot of the Two-Stage or Editorial Board Review in support for the trans-NIH initiative to enhance peer review. In the first stage, each grant application was evaluated by 2-3 specialist mail reviewers who were charged with evaluating the proposals based on technical merit. This step was followed by an in-person meeting of reviewers with broader expertise, who focused on the impact and significance of the science, emulating the function of journal editorial boards and giving the pilot its name (Stage 2). Stage 1 reviewers provided written critiques, but did not score applications, which took place at Stage 2. The Two-Stage format was expected to improve several aspects of traditional peer review through several mechanisms:[1]

- Providing specific expertise for a wide range of scientific areas
- Improving the discussion quality by using a small number of discussants
- Increasing the consistency of scoring
- Emphasizing the importance of overall significance and impact, while preserving the evaluation of technical merit
- Expanding the potential reviewer base by using mail reviewers
- Simplifying management of dyads and conflicts.

CSR piloted the Two-Stage review format on several types of funding mechanisms, including small business innovation research (SBIR) grants, transformative R01 (T-R01) grants, Challenge grants, and Bioengineering Research Partnership (BRP) grants. In February-May of 2009, CSR solicited feedback from Stage 1 and Stage 2 reviewers, Program Officers (POs), and Scientific Review Officers (SROs) who participated in the pilot, to gauge their level of satisfaction with the prebuttal process. The surveys explored several aspects of the process, including the rigor and quality of review, the level of burden on reviewers, the adequacy of instruction about the review process, the utility of prebuttal, and the overall participant satisfaction. This report is a summary of survey findings.

Methodology

A total of eight program officers (POs), 76 Stage 1 reviewers, and 20 Stage 2 reviewers responded to the surveys, resulting in the response rates of 12.5%, 65%, and 53%, respectively. Survey protocols contained a combination of multiple-choice and open-ended questions. Respondent feedback is reported by topic across groups.

[1] A. Kopstein. Center for Scientific Review. Presentation on June 8, 2009.

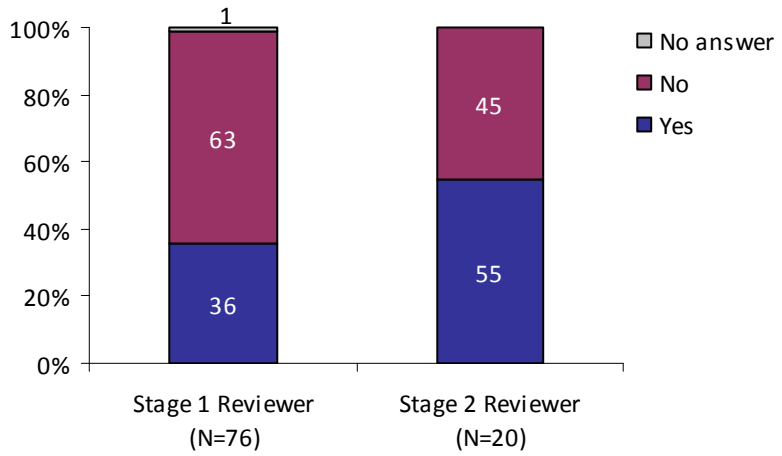
Findings

Experience with the Two-Stage Peer Review Format

Stage 2 reviewers appeared to be more familiar with the Editorial Board Review format than Stage 1 reviewers: more than half of Stage 2 respondents (55%) indicated that they have participated in a Two-Stage review process; for Stage 1 reviewers, this number was 36% (Figure 1). Among the Program Officers, half reported being "very knowledgeable" or "somewhat knowledgeable" about Two-Stage review (data not shown). Of the reviewers who had used this format in the past, the majority ranked their most recent experience as comparable to or better than their previous experience (data not shown).

Figure 1: Previous Participation

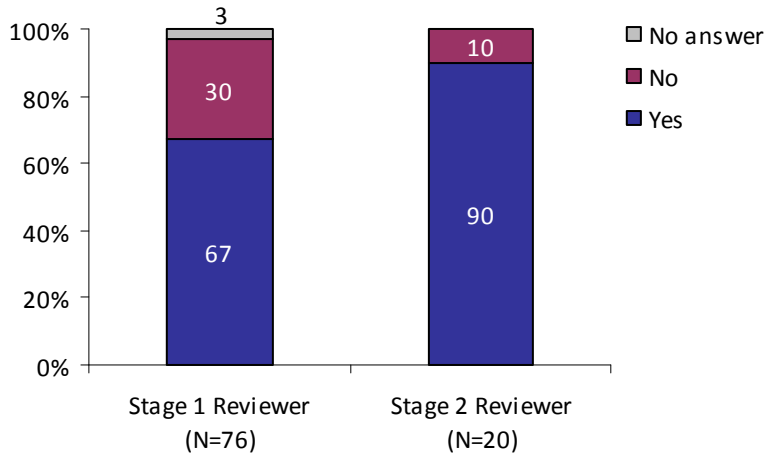
Have you done Two-Stage peer review before your most recent peer review experience?



Survey data also revealed that most survey respondents were experienced with face-to-face peer review: 67% of Stage 1 and 90% of Stage 2 reviewers indicated that they have participated in at least one in-person study section (Figure 2). Seventy-five percent of POs participated in six or more and 25% in 20 or more face-to-face study sections (data not shown). Familiarity with traditional study section format is important, as it allowed respondents to evaluate their Two-Stage review experience in the context of positive and negative aspects of a traditional review.

Figure 2: Familiarity with Traditional Peer Review

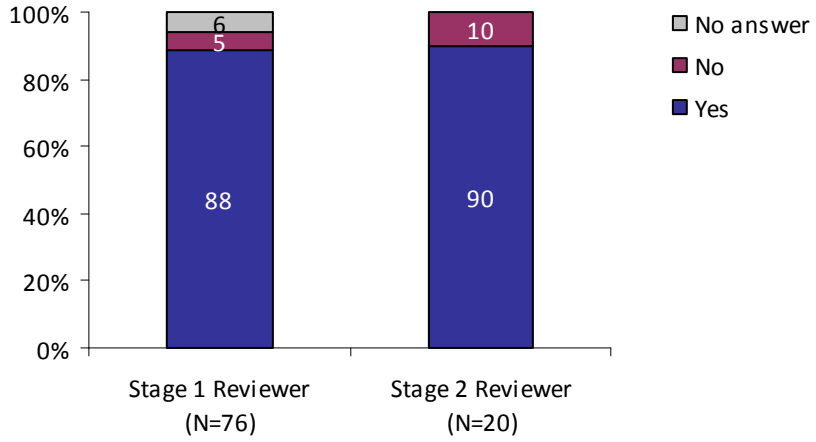
Have you previously participated in at least one traditional face-to-face review study section?



Most survey respondents appeared to be sufficiently prepared for participation in the Two-Stage pilot: 88% of Stage 1 reviewers and 90% of Stage 2 reviewers reported that the information they were provided about the review process was sufficient (Figure 3).

Figure 3: Level of Preparedness for Two-Stage Review

Did you receive sufficient information for participating in this pilot and being a reviewer?

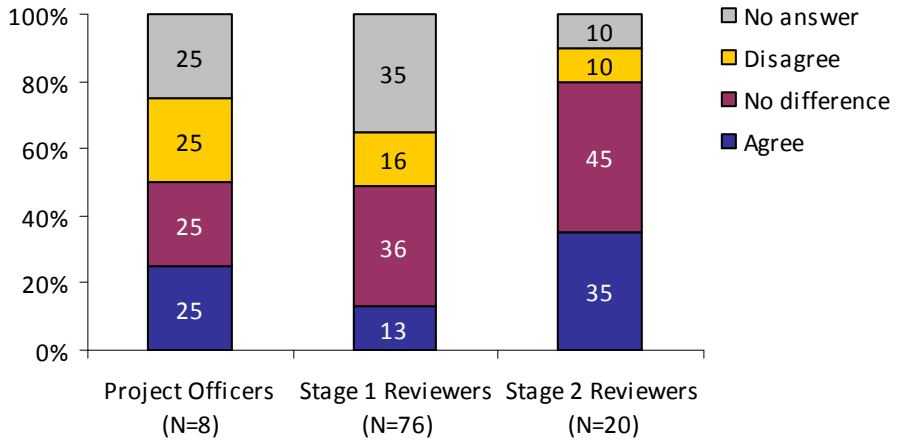


Review Quality and Rigor

Rigorous review process is vital to the identification and funding of best science. Therefore, several survey questions examined participants' satisfaction with the quality and rigor of the review. In one of the questions, survey subjects were asked how the level of emphasis on all five peer review criteria[2] in the pilot Two-Stage review compared to the traditional study section format. Of those who answered the question, more than half indicated the same or greater emphasis on all the criteria for the Two-Stage review (Figure 4).

Figure 4: Emphasis on Peer Review Criteria

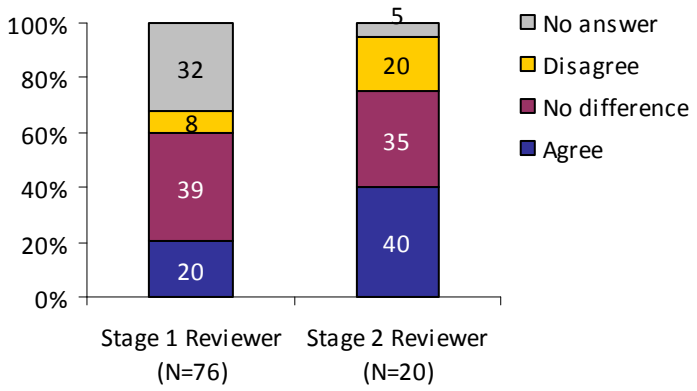
Compared to traditional face-to-face review, I feel this review format put more emphasis on all five peer review criteria.



Using mail reviews as the first step in a Two-Stage process was expected to enable the recruitment of larger number of better matched technical experts per application. The survey examined whether assignment of more reviewers resulted in a more effective review. The data revealed that 20% of Stage 1 and 40% of Stage 2 reviewers felt that the Two-Stage format was more effective in evaluating technical and scientific merit than a traditional study section; an additional 39% of Stage 1 and 35% of Stage 2 reviewers indicated that there was no difference in the two formats (Figure 5).

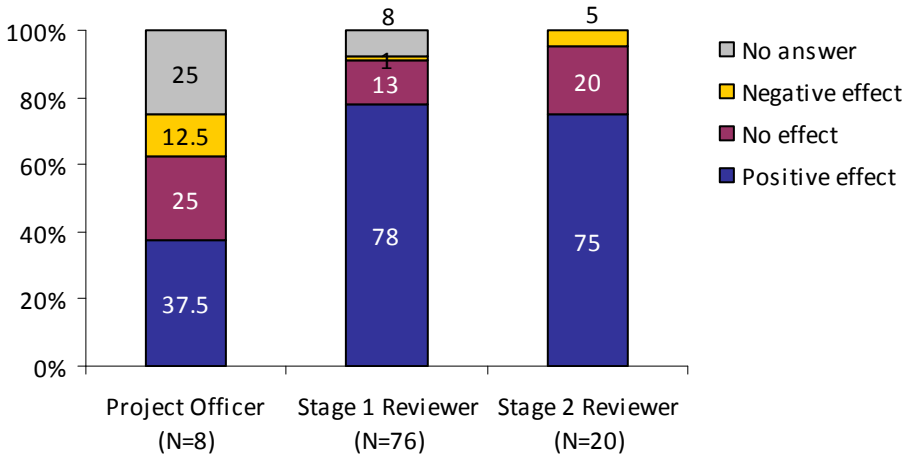
[2] Significance, approach, innovation, investigators, and environment

Figure 5: Effectiveness of Two-Stage Format in Evaluating Scientific and Technical Merit
Compared to traditional face-to face review, I found this review format to be more effective for identifying scientific and technical merit



Additionally, reviewers participating in both stages of the Editorial Board process reported that greater number of assigned reviewers per application resulted in access to more technical expertise (Figure 6). Finally, 85% of Stage 2 reviewers were satisfied with or neutral to the critiques received from Stage 1 reviewers (data not shown).

Figure 6: Effect of Greater Number of Assigner Reviewers on Technical Expertise
What effect do you feel having more assigned reviewers per application had on providing necessary scientific and technical expertise?



Appropriateness of Two-Stage Review Mechanism

While generally satisfied with the Two-Stage format, respondents indicated that it was not suitable for all funding mechanisms. SBIR grant reviews were seen as most amenable to this format (Figure 7). We do not know how many respondents had direct experience using the Two-Stage format for evaluating various types of grants, and thus cannot tell whether having a direct experience correlated with the preference for that mechanism.

Figure 7: Appropriateness of Two-Stage Review Mechanism

Percent respondents, who considered each mechanism appropriate

Grant Mechanism	Program Officers, % respondents	Stage 1 Reviewers, % respondents
R01	25	30
R21	25	30
SRBI Phase I	37.5	65
SRBI Phase II	59	54
T	12.5	13
K	25	18
P	12.5	18

Data were not collected from Stage 2 Reviewers.

Program Officers continued to prefer face-to-face review over all other mechanisms piloted by CSR: 60% of respondents in this group indicated an in-person review as their 1st or 2nd choice for their own applications (data not shown). Two-Stage Review, video- and teleconference were indicated as preferences by 40% of respondents and Internet Assisted Meetings (IAM) by 33% of respondents. None of the respondents reported mail reviews as their top two choices (data not shown).

Review Burden

Survey subjects were asked how many applications they could handle in their role as Stage 1 or Stage 2 reviewers. The majority of Stage 1 reviewers indicated that 3-4 applications were an appropriate number to review, whereas for Stage 2 reviewers this number was higher, between 6 and 13 applications (Figure 8). Several possible reasons may have contributed to this difference. First, evaluating applications for technical and scientific merit – the charge of Stage 1 reviewers – may be more time consuming than evaluating for significance and impact, the primary goal of Stage 2 reviews. Furthermore, Stage 2 reviewers were provided with input from Stage 1 reviewers, which should have facilitated their evaluation of applications, reducing the time necessary to complete the review. Finally, in contrast to Stage 1 reviewers who evaluated the applications on their own, Stage 2 reviewers met in person. It is possible that a communal discussion expedited the review process, by having researchers at hand who had the knowledge to respond to the questions that arose about the application. Thus, the differences in the functions for Stage 1 and 2 reviewers may have led to greater burden for Stage 1

reviewers and, consequently, to the preference by Stage 1 reviewers for a smaller number of applications.

Figure 8: Review Burden

How many applications can you handle in your role as Stage 1 or Stage 2 reviewer?

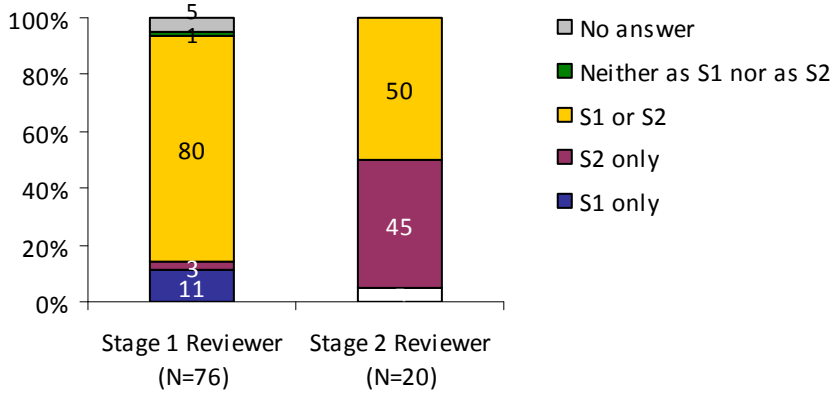
Stage 1, number of applications	Stage 1, % respondents	Stage 2, number of applications	Stage 2, % respondents
1	3	1-5	15
2	13	6-9	35
3	26	10-13	45
4	29	14+	0
5	12	0	0
6	7	0	0
7	0	0	0
8	1	0	0

Satisfaction with 2-Stage Review Process

A clear sign of participant satisfaction with any experience is their willingness to repeat it. The survey revealed that 99% of Stage 2 reviewers and 100% of Stage 1 reviewers would be willing to participate in the Two-Stage review process again (Figure 9). While some reviewers would prefer to participate in the same role only (11% of Stage 1 and 45% of Stage 2) or in the opposite role only (3% of Stage 1 and 5% of Stage 2), half or more were prepared to participate as either Stage 1 or Stage 2 reviewers (80% of Stage 1 and 50% of Stage 2, Figure 9).

Figure 9: Willingness to Serve as Two Stage Reviewers in the Future

For future meetings with Two Stage review, are you willing to serve as Stage 1, Stage 2, either Stage 1 or Stage 2, or not at all?



We found that the majority of reviewers were motivated to participate in the review by the desire to see the best science funded (Figure 10, note that most reviewers selected more than one answer option). Additional reasons given were personal scientific

enrichment, professional connections, and hopes of improving their own chances as applicants by gaining the knowledge of the NIH review process (Figure 10).

Figure 10: Reasons for Participation in Grant Review Process

The greatest intrinsic reward obtained as a grant reviewer (top 2 choices):

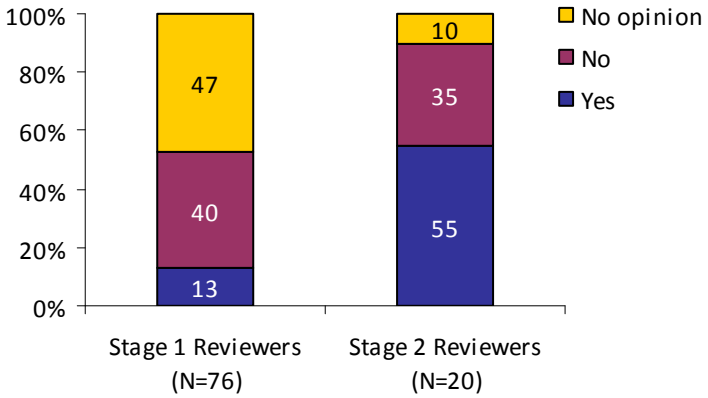
	Stage 1, % Reviewers	Stage 2, % Reviewers
Ensuring that the best approach receives a favorable score	70	60
Staying abreast of the most current research approaches	61	25
Developing networks with other researchers in my field	21	40
Improving my chances of achieving a favorable score	24	28
Better understanding of how a final score is derived	9	26
Benefiting my institution	9	5
Repaying the system	4	0
Serving the country	1	0
Setting an example by trying to be a thorough reviewer	1	0

Suggestions for Improvement

A sizable fraction of reviewers – 40% of Stage 1 and 35% of Stage 2 – indicated that they would make no changes to the Two-Stage review process (Figure 11). The suggestions that were made included: (1) giving Stage 1 reviewers a few days to read each others' critiques and comments, to make suggestions, and to revise their statements (26%); (2) sharing Stage 2 and final scores with Stage 1 reviewers (26%); and (3) providing Stage 1 and 2 reviewers with an opportunity to communicate prior or during the Stage 2 meeting (40%, data not shown).

Figure 11: Suggestions for Change

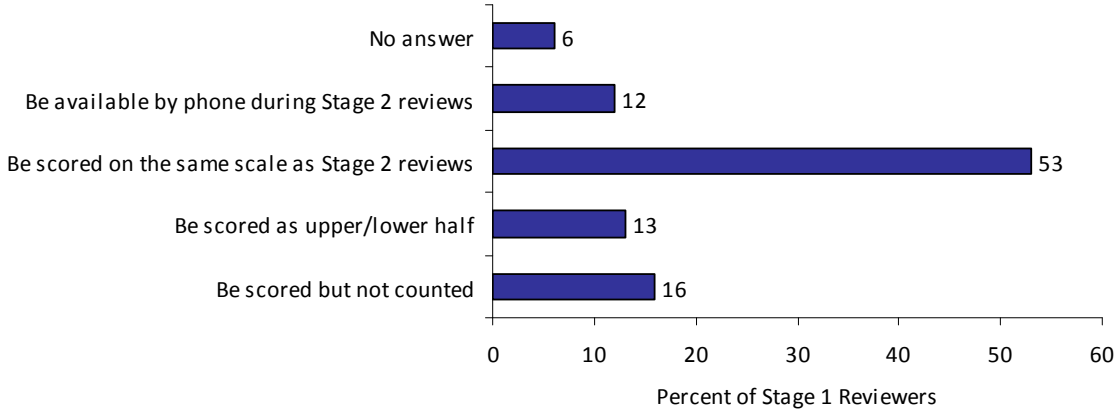
Would you change anything about 2-stage review?



In the pilot, Stage 1 reviewers could indicate their level of enthusiasm about the application and suggest the scores, but ultimately final scoring decisions were made by Stage 2 reviewers, who could disregard Stage 1 reviewer recommendations. Suggestions

made by Stage 1 Reviewers in multiple-choice (Figure 12) and in open-ended questions (summarized below) indicated that this group was somewhat unsatisfied with the role they played in the review process. The most common suggestion by far, made by 53% of Stage 1 reviewers, was that they should be permitted to score applications in the same way as Stage 2 reviewers (Figure 12).

Figure 12: Proposed Roles for Stage 1 Reviewers
Stage 1 Reviews should:



Survey respondents were offered an opportunity to elaborate on their answers by responding to several open-ended questions. On average, two Program Officers (25%), 20 Stage 1 reviewers (26%), and 11 Stage 2 reviewers (55%) submitted comments. Two Stage 2 reviewers used the space provided to communicate their dislike for the process; all other respondents suggested concrete changes or identified specific benefits and limitations of the Two-Stage review model. Concerns and associated recommendations were in two areas: respective roles for Stage 1 and Stage 2 reviewers and implementation of Two-Stage review.

The first area of some concern was related to the roles of Stage 1 and Stage 2 reviewers in the pilot review process. Stage 1 reviewers expressed a disappointment that their input may have had limited impact on the review outcome. This group had no access to the final scores and was not involved in Stage 2 discussion. Stage 1 reviewers made several suggestions on how to enhance their contribution to the review process. One reviewer recommended presenting a synopsis of comments about each application resulting from Stage 1 to the entire Stage 2 panel prior to the discussion of applications. Stage 1 reviewers also suggested that Stage 2 reviewers should provide them with feedback on their critiques, offer them access to the final scores, and engage them in discussions before and/or during Stage 2 review.

Stage 2 reviewers had their own concerns. Some respondents noted that Stage 1 reviewers did not offer sufficient detail and "varied enormously" in their input. Like Stage 1 reviewers, this group saw an advantage in having a discussion between the two groups

at some point during the review. In addition, some Stage 2 reviewers were uneasy about the level of familiarity their Stage 2 colleagues had displayed with the proposals being reviewed. Because Stage 2 reviewers were instructed to focus on the impact and overall significance of applications, some respondents took it to mean that they did not have to be knowledgeable about the technical content of application. One Stage 2 reviewer, for example, noted that some reviewers in his group "said that they don't have to read the grant." He then proceeded to say: "I cannot comprehend how someone can make a sound judgment without it. Based on my review of the grant, I differed significantly from S1 [Stage 1] reviewers on scientific/technical merit." One respondent suggested that in order to ensure that all reviewers are well prepared for the discussion CSR should make it clear that Stage 2 reviewers must be able to "speak to scientific and budgetary issues at the review session."

Respondents also offered suggestions for change in the implementation of Two-Stage review. Stage 1 reviewers would have preferred to have more time to review the applications, greater choice on which applications to review, and more reviewers per grant. Some also mentioned better financial reward. Stage 2 reviewers recommended that CSR provides more information on the Two-Stage process and on the scoring of applications and suggested that Stage 2 meetings are divided into smaller Internet Assisted Meetings, with 20-30 grants per meeting, so that "participation from a larger number of reviewers can be achieved for each application."

Conclusions

Findings from the analysis of survey data revealed that participants in the Two-Stage review were generally satisfied with the format and the outcomes of the process. More than 50% of all respondents indicated that the process resulted in comparable or greater emphasis on all five peer review criteria (Figure 4); 55% of Stage 1 and 75% of Stage 2 reviewers reported that Two-Stage review was as effective or more effective in evaluating technical and scientific merit of applications (Figure 5). Finally, 99% of Stage 1 reviewers and 100% of Stage 2 reviewers would be willing to serve in this capacity in the future (Figure 9).

While generally positive about the Two-Stage format, respondents felt that it was not appropriate for all funding mechanisms. Program Officers and Stage 1 reviewers were overwhelmingly in favor of using Two-Stage review for evaluating SBIR applications, but much less enthusiastic about other types of grants (Figure 7). The second area of concern communicated by respondents was related to the relative roles of Stage 1 and Stage 2 reviewers and the influence of Stage 1 reviewers on the review outcomes. The answers indicated that in some cases Stage 1 reviewers differed significantly in their views on a proposal and/or the views of Stage 2 reviewers differed significantly from those of Stage 1. Since the Two-Stage pilot did not include any communication between the two groups, these differences remained unresolved. Stage 1 reviewers would have preferred to score applications or at least to be informed on the final scores assigned during the second stage.

In summary, the majority of participants liked the Two-Stage process. Furthermore, the survey yielded evidence that several expected enhancements relative to traditional study section did occur, including greater emphasis on all review criteria and expanded access to technical expertise. These improvements came at no cost in terms of review burden.