Report No. 24-02
June 28, 2024

CSR Report

Reviewer Surveys – Peer Review Using a Hybrid Meeting Format

Office of the Director

Introduction

This report is one of a series of reports stemming from the Center for Scientific Review's ongoing effort to understand the impact of meeting format on the peer review process. The National Institutes of Health (NIH) depends on the Center for Scientific Review's (CSR) peer review process to ensure that all NIH grant applications receive fair, independent, expert, and timely reviews that are free from inappropriate influences. When the COVID-19 pandemic hit in mid-March 2020, CSR shifted its review meetings online, using the Zoom.gov video meeting platform. During Fall 2022, CSR reimplemented face-to-face meetings and held one third of its standing study sections, small business, and fellowship meetings face-to-face—with the remainder of meetings held virtually. This was CSR's first opportunity to hold a substantial amount of its regular review meetings both virtually and face-to-face at the same time. During Summer and Fall 2023 and Winter 2024, CSR held a small number of review meetings in a hybrid format (i.e., a meeting where some reviewers attended in-person and some attended remotely). The purpose of this survey was to assess participants' experiences and perceptions of those meetings, and whether views differ for those who attended hybrids in-person versus those who attended remotely. To facilitate comparison with other CSR reports on the effects of meeting format, data from reviewers at prior face-to-face meetings are included in some figures for comparison.

This is an interim report; as CSR conducts more hybrid meetings and as more data become available updates will be issued.

Key findings

The data show that hybrid reviewers perceived hybrid meetings as highly effective; over 90% of reviewers, regardless of whether they attended in-person or remotely, rated discussions, reviewer engagement, and overall meeting quality as excellent or good. Small differences were observed between in-person and remote reviewers' ratings of personal participation and experience in the meetings. Reviewers' preferred meeting format differed substantially depending on whether they attended in-person or remotely.

- 94% of reviewers thought that their review panels were able to prioritize applications according to their impact and scientific merit and that the scientific discussions helped the panel evaluate the applications. There were no significant differences in these ratings among reviewers who attended in-person and remotely.
- 97% of reviewers rated the overall quality of review at their meetings as excellent or good. There were no significant differences in ratings of review quality (across all measures, including engagement) between reviewers who attended in-person and those who attended remotely.
- Reviewers reported high levels of self-participation and a positive experience at their review meetings. Ratings of participation and engagement were significantly higher from reviewers who attended in-person vs. remotely, although the magnitude of these effects are small.
- Reviewers who attended in-person reported being able to sustain attention longer than did remote reviewers.
- A majority of hybrid reviewers prefer meeting in-person, but results differ greatly according to how they attended. In-person reviewers favor face-to-face meetings by 84% to 4%, while remote reviewers favor virtual meetings by 49% to 23%.

See *Appendix* for Methods

Results

The survey was administered to 7,577 reviewers, of which 3,605 completed the survey for a response rate of 48%. Of the respondents, 52% attended virtual review meetings (n = 1,864), 33% attended face-to-face review meetings (n = 1,204), and 15% attended hybrid meetings (n = 537)[1]. Among those who attended hybrid meetings, 46% (n = 246) attended remotely and 54% (n = 291) attended in-person. See Table 1 for reviewer characteristics. The analyses that follow are confined to hybrid meetings only. However, to add context, data from CSR's Summer 2023 survey of face-to-face meetings are displayed in many of the figures.

Table 1. Characteristics of Survey Respondents

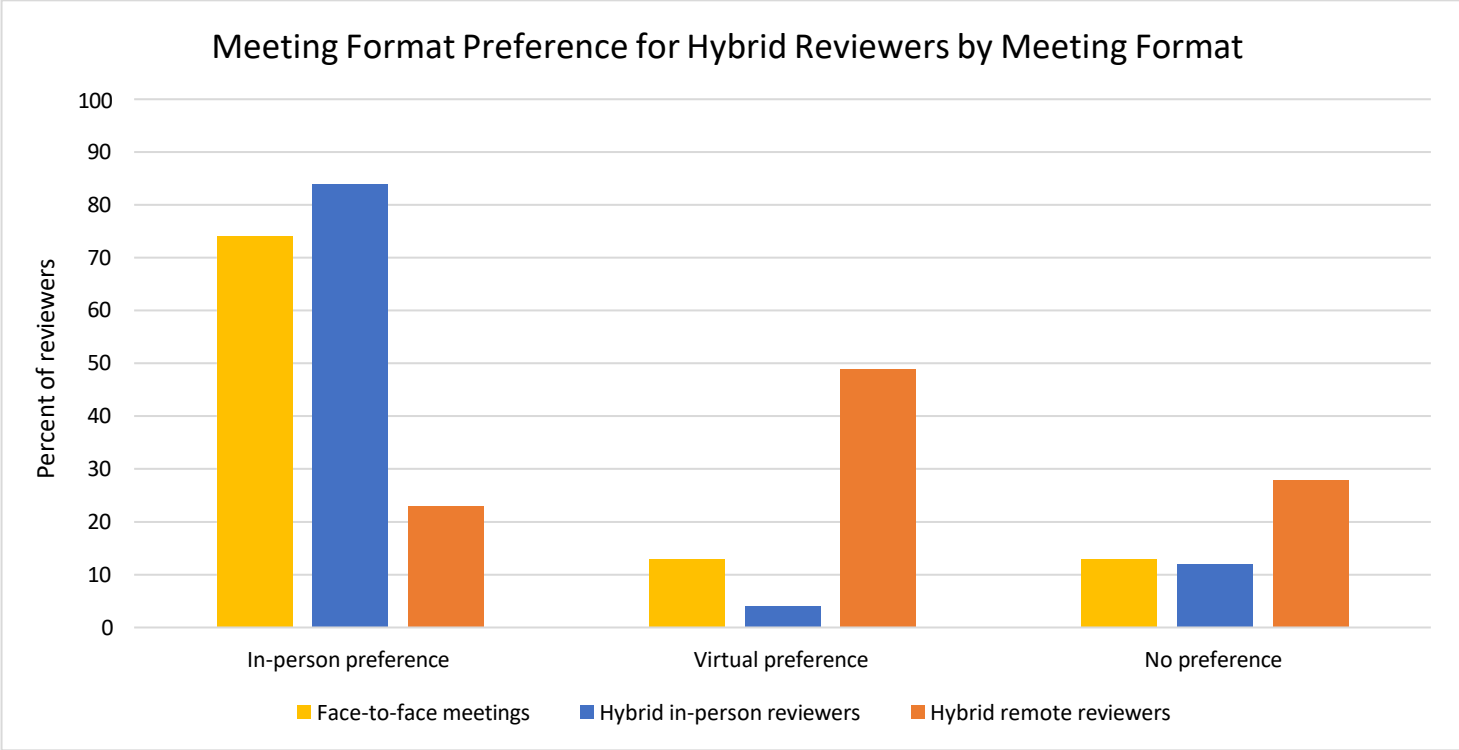
Reviewer Characteristics		% Survey Respondents (n = 3,605) * includes hybrids	% Hybrid Survey Respondents (n = 537)
Gender			
	Male	52	51
	Female	45	46
	Withheld	3	3
Race			
	American Indian or Alaskan	< 1	< 1
	Asian	22	21
	Black or African American	4	4
	More than one race	3	3
	Native Hawaiian or Pacific Islander	< 1	0
	White	62	63
	Withheld	9	9
Ethnicity			
	Hispanic/Latino	9	11
	Non-Hispanic	86	84
	Withheld	5	5
URM			
	No	78	71
	Yes	14	21
	Withheld	8	8
Career Stage			
	Professor	45	47
	Associate Professor	33	27
	Assistant Professor	17	18
	Other	5	8

[1] To increase sample size, participants who attended hybrid meetings came from three different rounds of review meetings—155 participants attended the Summer 2023 meetings, 95 participants attended the Fall 2023 meetings, and 287 participants attended the Winter 2024 meetings.

Meeting Format Preferences of Reviewers from Hybrid Meeting

- Figure 1 shows the preferences of reviewers from hybrid meetings[2]. Preferences of those reviewers attending face-to-face review meetings (from prior surveys) are also shown for comparison.
- Overall, hybrid reviewers prefer face-to-face meetings, but the degree of preference varies based on the format of the meeting they attended. In-person reviewers strongly prefer face-to-face meetings, whereas those who attended remotely favored virtual meetings overall, with many declaring no preference.
- In-person reviewers were significantly more likely than remote reviewers to prefer face-to-face meetings and remote reviewers were significantly more likely to prefer virtual meetings than were in-person reviewers.

Figure 1. Hybrid Reviewers Meeting Format Preference by Meeting Format



There was a significant association between the format of reviewers' meetings and their meeting format preferences ($\chi 2$ (2) = 212.830, p <.001), with a moderate to large effect size (ϕ_c = .630, p <.001). Post hoc z-tests show that the proportion of hybrid in-person reviewers who preferred face-to-face meetings was significantly more than the proportion of hybrid remote reviewers who preferred face-to-face meetings and 2) the proportion of hybrid in-person reviewers who preferred virtual meetings or had no meeting preference was significantly less than the proportion of hybrid remote reviewers who preferred virtual meetings or had no meeting preference.

[2] Scientific Review Officers (SROs) volunteered to hold hybrid meetings, taking into consideration the preferences of their reviewers on whether they wanted to attend in-person or remotely. Thus, these findings could reflect those reviewers who, a priori, preferred to attend in-person or remotely.

Quality of Review for Hybrid Meetings for in-person vs. remote reviewers

- Figure 2 shows reviewers' perceptions of the quality of the review meeting by meeting format.
- 97% of all reviewers thought the overall quality of the meetings were good or excellent.
- There were no significant differences in ratings of review quality (across all measures) between reviewers who attended in-person and those who attended remotely (see Table 1).
- In comparison to reviewers at face-to-face meetings, in-person hybrid reviewers rated, on average, the quality of the review (across all measures) lower (see Table 1).

Figure 2. Reviewer Assessment of Peer Review Quality

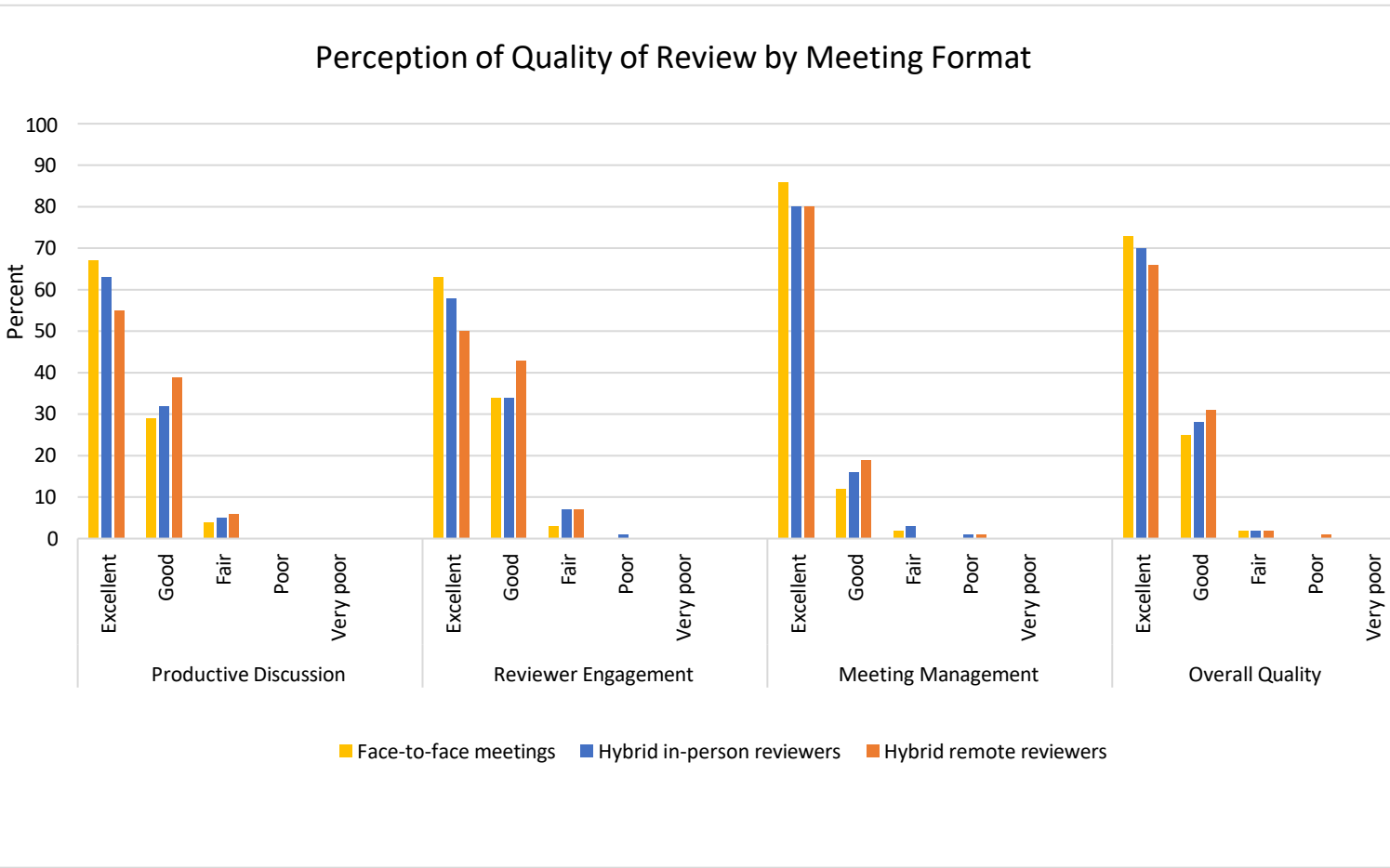


Table 1. Reviewer Perception of the Quality of the Review by Meeting Format

Reviewer Ratings of Review Quality by Meeting Format				*For comparison between hybrid in-person and face-to-face meetings*
	Hybrid Remote Reviewers (*M, SE*)	Hybrid In-person Reviewers (*M, SE*)	Independent t-test Statistic Between Hybrid Reviewers	Reviewers in face-to-face meetings (*M, SE*)
Productive Discussion	4.48 (.04)	4.58 (.04)	$t(535) = -1.927$, $p = .055$	4.61 (.02)
Reviewer Engagement	4.41 (.04)	4.48 (.04)	$t(535) = -1.155$, $p = .249$	4.59 (.02)
Meeting Management	4.78 (.03)	4.75 (.03)	$t(535) = .581$, $p = .562$	4.84 (.01)
Overall Quality of Review	4.62 (.45)	4.68 (.03)	$t(495.423) = -1.329$, $p = .189$	4.70 (.02)

Meeting Experience and Participation for Hybrid Meetings

- Figure 3 shows reviewer perceptions of their meeting experience and participation by meeting format. Data for those who attended face-to-face meetings (i.e., held entirely in-person) is included as a point of reference.
- Around 90% of all reviewers always or often felt confident voicing their opinions, thought others were receptive and responsive to their feedback, and were able to clearly communicate their opinions.
- Reviewers' contributions to the discussion and comfort voting outside the range were the largest reported difference between those who attended in-person and those who attended remotely—71% of reviewers who attended in-person and 46% of reviewers who attended remotely reported that they always or often contributed to the discussion; 84% of reviewers who attended in-person and 73% of reviewers who attended remotely reported that they always or often felt comfortable voting outside the range.
- There were significant differences among reviewers across all measures of meeting experience and participation, with those who attended in-person rating their experience and participation higher (see Table 2). The magnitude of these effects is small.
- In comparison to face-to-face meetings, in-person hybrid reviewers rated, on average, their meeting experience and participation higher (see Table 2).

Figure 3. Reviewer Experience and Participation by Meeting Format

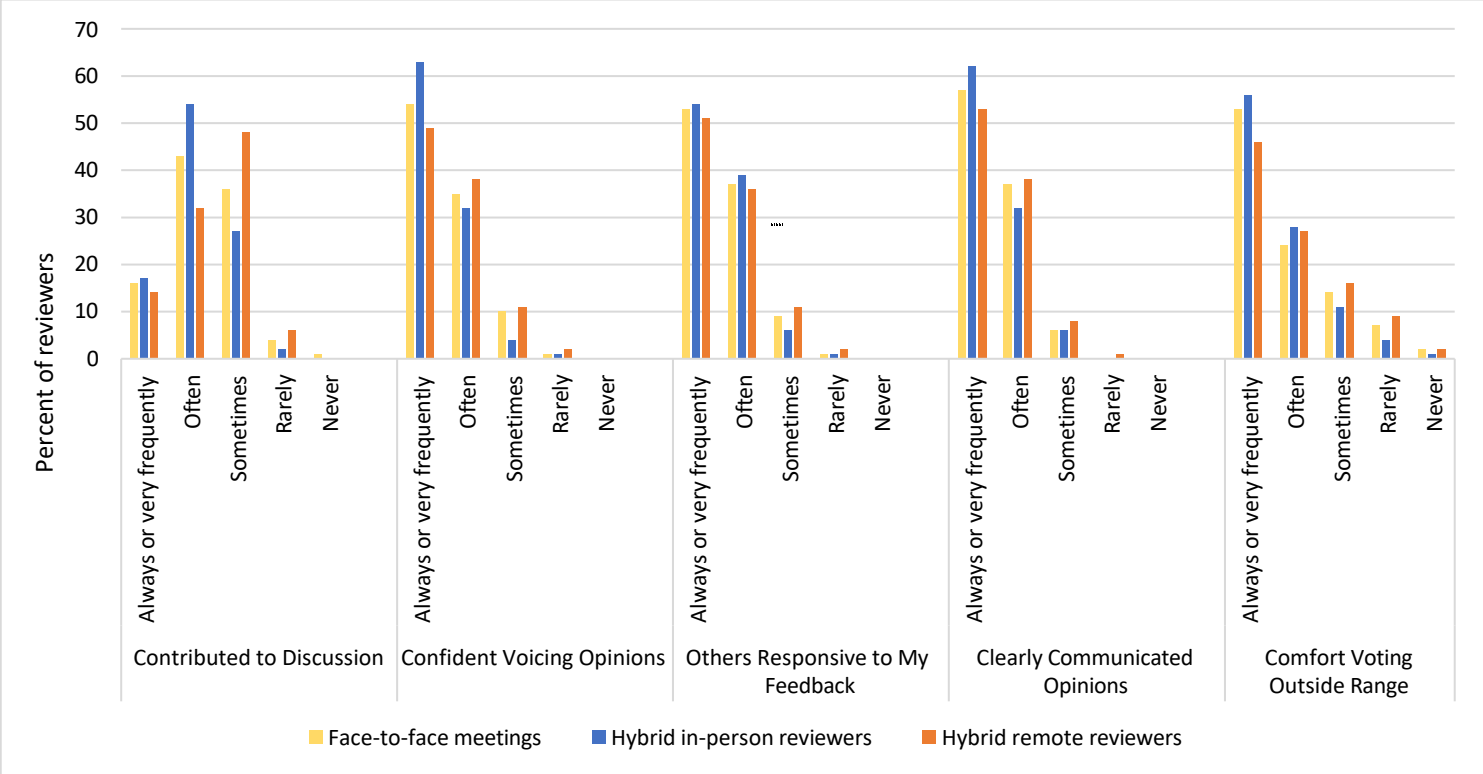


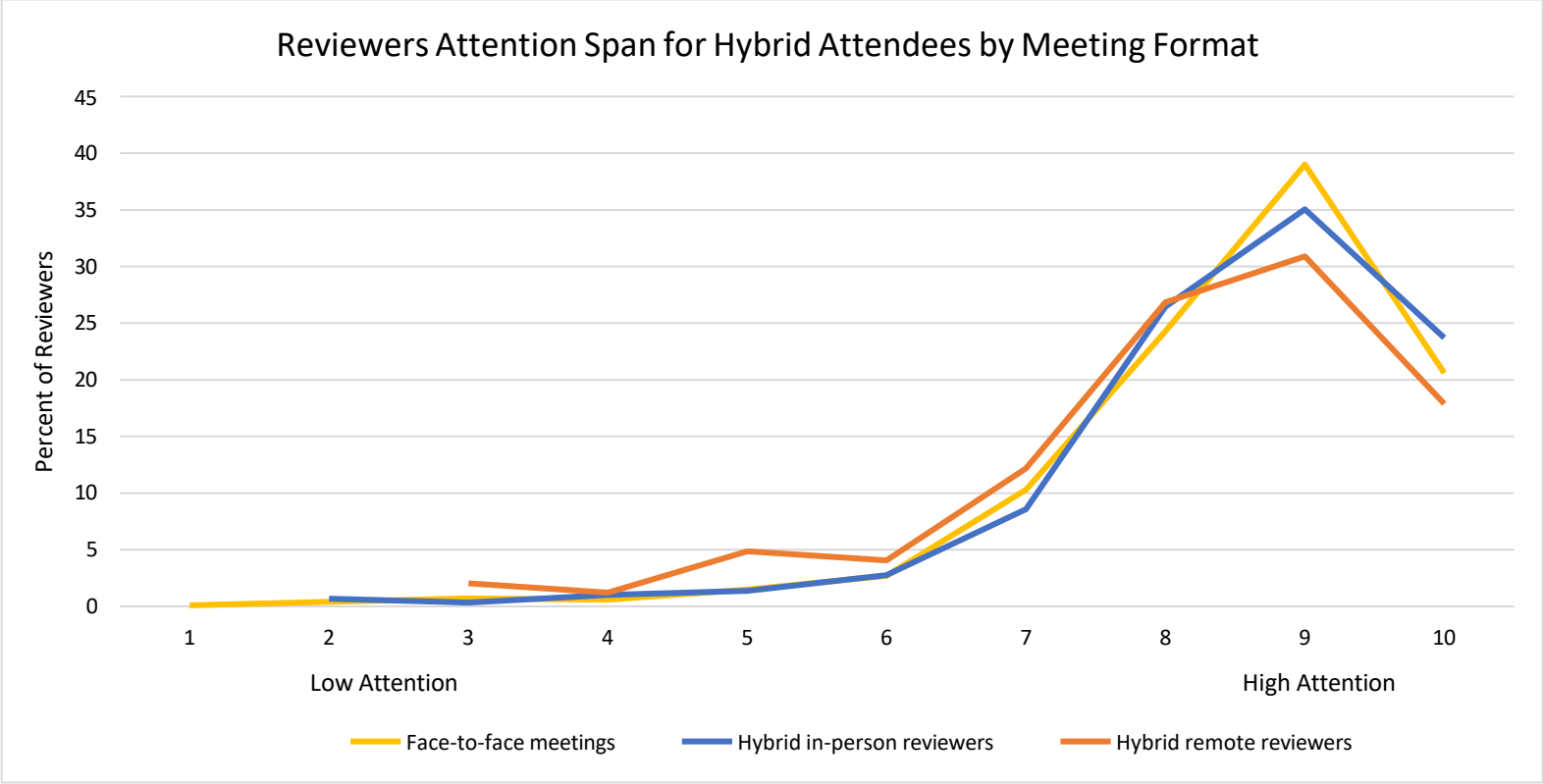
Table 2. Reviewer Experience and Participation by Meeting Format

Experience and Participation of Reviewers Attending Hybrid Meetings				*For comparison between hybrid in-person and face-to-face meetings*
	Hybrid Remote Reviewers (*M, SE*)	Hybrid In-person Reviewers (*M, SE*)	Independent t-test Statistic and *r* Between Hybrid Reviewers	Reviewers in face-to-face meetings (*M, SE*)
Contributed to Discussion	3.54 (.05)	3.86 (.04)	$t(492.50) = -4.778$, $p < .001$; $d = .42$	3.70 (.80)
Confident Voicing Opinions	4.34 (.05)	4.56 (.04)	$t(448.41) = -3.611$, $p < .001$; $d = .32$	4.41 (.73)
Others Receptive and Responsive to Feedback	4.35 (.05)	4.47 (.04)	$t(468.63) = -1.928$, $p = .054$; $d = .17$	4.43 (.69)
Clearly Communicated Opinions	4.42 (.04)	4.55 (.04)	$t(498.90) = -2.175$, $p = .030$; $d = .19$	4.50 (.62)
Comfortable Voting Outside Range	4.05 (.07)	4.35 (.05)	$t(397.57) = -3.261$, $p = .001$; $d = .31$	4.21(1.03)

Reviewer Attention Span in Hybrid Meetings

- Figure 4 shows hybrid reviewers' report of their ability to pay attention throughout the meeting, by meeting format.
- Around 3% of remote reviewers and 2% of in-person reviewers had difficulty concentrating at the meeting (i.e., less than 5 on attention scale).
- Around 76% of remote reviewers and 85% of in-person reviewers were able to sustain their attention throughout the meeting (i.e., 8 or more on the attention scale).
- There was a significant difference in reviewers' ability to pay attention at the meeting, with reviewers who attended in-person paying more attention or having longer attention spans than those who attended remotely. The magnitude of this effect is small.
- In comparison to face-to-face meetings, in-person hybrid reviewers rated, on average, their attention span about the same (M = 8.54 for hybrid in-person, *M* = 8.51 for face-to-face meetings).

Figure 4. Hybrid Reviewers Attention Span at the Meeting

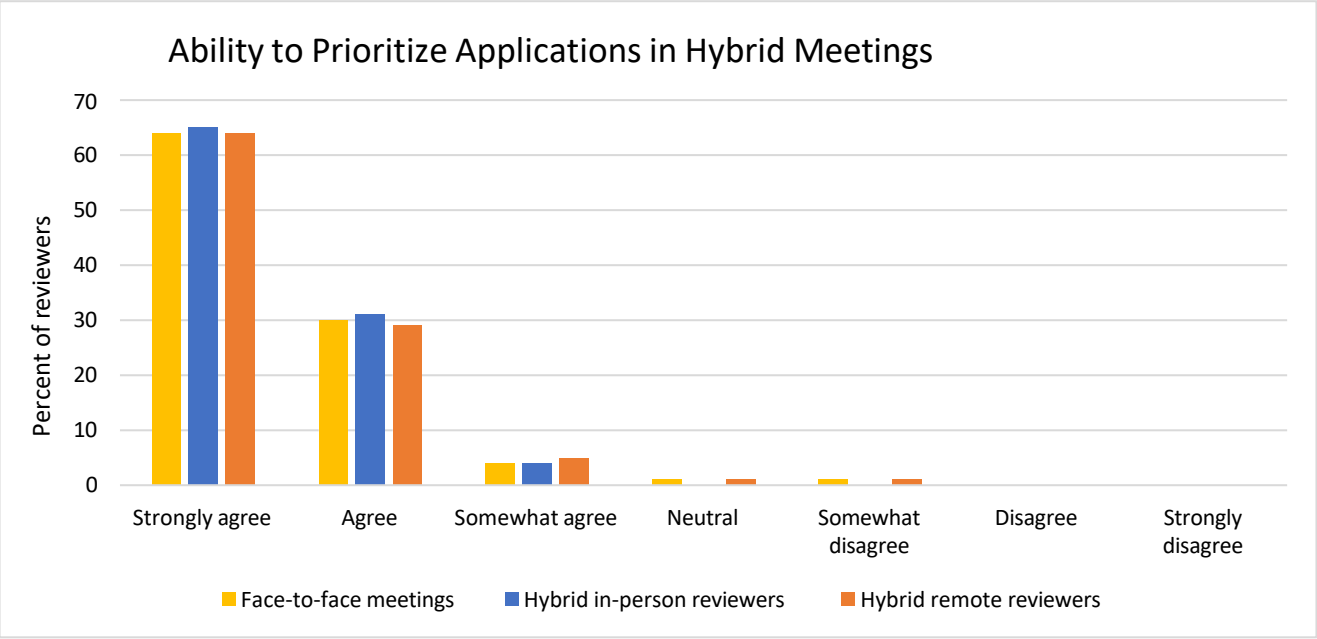


There was a significant difference between reviewers who attended the meeting in-person (*M* = 8.54, *SE* = .08) and those who attended remotely (*M* = 8.17, *SE* = .10) in their attention span or ability to concentrate at the meeting. This difference, -.376, *CI* [-.623, -.129], was significant *t*(535) = -2.994, *p* = .003; and represented a small-sized effect, *d* = .26.

Prioritizing Applications in Hybrid Meetings

- Figure 5 shows data capturing hybrid reviewers' perceptions of their panel's ability to prioritize applications, by meeting format.
- 94% of all reviewers believed that the panel was able to prioritize applications according to their impact and scientific merit.
- There was no significant difference in reviewers' ability to prioritize applications between those who attended remotely and those who attended in-person.
- In comparison to face-to-face meetings, in-person hybrid reviewers rated, on average, their ability to prioritize applications about the same (*M* = 1.43 for hybrid in-person and *M* = 1.45 for face-to-face meetings; see note below graph).

Figure 5. Hybrid Reviewers' Assessments of their Panel's Ability to Prioritize Applications



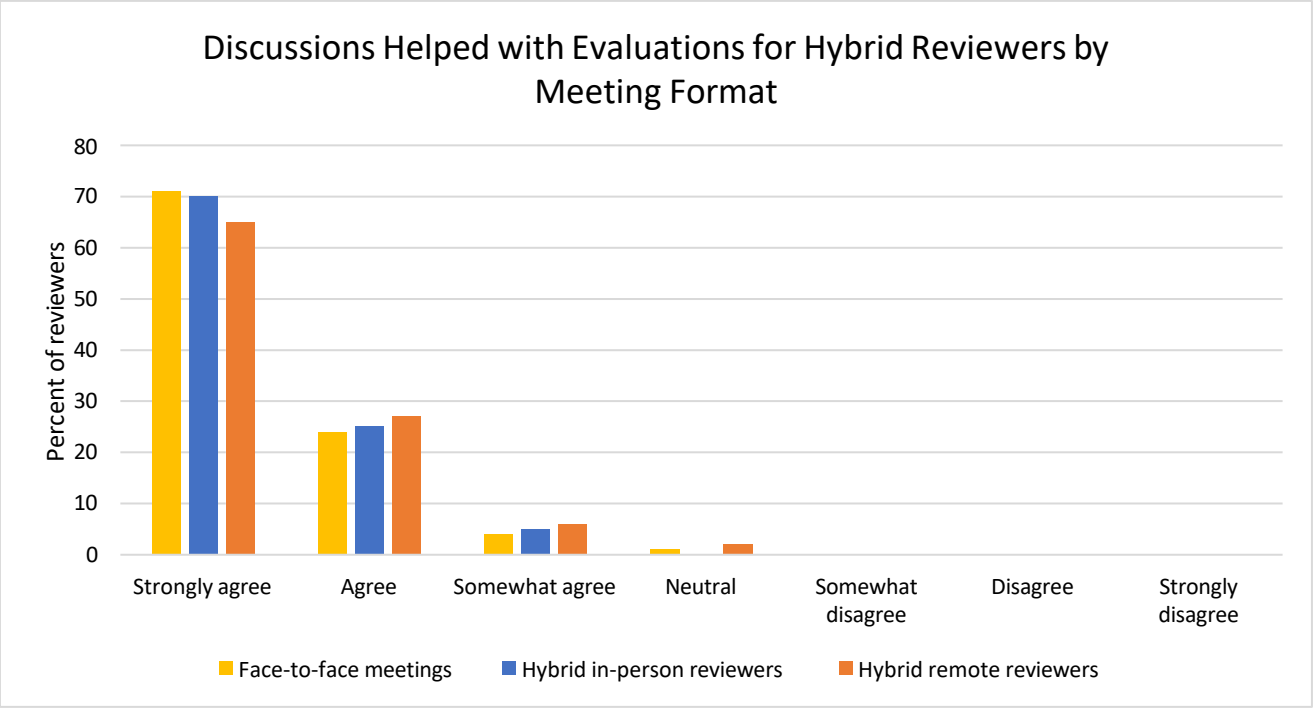
Note: Survey questions were on a scale from 1 (strongly agree) to 7 (strongly disagree), with lower scores indicating a more favorable view of the review meeting.

There was no significant difference in ability to prioritize applications between those who attended the meeting in-person (*M* = 1.43, *SE* = .04) and those who attended remotely (*M* = 1.46, *SE* = .05) (.022, *CI* [-.099, .144]; *t*(535) = .361, *p* = .718).

Perceptions of Discussion Quality for Reviewers in Hybrid Meetings

- Figure 6 shows data capturing hybrid reviewers' perceptions of their panel's ability to discuss and evaluate applications by meeting format.
- 94% of all reviewers believed that the scientific discussions helped the panel evaluate the applications being reviewed.
- There was no significant difference between reviewers who attended in-person and those who attended remotely in their opinions on whether the discussions helped to evaluate the applications.
- In comparison to traditional face-to-face meetings, in-person hybrid reviewers rated, on average, the discussions help with the evaluations about the same (*M* = 1.38 for hybrid in-person and *M* = 1.37 for face-to-face meetings; see note below graph).

Figure 6. Hybrid Reviewers' Assessments of the Utility of Discussion for Evaluations



Note: Survey questions were on a scale from 1 (strongly agree) to 7 (strongly disagree), with lower scores indicating a more favorable view of the review meeting

There was no significant difference between reviewers who attended the meeting in-person (*M* = 1.38, *SE* = .04) and those who attended remotely (*M* = 1.43, *SE* = .04) in their perceptions on whether the discussions helped the panel evaluate the applications (.054, *CI* [-.063, .170]; *t*(535) = .902, *p* = .368).

Appendix

Methods

Participants

Surveys were completed by reviewers who participated in 241 CSR study section meetings (n = 3,605) between May 31, 2023 to March 28th, 2024. The study section meetings included chartered panels, recurring small business and fellowship special emphasis panels (SEPs), and a few additional SEPs that held hybrid meetings. Reviewers who attended hybrid meetings came from 34 study sections.

Survey Administration

Reviewers were asked for their participation in a survey via email on the last day of the study section meeting. The email contained a weblink to the survey. Reviewers were told in the email that their responses would be kept confidential and that the survey would take less than five minutes to complete. All surveys returned by April 20th, 2024 were included for analysis.

Measures

Application Evaluation

Two items asked participants to rate on a scale from 1 (strongly agree) to 7 (strongly disagree) the panel's ability to evaluate the applications: 1) the panel was able to prioritize applications according to their impact and scientific merit, and 2) the scientific discussion helped the panel evaluate the applications being reviewed.

Peer Review Quality

Four items asked participants to rate on a scale from 1 (very poor) to 5 (excellent) the following items: 1) overall quality of review, 2) productivity of discussions, 3) level of reviewer engagement, and 4) meeting management.

Reviewer Meeting Experience and Participation

Five items asked participants to rate on a scale from 1 (never) to 5 (always or very frequently) the following items: 1) I contributed to the discussion, 2) I felt confident voicing my opinions, 3) I felt others were receptive and responsive to my feedback, 4) I was able to clearly communicate opinions, and 5) I felt comfortable voting outside the range.

Attention Span

One item asked participants to rate their attention span at the review meeting using a scale from 1-10, with 1 being "really struggled to concentrate" and 10 being "no problem concentrating at all".

Format Preference

One question asked participants if there were no or minimal health risks from COVID-19, would they be more likely to participate in a review meeting if it was held face to face or over Zoom/video? Response options included: face-to face, Zoom/video, and no preference.

Demographic Information

Four questions were used to collect the demographic characteristics of respondents. 1) *Gender*: male, female, I prefer not to respond; 2) *Race:* American Indian or Alaskan Native, Asian, Black or African American, Native Hawaiian or Pacific Islander, White, I prefer not to respond; 3) *Ethnicity*: Are you Hispanic? Yes, No, I prefer not to respond; 4) *Career stage*: Assistant Professor, Associate Professor, Professor, Other.

Participants' race and ethnicity were used to determine whether they were an underrepresented minority or not. Non-Hispanic Asians and Non-Hispanic whites were coded as "not URM" and all other participants were coded as "URM". For participants who identified with more than one racial group, if one racial identity was not white or Asian, they were coded as "URM". Participants who identified as both white and Asian were coded as "not URM"[2].

Open-ended Response Options

In an open-ended text box, participants were asked to please share any comments (positive or negative) about their experience or general thoughts on their recent review meeting.

[2] According to OMB standards, individuals who identify with an Asian racial group other than Chinese, Filipino, Japanese, Korean, Asian Indian or Thai are considered an under-represented minority (URM). However, the current data does not allow for this level of group specificity, and therefore only Asian participants who identify as Asian and another racial group (other than White) or as Asian and Hispanic are coded as URM.