NIH Center for Scientific Review — Impact of Zoom Format on CSR Review Meetings

Executive Summary

This document compiles data from 4 sources: 1) survey data from 3288 reviewers who participated in 276 CSR study section meetings; 2) quantitative data from 143 Zoom meetings; 3) survey data on 230 Zoom meetings from 128 CSR scientific review officers (SRO); 4) survey data from 73 extramural support assistants (ESA). All concerned video review meetings in the August/October 2020 round. Main findings from each follow. The data are not exhaustive and CSR plans to continue evaluation of review conducted in the Zoom platform. Thus far, the data suggest that use of the Zoom virtual format has not negatively impacted review. Two consistent points of concern are difficulties maintaining reviewer engagement and with sustaining attention.

REVIEWER SURVEY FINDINGS

1. Reviewers mostly rated review quality in Zoom meetings the same as face-to-face (F2F), but there are some concerns:
   - 60% of reviewers rated overall quality of review in Zoom meetings the "same" as F2F, and 50% rated the quality of discussion the same as F2F.
   - 36% rated discussions slightly or much worse and 51% rated reviewer engagement as worse.
2. Reviewers generally felt as able to interact in Zoom meetings as they had in F2F: The percent who rated Zoom and F2F the same on the items "confident voicing opinions", "others responsive to my feedback", "comfort voting outside range" was 76%, 74%, and 79%, respectively.
3. Nearly half (46%) said their attention span was shorter in Zoom meetings.
4. Reviewers overall preferred the face to face (F2F) meeting format over Zoom (43% to 31%). However, a plurality of reviewers for SBIR and for member conflict SEPs preferred Zoom.
5. A small minority of reviewers (~10-15%) rated Zoom meetings better than F2F on nearly every measure of review quality and reviewer participation.
6. Ratings did not differ substantially according to demographics or career stage.

QUANTITATIVE ANALYSIS OF REVIEW MEETINGS DONE WITH ZOOM

Meeting duration and roster characteristics from a sample of 143 August/October council Zoom meetings were compared to their last in-person equivalents.

1. For Zoom review meetings, per application discussion time averaged 20 minutes, and total meeting time averaged between 15 and 20 hours for medium and large meetings. No comparable data are available for F2F, but CSR believes both durations are modestly longer than F2F.
2. Significant variation in practices regarding breaks was observed.
3. Reviewer workloads were very slightly reduced in the Zoom round.
4. There was no consistent substantial change in female or minority representation on rosters for meetings held by Zoom vs F2F.
5. Distributions of reviewers according to career stage did not change substantially with the introduction of Zoom meetings.
6. The proportion of new or nearly new reviewers did not change in the Zoom meetings.

SRO SURVEY FINDINGS

1. SROs overall preferred F2F over Zoom 44% to 36%, and preferred F2F for chartered panels even more strongly, 52% to 26%.
2. SROs perceived the quality of review in Zoom meetings generally the "same": 85% for Overall Quality, 81% for Quality of Discussion, 67% for Reviewer Engagement.
3. 70-90% of SROs said meeting format did not change demographic diversity of study sections. At the same time, SROs reported it was slightly or much easier to recruit women for about a third of meetings.
4. SROs reported roughly 60% of meetings ran longer; virtually none reported shorter meetings.

ESA SURVEY FINDINGS

1. Overwhelmingly, ESAs thought support tasks were the same or easier to manage in Zoom.
    a. 54% said last minute changes were easier, 29% said it was the same
    b. 35% said handling COIs was easier, 62% said it was the same as F2F
    c. 46% said technical troubleshooting was easier, 31% said it was the same
    d. 48% said solving problems in general was easier
2. A majority said Zoom took no extra support time or less than one hour additional and only ~15% said it took 3 or more additional hours
3. 63% reported that solving technical problems required less than 30 min
4. 79% of respondents had prior experience supporting video-assisted review meetings.

NEXT STEPS

Analyses of the review process when conducted using Zoom are ongoing. CSR plans to administer surveys to stakeholders during the February and March 2021 review meetings. Additional quantitative analyses are planned, particularly with regards to scoring behavior.

I.   Reviewer Survey Findings

II.   Quantitative Analyses of CSR's First All-Virtual Review Cycle Using Zoom

III.   Scientific Review Officer Survey Findings

IV.   Extramural Support Assistants Survey Findings

V.   Appendices – Surveys Administered

Reviewer Surveys - Experiences with Review Using Zoom

Introduction

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 most review meetings online, using the Zoom.gov video meeting platform. The purpose of this survey is to assess the reviewers' observations on the quality of the review and their experiences as reviewers, compared to the normal face-to-face review meetings.

Methods

Participants

Reviewers who participated in 276 CSR study section meetings (n = 3403; n for those using Zoom = 3288) between May 27th to August 4th, 2020. The study section meetings included chartered panels, recurring special emphasis panels (SEP) such as small business and fellowship reviews, as well as member conflict SEPs.

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 August 6th were included for analysis. A copy of the survey is appended.

Measures

Peer Review Quality

Four survey items asked participants to rate on a scale from 1 (much better) to 5 (much worse) how their Zoom review meeting compared to their normal face-to-face review meetings for the following items: 1) overall quality of review, 2) productivity of discussions, 3) level of reviewer engagement, and 4) meeting management.

Individual Participation and Perceptions

Six survey items asked participants to rate on a scale from 1 (much more) to 5 (much less) how their Zoom review meeting compared to their normal face-to-face review meetings for 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, 5) I felt comfortable voting outside the range, and 6) My attention span at the meeting lasted

Experience with Face to Face Meetings

One survey question asked participants to indicate approximately how many face-to-face (F2F) review meetings they have participated in for NIH. Response options included: 0, 1, 2-5, 6-15, and 16+

<u>Format Preference</u>

One survey 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.

<u>Demographic Information</u>

Four survey 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"[1].

<u>Open-ended Response Options</u>

Reviewers were also given the opportunity to provide any additional comments on overall quality of Zoom review meeting and any comments (positive or negative) about their experience in two text boxes: 1) How was the overall quality of your virtual review meeting compared to your normal face-to-face meetings?, 2) Please share any comments (positive or negative) about your experience or general thoughts on having your review meeting over Zoom?

Results
Respondent characteristics

- The survey was administered to 8,083 reviewers, 3403 (43%) completed the survey, and of those, 3288 had participated in a Zoom meeting. Analysis was confined to those 3288 Zoom meeting respondents.
- See Table 1 for survey responder and non-responder demographics; the data source for respondents is the survey and for non-respondents, IMPAC II. Note that compared to non-responders, responders include more women and full professors, and fewer reviewers who do not hold traditional academic ladder positions, presumably people from industry.
- Table 2 shows the count of meetings represented according to type and number of survey respondents. The bulk of responses come from chartered study section meetings; member conflict SEPs, SBIR, and fellowship meetings combined account for ~40% of responses.
- Figure 1 contains the distribution of respondents according to prior review service.

---

[1] 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.

Table 1. Reviewer Characteristics of Survey Respondents who used Zoom and Non-respondents

| Reviewer Characteristics		Survey Respondents (n = 3288)	Survey Non-respondents (n = 4680)
Gender			
	Male	59%	62.8%
	Female	39%	36.4%
	Unknown	--	0.8%
	Withheld	2%	<11
Race			
	American Indian or Alaskan	<1%	<11
	Asian	19%	25.2%
	Black or African American	2%	2.7%
	More than one race	<1%	3.8%
	Native Hawaiian or Pacific Islander	<1%	<11
	White	69%	64.8%
	Unknown	--	0.5%
	Withheld	8%	2.8%
Ethnicity			
	Hispanic/Latino	5%	4.9%
	Non-Hispanic	90%	89.8%
	Unknown	--	2.5%
	Withheld	5%	2.8%
URM			
	No	84%	88.6%
	Yes	8%	9.2%
	Withheld	8%	2.3%
Career Stage			
	Professor	56%	46.3%
	Associate Professor	30%	29.5%
	Assistant Professor	9%	8.8%
	Other	5%	15.4%

Survey Respondents

Table 2. Study Section Type and Number of Respondents

Study Section Type	Number of Study Sections	Number of Respondents
Chartered	154	2166
Member Conflict	42	283
Recurring Small Business	36	534
Recurring Fellowships	21	277
Total	253*	3260*

*Only 28 reviewers came from two *Recurring Special Topics* study sections bringing the total number of study sections to 255 and the total number of respondents to 3288. Due to these low numbers, separate analyses for this type of study section were not examined, but these reviewers are included with the main analyses examining the full sample.

Experience with Face to Face Review Meetings at NIH

Figure 1. Number of Face to Face Meetings at NIH
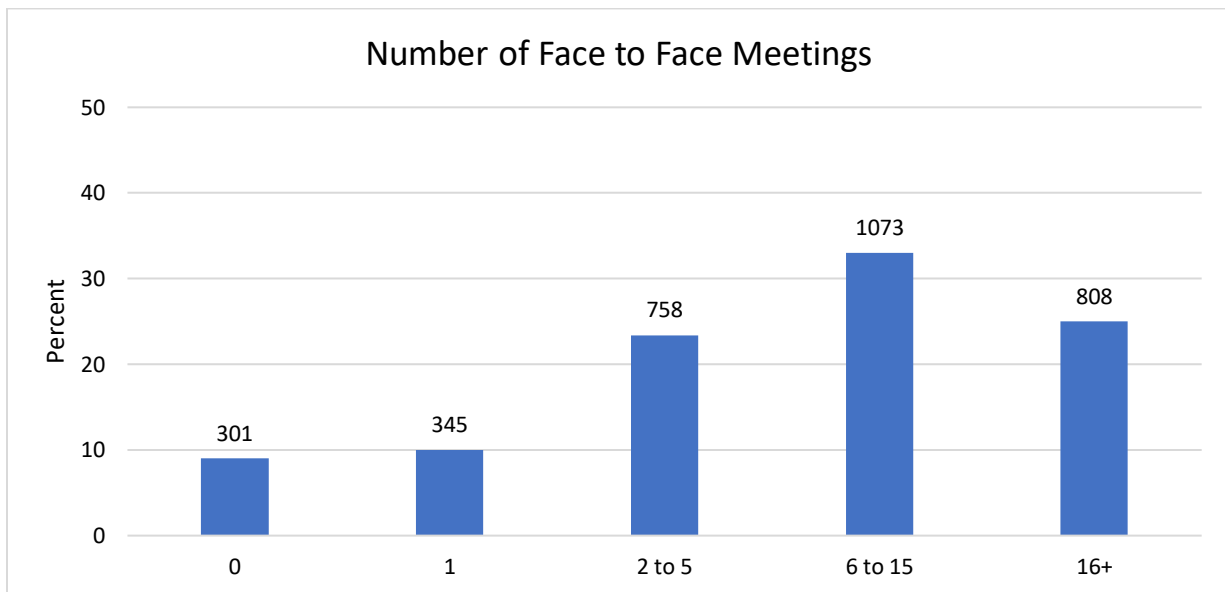
Meeting Format Preferences and Perceived Quality of Review

- As shown in Figure 2, reviewers overall prefer F2F meetings, but preference varies with meeting type. The results also show that 15% of all reviewers surveyed experienced technical difficulties with their Zoom meeting.
- Table 4 shows that gender and career stage are associated with small differences in relative preference for Zoom vs. F2F meetings.

- Figure 3 shows overall reviewer perceptions of the quality of Zoom meetings compared to F2F.
- Figures 4-7 show the data broken down according to meeting type. Focusing on the overall responses and on responses for chartered study sections, the data show that about 60% of reviewers thought the quality of Zoom meeting was the same as that of F2F. Potential concerns are raised by the finding that roughly 35-50% of reviewers rated quality of discussion and level of reviewer engagement worse compared to that in F2F meetings.

Figure 2. Meeting Format Preference by Meeting Type
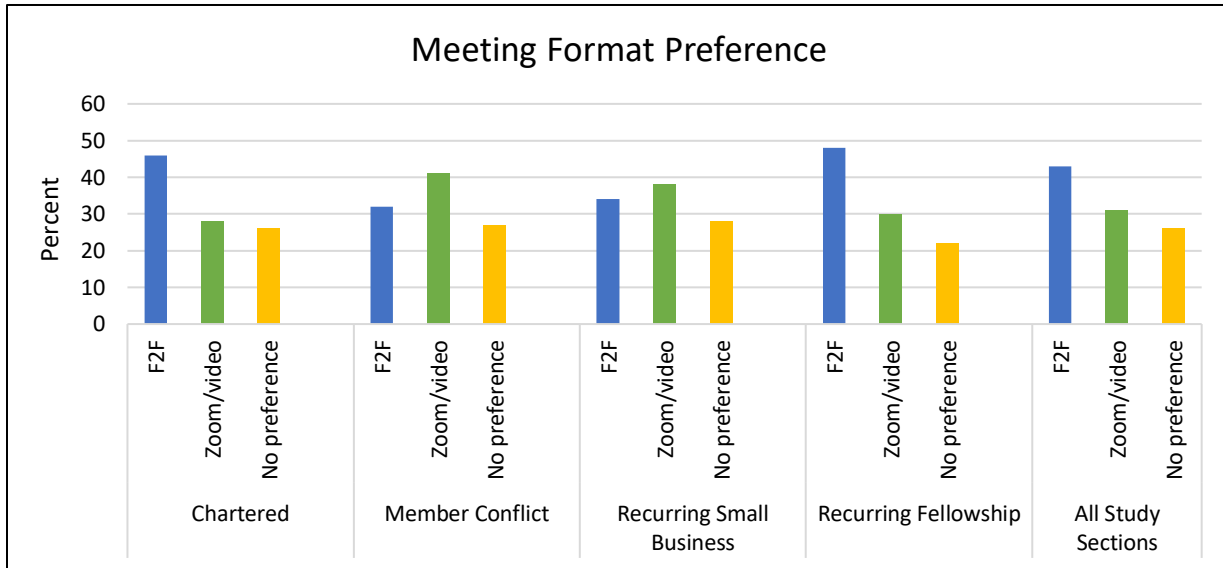
Table 3. Meeting Format Preference by Reviewer Characteristics

Reviewer Characteristics		Meeting Format Preference					
		Face-to-Face		Zoom		No Preference	
		n	%	n	%	n	%
Gender							
	Male	881	46.0	578	30.0	462	24.0
	Female	501	39.0	408	32.0	368	29.0
URM							
	No	1191	43.0	858	31.0	724	26.0
	Yes	108	41.0	81	31.0	72	28.0
Career Stage							
	Assistant Professor	117	41.0	84	29.0	87	30.0
	Associate Professor	420	42.0	304	30.0	275	28.0
	Professor	835	46.0	548	30.0	433	24.0
	Other	49	27.0	75	42.0	55	31.0

Figure 3. Reviewers' Assessments of Peer Review Quality for All Study Sections

Quality of Review for All Study Sections
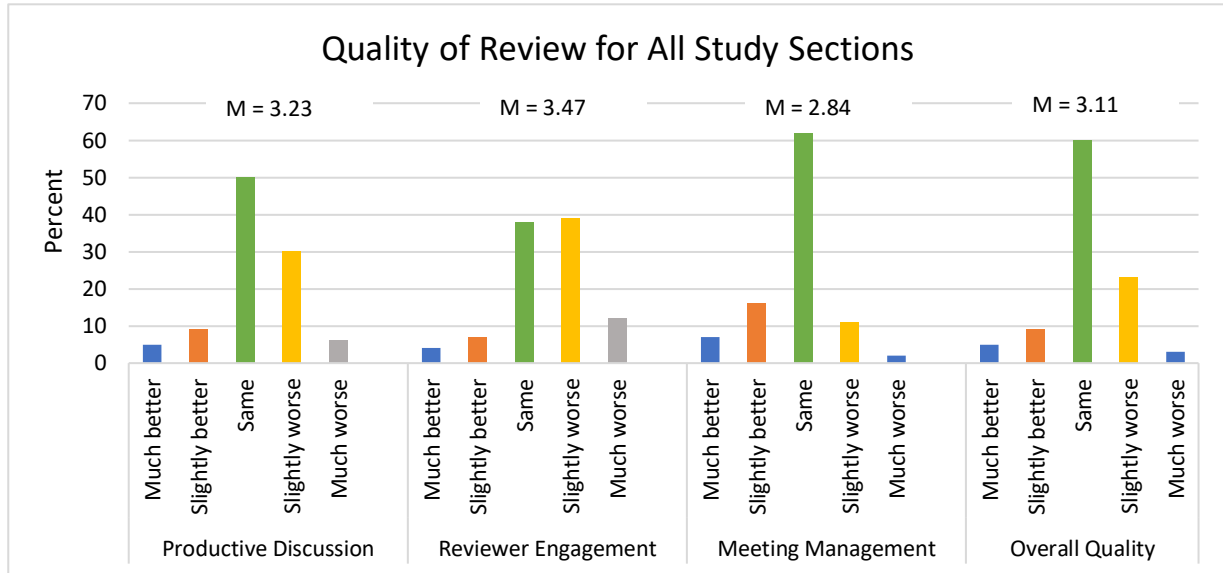
Figure 4. Reviewers' Assessments of Peer Review Quality
    for Chartered Meetings

Quality of Review of Chartered Meetings
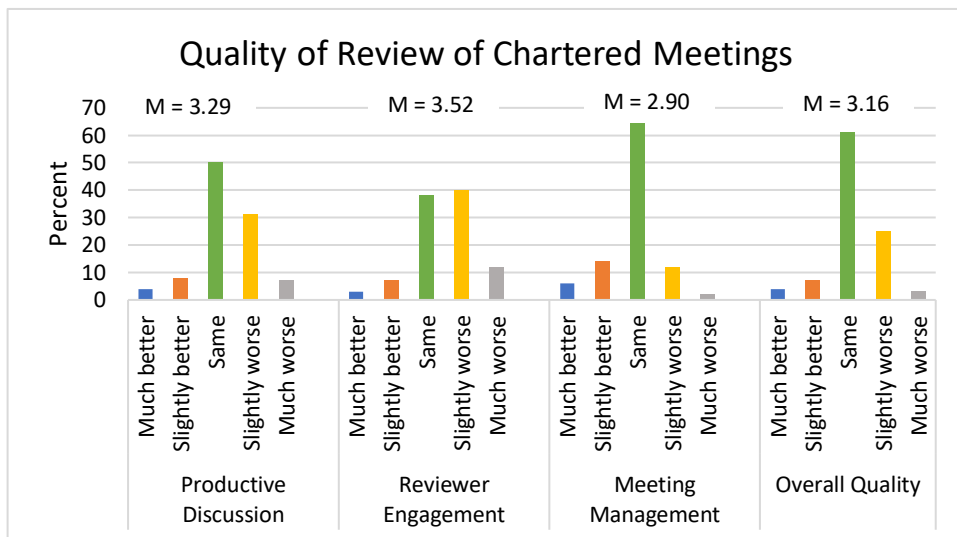
M = 3.29   M = 3.52   M = 2.90   M = 3.16

Figure 5. Reviewers' Assessments of Peer Review Quality
    for Member Conflict Meetings

Quality of Review of Member Conflict Meetings

M = 3.17   M = 3.32   M = 2.67   M = 3.00
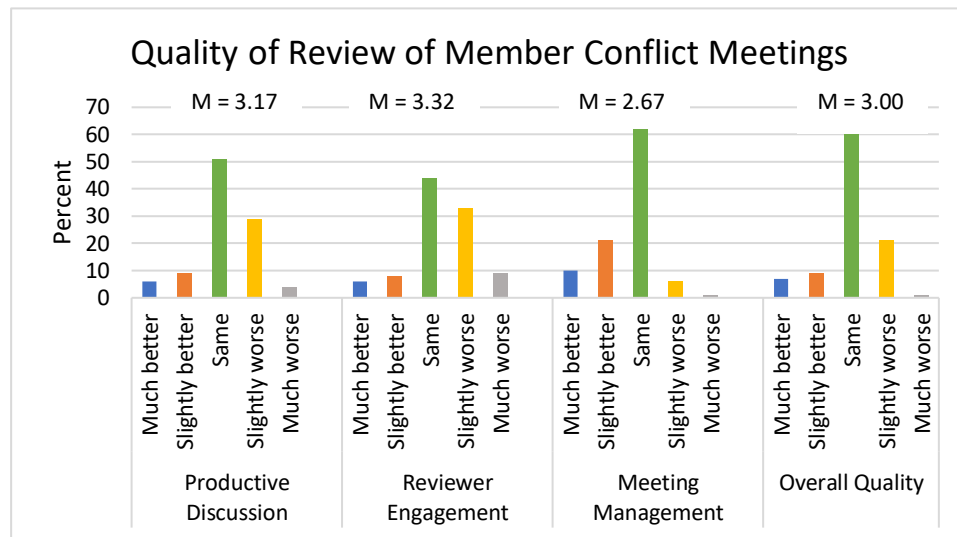
Figure 6. Reviewers' Assessments of Peer Review Quality
    for Recurring Small Business Meetings

Quality of Review of Recurring Small Business Meetings

M = 3.02   M = 3.32   M = 2.73   M = 2.96
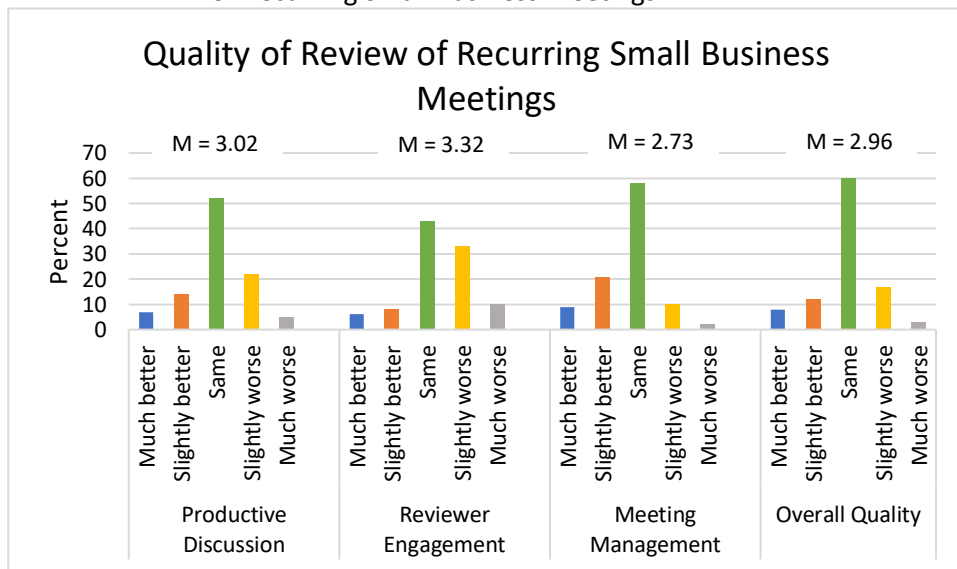
Figure 7. Reviewers' Assessments of Peer Review Quality
    for Recurring Fellowship Meetings

Quality of Review of Recurring Fellowships

M = 3.26   M = 3.53   M = 2.81   M = 3.09
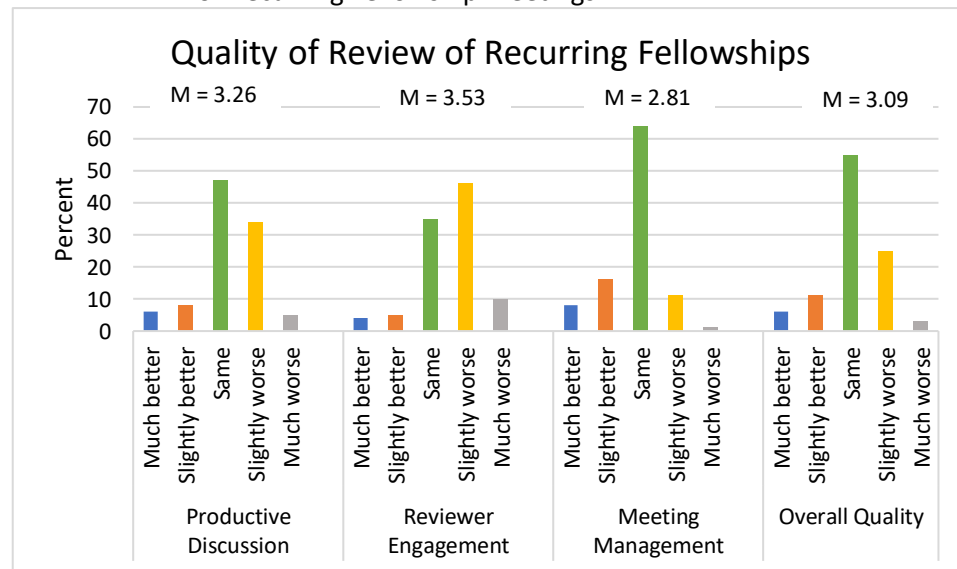
Self-report of Meeting Participation

- Figure 8 shows overall reviewer ratings of their own ability to participate in Zoom meetings compared to F2F.
- Figures 9-12 show the data broken down according to meeting type. The overall ratings show reviewers generally felt able to participate in Zoom meetings similarly to F2F.
- Possible points of concern are that about a third said they contributed less to discussion, and about half said their attention faded more quickly in Zoom meetings, a pattern that was mirrored in ratings of across all kinds of meetings, and especially for meetings of chartered study sections.

Figure 8. Reviewers' Meeting Participation for All Study Sections
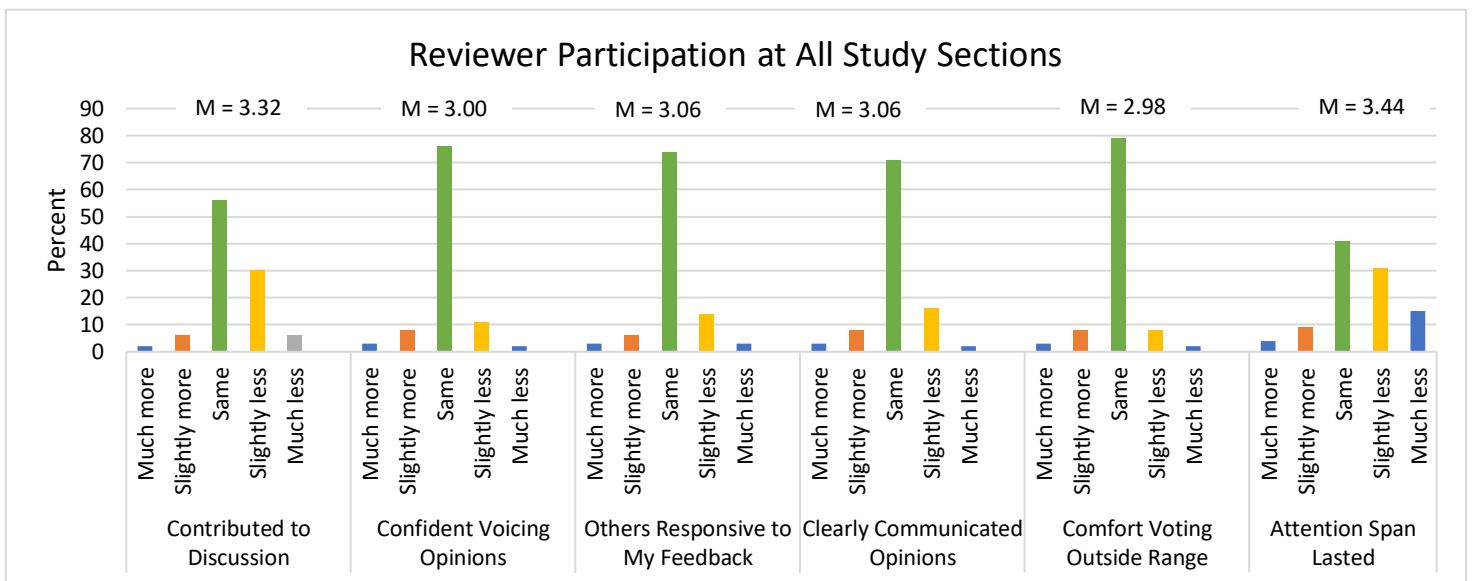
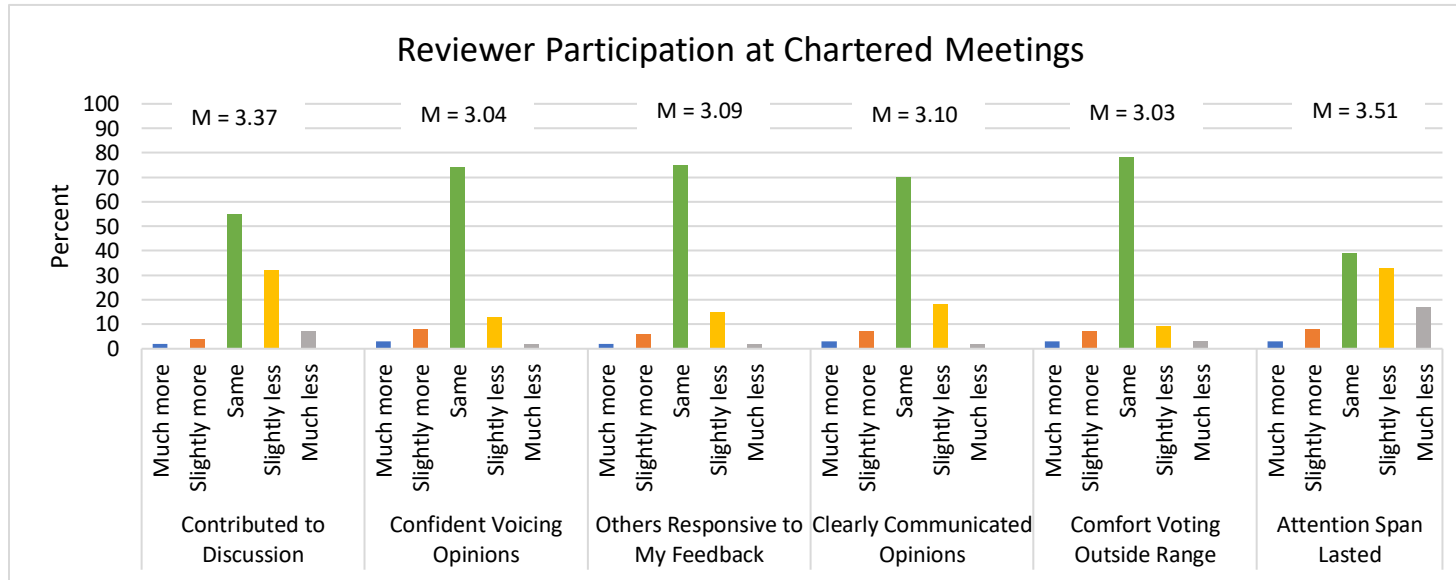Figure 9. Reviewers' Meeting Participation for Chartered Meetings

Figure 10. Reviewers' Meeting Participation for Member Conflict Meetings
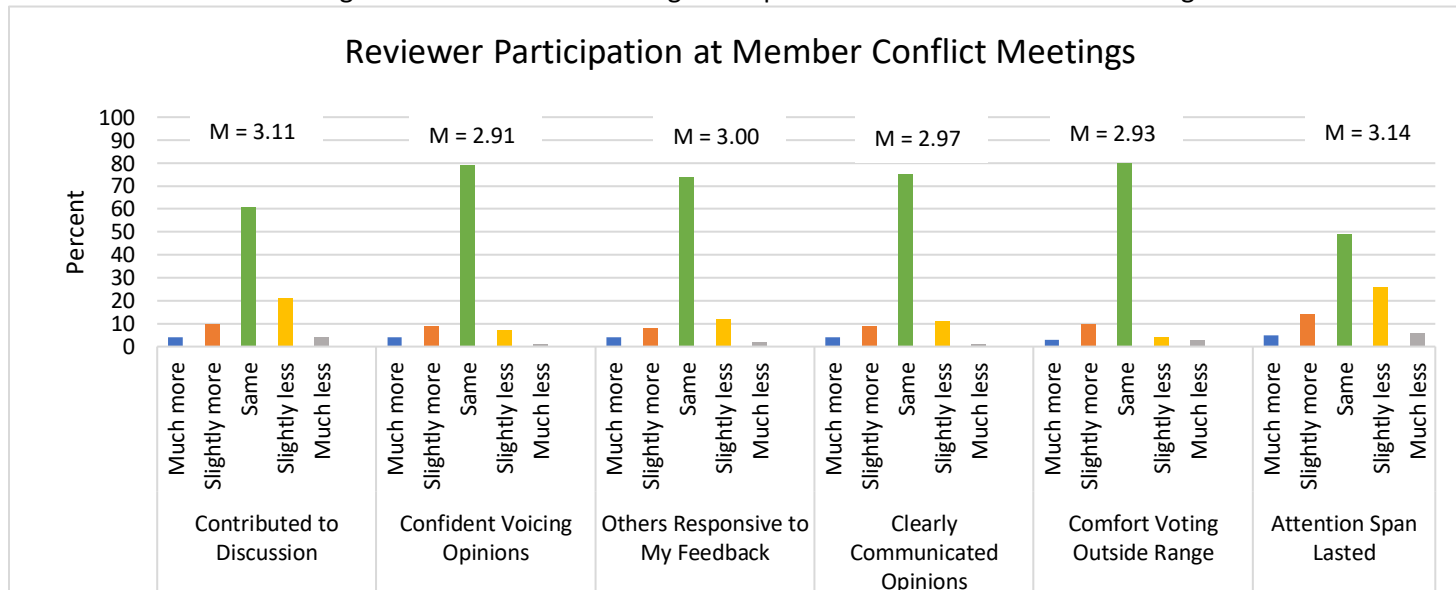
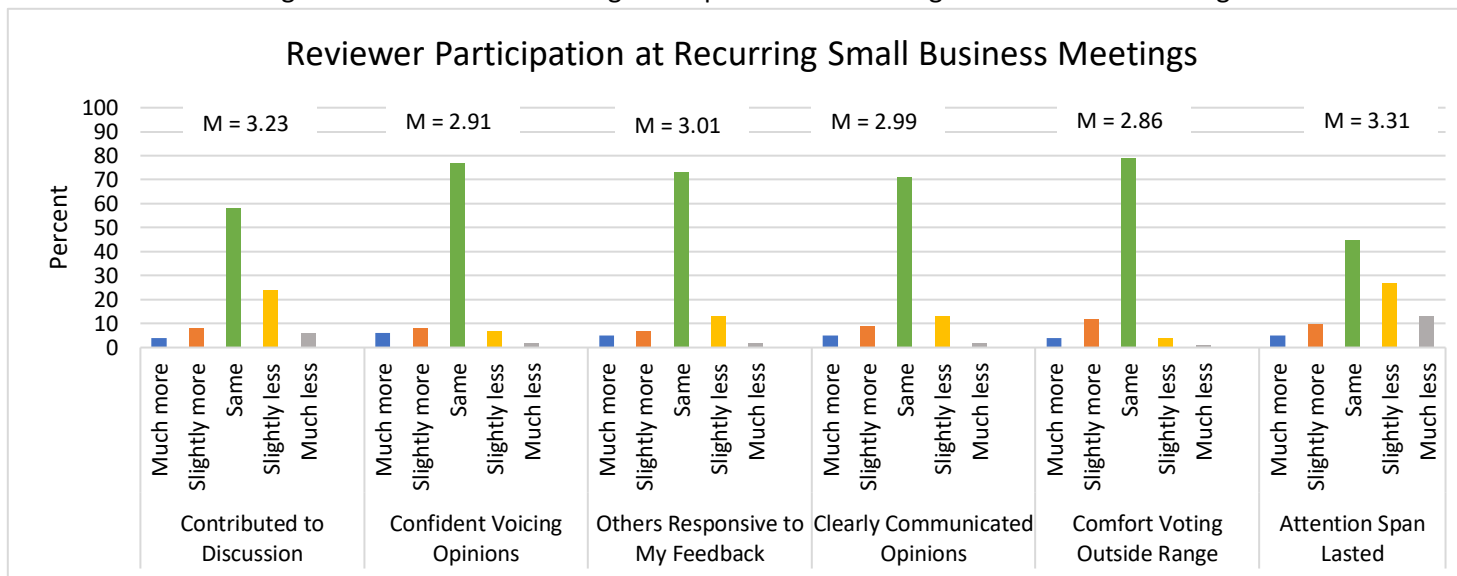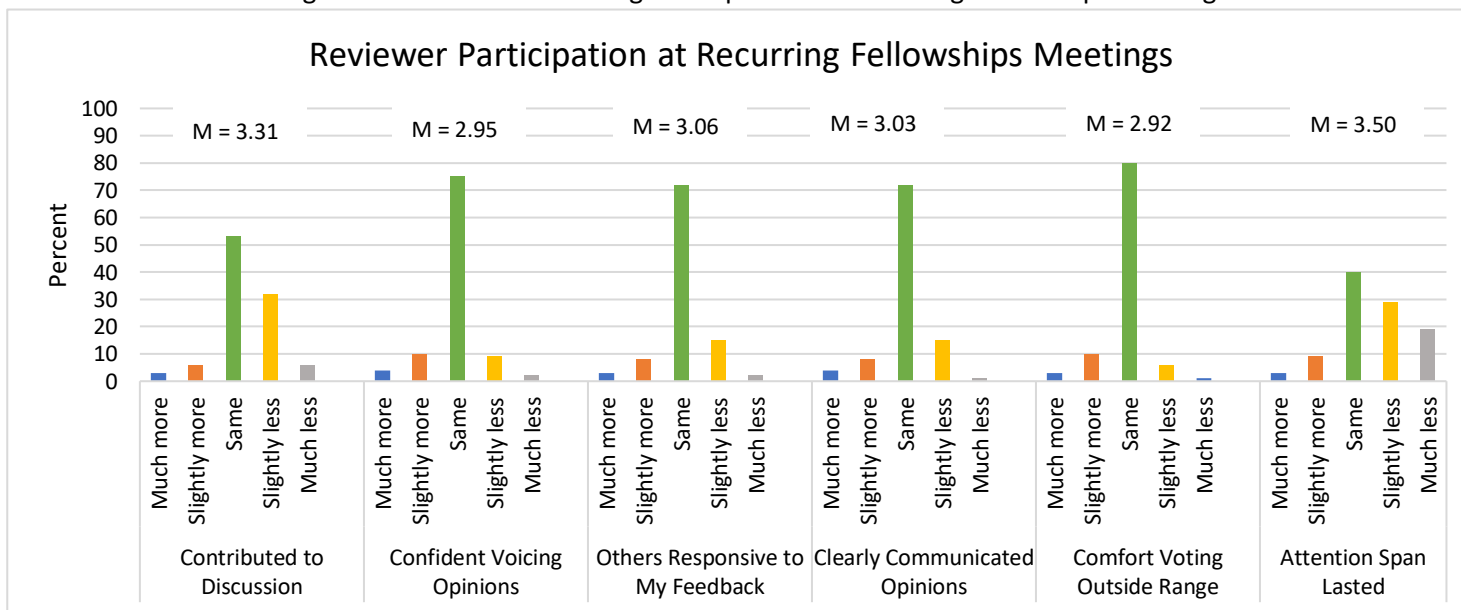Figure 11. Reviewers' Meeting Participation for Recurring Small Business Meetings

Figure 12. Reviewers' Meeting Participation for Recurring Fellowships Meetings

Associations between reviewer characteristics and views of Zoom meetings

We examined associations between reviewer characteristics and reviewer ratings of their participation in meetings. Tables 4, 5, and 6 show a number of statistically significant, but small effect size differences. **When interpreting them, remember that the rating is a comparison of Zoom to F2F meetings, not an absolute rating**. Thus, the statistically significant finding that URMs differed from non-URMs on confidence expressing opinions, means that URMs, on average said they were slightly more confident expressing opinions in Zoom meetings compared to F2F compared to non-URMs ratings of their confidence expressing opinions in Zoom meetings compared to F2F.

Group Differences and Reviewer Participation

Table 4. Gender Differences and Reviewer Participation

Gender			
Participation Items	**Males** **M (SD)**	**Females** **M (SD)**	**t-test and *r***
Contribute to Discussion	3.30 (.79)	3.34 (.75)	$t(3034) = -1.26$, $p = .208$
Confident Voicing Opinions	2.97 (.64)	3.04 (.64)	$t(3032) = -2.83$, **$p = .005$, $r = .05$**
Others Responsive to Feedback	3.06 (.68)	3.06 (.60)	$t(2721.90) = .237$, $p = .813$
Clearly Communicated Opinions	3.05 (.70)	3.08 (.64)	$t(3029) = -1.08$, $p = .279$
Comfort Voting Outside Range	2.96 (.62)	3.02 (.60)	$t(2911) = -2.66$, **$p = .008$, $r = .05$**
Attention Span Lasted	3.40 (.99)	3.49 (.98)	$t(3025) = -2.45$, **$p = .014$, $r = .04$**

Note: 1) All compared to their normal face to face meetings; 2) the higher the mean value, the less the participation

Table 5. Under-represented Minority Status and Reviewer Participation

URM			
Participation Items	**URM** **M(SD)**	**Not URM** **M(SD)**	**t-test and *r***
Contribute to Discussion	3.26 (.75)	3.32 (.76)	$t(2876) = -1.22$, $p = .224$
Confident Voicing Opinions	2.90 (.60)	3.00 (.63)	$t(2875) = -2.55$, **$p = .011$, $r = .05$**
Others Responsive to Feedback	2.98 (.68)	3.06 (.63)	$t(2844) = -2.01$, **$p = .045$, $r = .04$**
Clearly Communicated Opinions	2.96 (.69)	3.07 (.66)	$t(2873) = -2.44$, **$p = .015$, $r = .05$**
Comfort Voting Outside Range	2.93 (.66)	2.98 (.60)	$t(260.68) = -1.10$, $p = .272$
Attention Span Lasted	3.44 (1.03)	3.44 (.97)	$t(2868) = .044$, $p = .965$

Note: 1) All compared to their normal face to face meetings; 2) the higher the mean value, the less the participation

Table 6. Career Stage and Reviewer Participation

Career Stage					
Participation Items	Assistant Prof. M(SD)	Associate Prof. M(SD)	Professor M(SD)	Other M (SD)	One-way ANOVA and Omega^2
Contribute to Discussion	3.22 (.84)	3.31 (.74)	3.34 (.79)	3.25 (.80)	F (3, 3114) = 1.86, p = .134
Confident Voicing Opinions	2.97 (.78)	3.01 (.62)	3.01 (.64)	2.84 (.61)	*Welch's* F (3, 459.41) = 4.16, **p = .006**, **ω² = .003.**
Others Responsive to Feedback	3.01 (.66)	3.07 (.63)	3.07 (.67)	2.96 (.61)	F (3, 3077) = 2.17, p = .089
Clearly Communicated Opinions	3.02 (.74)	3.08 (.64)	3.07 (.69)	2.97 (.65)	F (3, 3108) = 1.46, p = .225
Comfort Voting Outside Range	2.97 (.72)	2.97 (.64)	2.99 (.59)	2.93 (.58)	*Welch's* F (3, 438.13) = .550, p = .648
Attention Span Lasted	3.44 (1.00)	3.45 (.95)	3.45 (1.00)	3.30 (1.00)	F (3, 3105) = 1.11, p = .346

Note: 1) All compared to their normal face to face meetings; 2) the higher the mean value, the less the participation; 3) Post hoc analyses showed that reviewers who hold "other" career titles were significantly more confident voicing their opinions than were associate professors and full professors (compared to their normal face to face meetings).

Open-ended, Qualitative Responses

The major themes from respondent comments in free fields asking for positive and negative aspects of Zoom review meetings are summarized in Table 8.

Table 8. Major Themes

Positive	Negative
No travel	No social interactions/camaraderie missing
Saves time, cost and energy	Prefer face-to-face meeting
Reduces carbon footprint	Less discussion
Better than teleconference or virtual meeting (Chat-based)	Physical cues missing

Recommendations from Reviewers

- Have one face-to-face meeting out of 3 meetings for social interactions and camaraderie
- Have a timer for discussions to avoid long meetings
- Provide more breaks to reduce Zoom fatigue
- Include a social hour for networking

Quantitative Analyses of CSR's First All-Virtual Review Cycle Using Zoom

Executive Summary

This ongoing analysis is examining the efficiency of Zoom as a format for peer review meetings. The present analyses examine meeting duration, roster composition and reviewer workloads. For all components, except for meeting duration, information from a sample of meetings from the most recent all-Zoom round (2020/08 or 2020/10 advisory council), necessitated by the COVID-19 pandemic, was compared to their in-person equivalents from the 2020/05 or 2020/01 advisory council round. The results indicate that the quality of review has been sustained in the transition from in-person review to Zoom, although there may be some concern over the longer length of meetings and shorter break times when using Zoom:

- Overall, average active discussion time per application did not vary by meeting size and meeting type (19-21 minutes), although discussions were slightly longer than the target of 15 minutes per application. Of the sample of small business panels, fellowship, and standing study sections, the majority of meetings started at 9:00 AM EST (67 meetings) and at 10:00 AM EST (33 meetings); on day 1 the majority of these meetings ended between 6-6:59 PM EST with three meetings ending between 8-8:30 PM EST. Five panels in the sample took no breaks during the meeting.
- While negligible trends in applications loads were observed across the sample, fluctuations in meeting application counts can contribute to trends observed in roster composition. Average reviewer workloads slightly decreased from in-person rounds to Zoom rounds, although generally remain in an appropriate range.
- Roster composition analyses showed that standing member participation was retained at an equivalent level. There were no notable differences in female and minority representation between in-person and Zoom rounds.
  - There were slight increases in the average representation of ad hoc reviewers who reviewed regularly (13-24 reviewers) on panels and slight decreases in the average representation of ad hoc reviewers who were newly reviewing (0 reviews or 1-5 reviews) in relation to the entire panel.
  - In the sample, 85.7% of panels in the in-person round and 89.9% of panels in the Zoom round had no instances of reviewer overuse (reviewers with over 45 meetings in the last 12 years). Instances of reviewer overuse have decreased from 24 instances in the in-person round to 13 instances in the most recent Zoom round.

Limitations include not explicitly accounting for meeting application count fluctuations between rounds, and the small number of meetings in the sample for fellowships and small business panels for comparative analyses.

Data Summary and Methods

Data for a sample of approximately six meetings per integrated review group (IRG), composed of a mixture of meeting types (n=143; standing study sections, fellowships, small business panels, recurring special emphasis panel (SEP), etc.) were collected for this analysis. The full sample was used for meeting duration analyses. Meeting duration information was collected through a data collection sheet distributed to the extramural support assistants (ESAs) on meeting start and finish times, start times of the first application, and detailed information about break times. A subset of these 119 meetings, comprised of standing study sections, fellowships, and small business panels only, were further analyzed for roster composition and reviewer workloads by matching them with their in-person meeting equivalents from previous rounds. Most meetings in the comparison subset are matched to their 2020/05 in-person equivalent, but in order to increase the comparison base for fellowships and small business panels, some are matched to their 2020/01 in-person equivalent. For the purposes for this analysis, "Zoom" will refer to the 2020/10 or 2020/08 meetings, and "in-person" will refer to the matched 2020/05 or 2020/01 equivalents.

Table 1 displays meeting counts by scientific division and meeting type, and the proportion of meetings that were included or excluded for the comparison subset (roster composition, reviewer workload). Exclusion criteria for the

comparison analyses included meetings that ran in formats other than in-person in the previous rounds, recurring SEPs as only one meeting had a recent in-person match, and member conflict SEPs as their reviewer composition is more variable from round to round. Appendix A shows a graphical representation of included/excluded meetings for comparison analyses by division and meeting type. Table 2 displays the parameters for meeting size, categorization for which is based on the meeting size of the Zoom round, and the counts of meetings in each category. It should be noted that for comparison analyses, fellowships and small buisness panels were rarely divided up by meeting size due to the small sample of meetings represented.

Table 1. Meeting Types by Division	DABP	DBIB	DNDA	DPPS	DTCS	Grand Total
Excluded (n=24)	8	5	4	3	4	24
Included (n=119)	21	31	20	21	26	119
Fellowship	2	3	2	2	3	12
Small Business	1	3	1	4	3	12
Standing Study Section	18	25	17	15	20	95
Grand Total per Division	29	36	24	24	30	143

DABP – Division of AIDS, Behavior and Population Sciences; DBIB – Division of Basic and Integrative Biological Sciences; DNDA – Division of Neuroscience, Development and Aging; DPPS – Division of Physiological and Pathological Sciences; DTCS – Division of Translational and Clinical Sciences

Table 2. Meeting Types by Meeting Categorization	Large	Medium	Small	Grand Total
Meeting Size Parameters	90+ apps	56-89 apps	1-55 apps	
Excluded (n=24)	1	7	16	24
Included (n=119)	35	65	19	119
Fellowship	3	6	3	12
Small Business	5	5	2	12
Standing Study Section	27	54	14	95
Grand Total per Meeting Size	36	72	35	143

Information for comparative analyses were primarily extracted from Query View Report (QVR) and the Committee Management Module (CMM) of the IMPAC II database. Mail reviewers were excluded for all roster analyses, while early career reviewers (ECRs) were included. Information regarding female and minority representation were extracted from CMM, while information regarding title rank was manually extracted from roster reports. CSR's Informatics Team provided cross-sectional reviewer-level data on prior extent of service, adjusted for one week before the meeting to get the most accurate data at time of recruitment. Meeting counts for prior extent of service include NIH review meetings (chartered and SEPs) and NIH Institute/Center National Advisory Council meetings for application funding, as well as telephone meetings in the last 12 years. Non-FACA (Federal Advisory Committee Act) meetings, mail reviews, and CSR rump SEPs (meetings divided from a primary meeting for administrative purposes or management of conflicts) were excluded from meeting counts.

Analyses

Meeting Duration during Zoom Rounds

<u>Introductions</u>

Length of introduction times for the Zoom round varied slightly, with an average introduction lasting approximately 35 minutes and the range of introduction times by meeting type ranging from 24-38 minutes. Duration of introductions

fluctuated slightly more when assessed by both meeting size and meeting type (Figure 1); distribution of introduction time by roster size had significant variability with large rosters (31+ reviewers) taking a mean introduction time of 40 minutes (standard deviation: 12.4), medium rosters (21-30 reviewers) taking a mean time of 33 minutes (standard deviation: 12.4), and small rosters (14-20 reviewers) taking 25.6 minutes (standard deviation: 11.5) (Appendix B). When assessed by meeting size alone, there was significant variability in introduction times (Appendix C). Large meetings had a mean of 39.8 minutes (standard deviation: 9.9) with a range of 48 minutes, medium-sized meetings had a mean of 37.5 minutes (standard deviation: 13) with a range of 64 minutes, and small meetings had a mean of 25.5 minutes (standard deviation: 9.0) with a range of 46 minutes. The data indicate that out of the three meetings sizes, medium-sized meetings had the largest amount of variability in introduction times. High variability in medium-sized meetings could in part be attributed to a larger number of meetings n= 72 as opposed to large meetings (n=36) and small meetings (n=35). Longer introduction times can be attributed to larger roster sizes and thus, longer panel introductions.

Figure 1. Distribution of Average Introduction Times, by Meeting Size and Meeting Type (n=143)
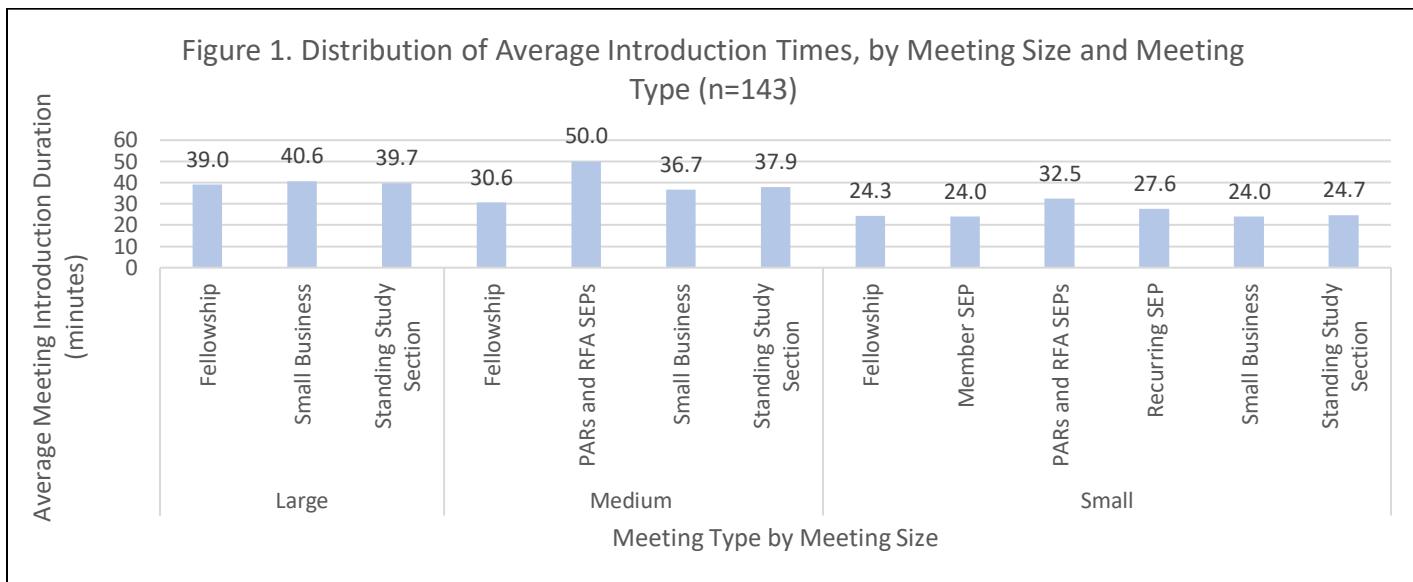
Discussion times

Overall, average active discussion times for time per application did not vary much by meeting size and ranged from 19-21 minutes (Table 3). Average active discussion time by meeting size and meeting type also did not fluctuate much and ranged from 19-21 minutes, although was slightly longer than the target 15 minutes per application (Figure 2). Distributions for discussion time by meeting size indicate low variability across large, medium,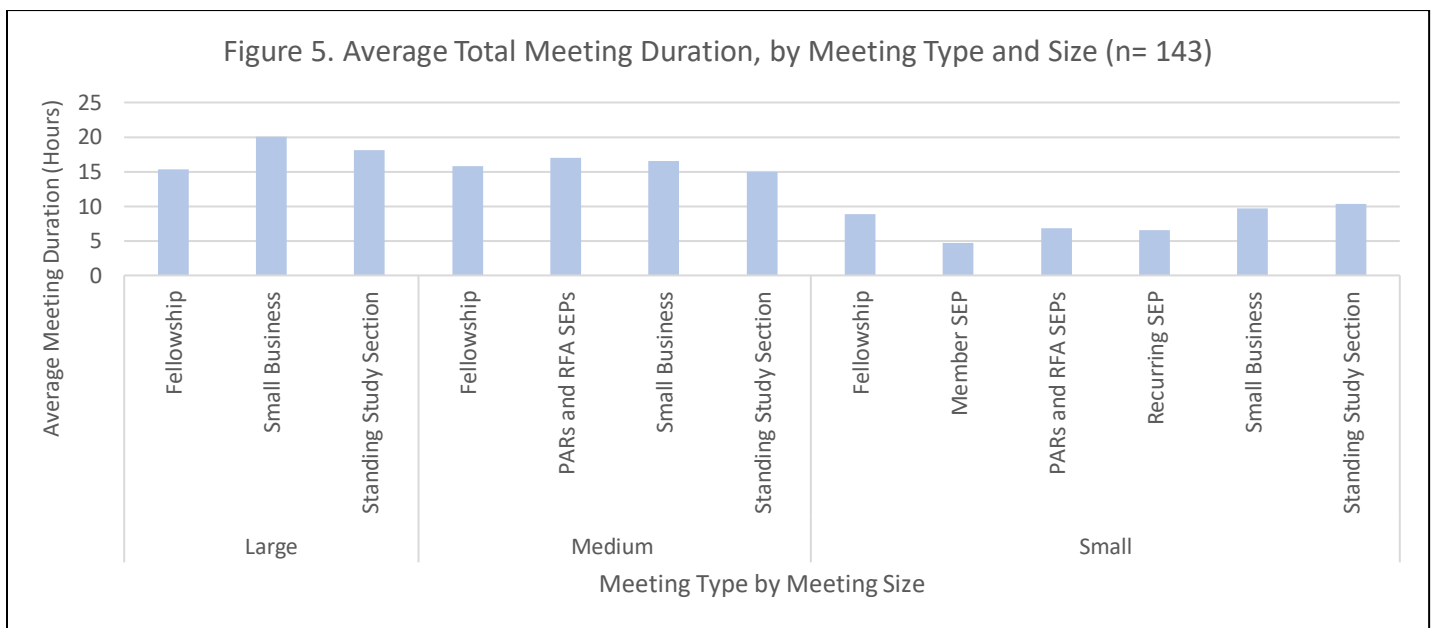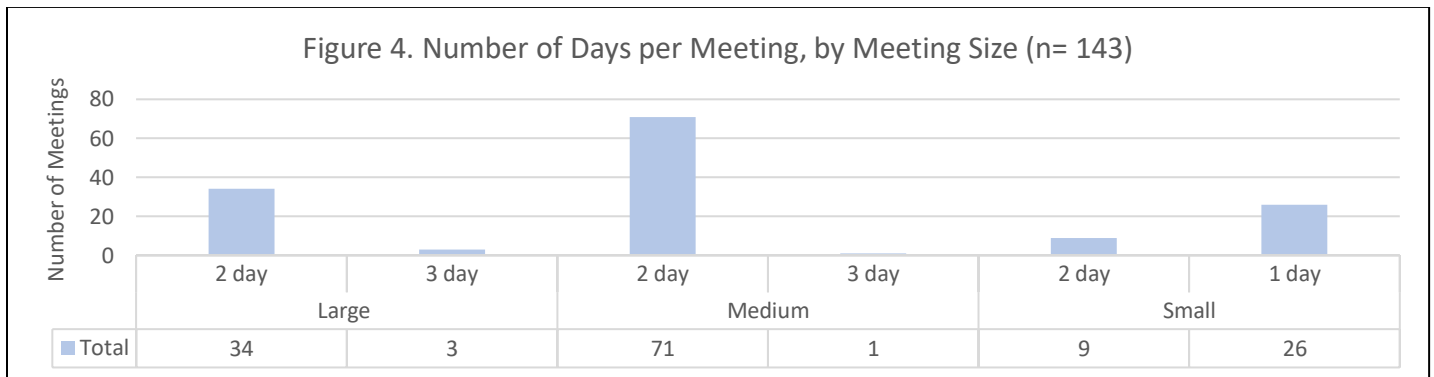 and small meetings all with standard deviations between 2.3-3.3; full distributions can be found in Appendix D. Figure 3 demonstrates the association between meeting application counts and average active discussion time per application for small business panels, fellowships and standing study sections only; as shown by the data there is variability between the two variables. The Pearson correlation coefficient for these data is -0.33, indicating that there is a very small, weak association between the two. In this instance it cannot be concluded that there is a correlation between discussion times and meeting application counts.

Table 3: Active Discussion Time per Application	
Meeting Size	**Average Discussion Time (minutes)**
Large	19.3
Medium	20.7
Small	21.2
All Meeting Average	20.5

Figure 2. Average Active Discussion Time per Application, by Meeting Type and Size (n=143)

Figure 3. Active Discussion Time per Application (mins) and Meeting Application Counts (n=143)
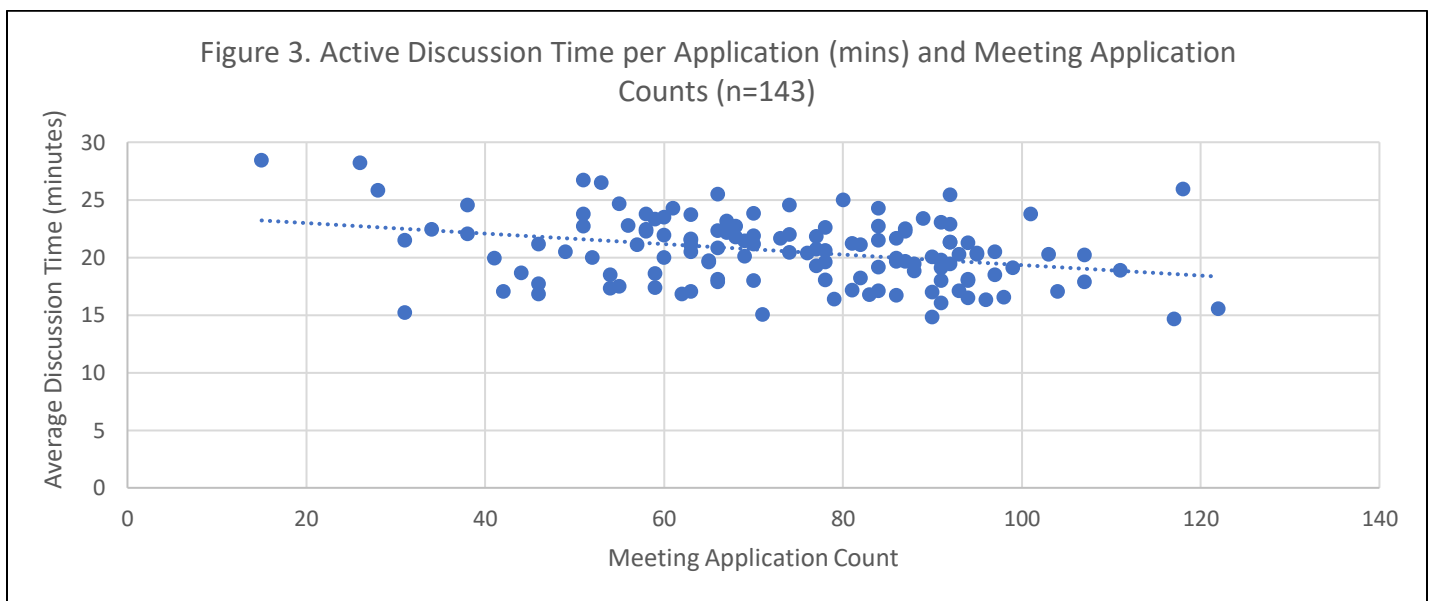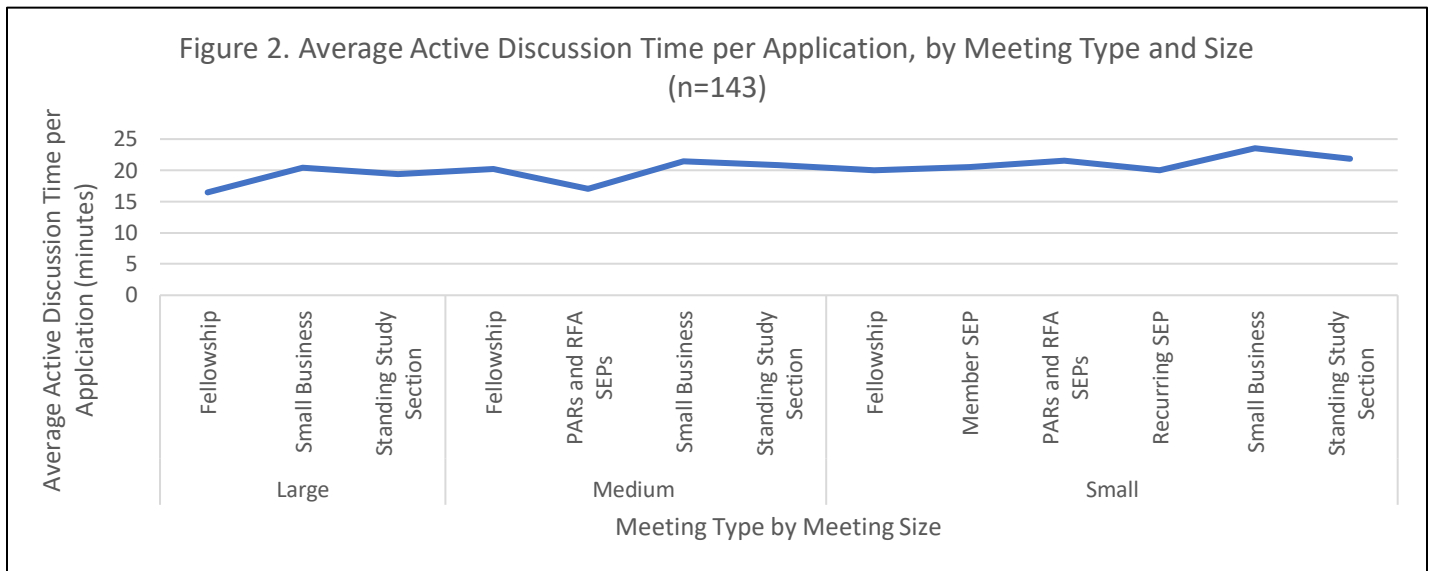
Total Meeting Duration

Total meeting duration varied only slightly within each meeting size category; for small meetings duration ranged from 4-10 hours, for medium meetings duration ranged from 14-17 hours and for large meetings duration ranged from 15-20 hours. Figure 4 displays the number of days per meeting, by meeting size. On average for 2-day meetings, the first day for large meetings ran 11.6 hours, for medium meetings ran 11.02 hours and for small meetings ran 10.3 hours. On the second day for the same meetings, on average meeting duration ran 8.1 hours for large meetings, 6.0 hours for medium-sized meetings and 3.8 hours for small meetings. The first day for 2-day meetings did not run longer than 11.5 hours and the second day for the same meetings similarly did not run longer than 11 hours. Through analyzing distributions for the first day versus the second day for 2-day meetings is can be seen that there is a large amount of variability for both days (day 1: mean: 556.6 mins, ~9.2 hours, standard deviation: 60.3 mins; day 2: mean: 388.6 mins, ~6.5 hours, standard deviation: 126.0 mins). Variability for the first day versus the second day for 2-day meetings can be observed in Appendix E. Average meeting duration both by meeting type and size can be observed in Figure 5. Variation of meeting duration within each meeting category size could be attributed to large range of meeting application counts that each categorization covers. Of the sample of small business panels, fellowships, and standing study sections the majority of

meetings started between 9:00-9:30 AM EST (67 meetings) and 10:00-10:30 AM EST (33 meetings); on day 1 the majority of these meetings ended between 6-6:59 PM EST with three meetings ending between 8-8:30 PM EST.

Figure 4. Number of Days per Meeting, by Meeting Size (n= 143)

	Large		Medium		Small	
	2 day	3 day	2 day	3 day	2 day	1 day
Total	34	3	71	1	9	26

Figure 5. Average Total Meeting Duration, by Meeting Type and Size (n= 143)

Meeting Type by Meeting Size

Breaks

Total average break times across all days for meetings varied significantly both by meeting size and meeting type. Average break times by meeting type, not taking into account meeting size, ranged from 9-85 minutes (1 hour, 25 minutes), with member conflict SEPs falling on the lower end of the range due to a smaller meeting application counts and small business panels falling on the higher end of the range. There were 5 meetings that did not take any breaks during the meeting – 2 medium-sized study sections, 2 fellowships (one medium and one small), and 1 small PAR/RFA. Average total break times by meeting size and type varied significantly between large/medium and small meetings and varied slightly between large and medium meetings (Figure 6). Break times by meeting size varied significantly; the mean total breaks for large meetings was 88.22 minutes (standard deviation: 31.1), for medium-sized meetings was 77.8 mins (standard deviation: 32.4), and for small meetings was 41.3 mins (standard deviation: 4.6) (Appendix F). Average break times were also classified in relation to total meeting duration; for these purposes short meetings were classified as lasting 1-6 hours, medium-length meetings were classified as 7-15 hours, and long meetings were classified as lasting 16+ hours. Of the meetings in the Zoom round 56 meetings were classified as long meetings, 76 were classified as medium-length meetings, and 11 meetings were classified as short meetings. On average, long meetings took 90 minutes' worth of break, medium lasting meetings took an average of 63 minutes' worth of breaks, and short meetings took an average of 27 minutes' worth of breaks (Appendix G).

Figure 6. Distriubution of Average Break Times, by Meeting Type and Size (n=143)
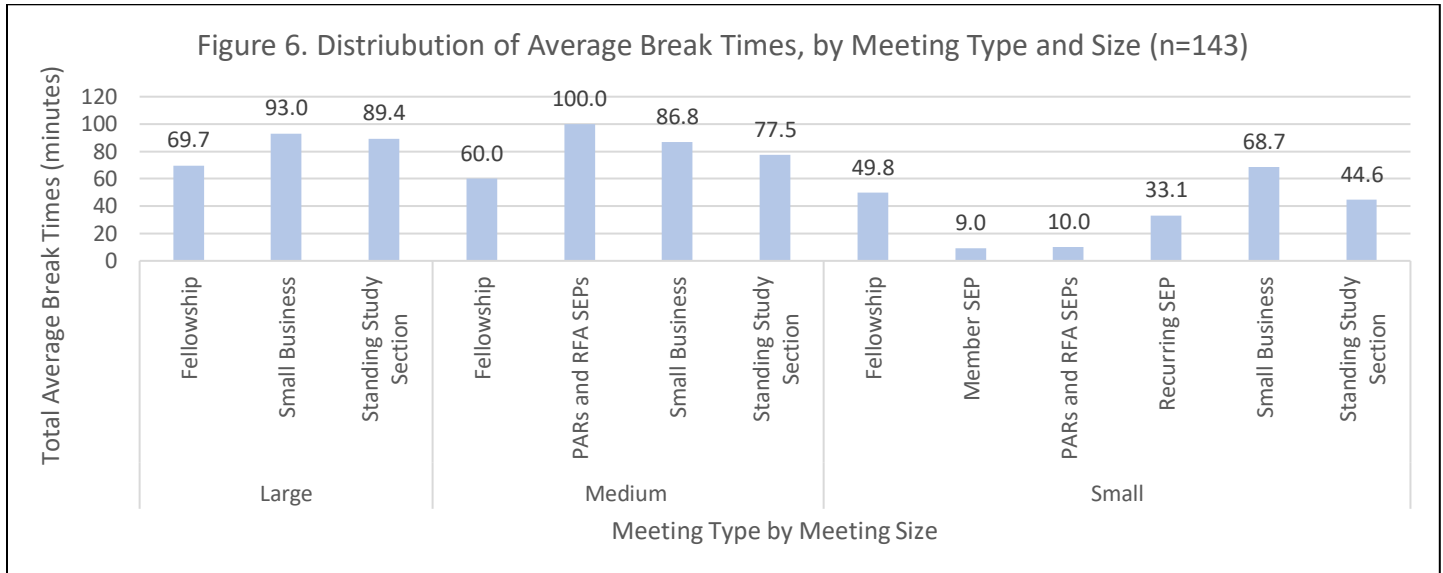
Special Examination of 3-day Meetings

There were three 3-day meetings in the original sample of 143 meetings. Meeting application counts for these three meetings ranged from 89-118 applications (medium to large meetings). Average discussion time per application for 3-day meetings was slightly higher than other meetings and averaged at about 23.5 minutes. On average total break times for all three days averaged about 118 minutes (1 hour, 58 minutes) and introductions averaged approximately 64 minutes (1 hour, 4 minutes).

Meeting Application Counts, Roster Sizes, and Reviewer Workload Trends from In-Person to Zoom Rounds

Comparative Meeting Application Counts

Applications loads were compared for the Zoom round and its in-person equivalent meetings. Appendix H displays descriptive information about the distribution of meeting application counts across the sample for both rounds. Appendix I breaks the distributions down further into meeting type (standing, small business, and fellowships). While negligible trends are accounted for across the entire sample (as well as the subsample of standing study sections), there were a few notable trends in small business and fellowship panels, albeit small sample sizes. In small business panels, most panels in the sample increased in size by an average magnitude of +20 applications (n= 9, including one outlier of +59 applications). The other three small business panels decreased in size by an average magnitude of -3 applications. On the contrary, most fellowship panels in the sample decreased on an average magnitude of -18 applications (n=8, including one outlier of -50 applications). The few fellowship panels that increased in size increased on an average magnitude of +4 applications.
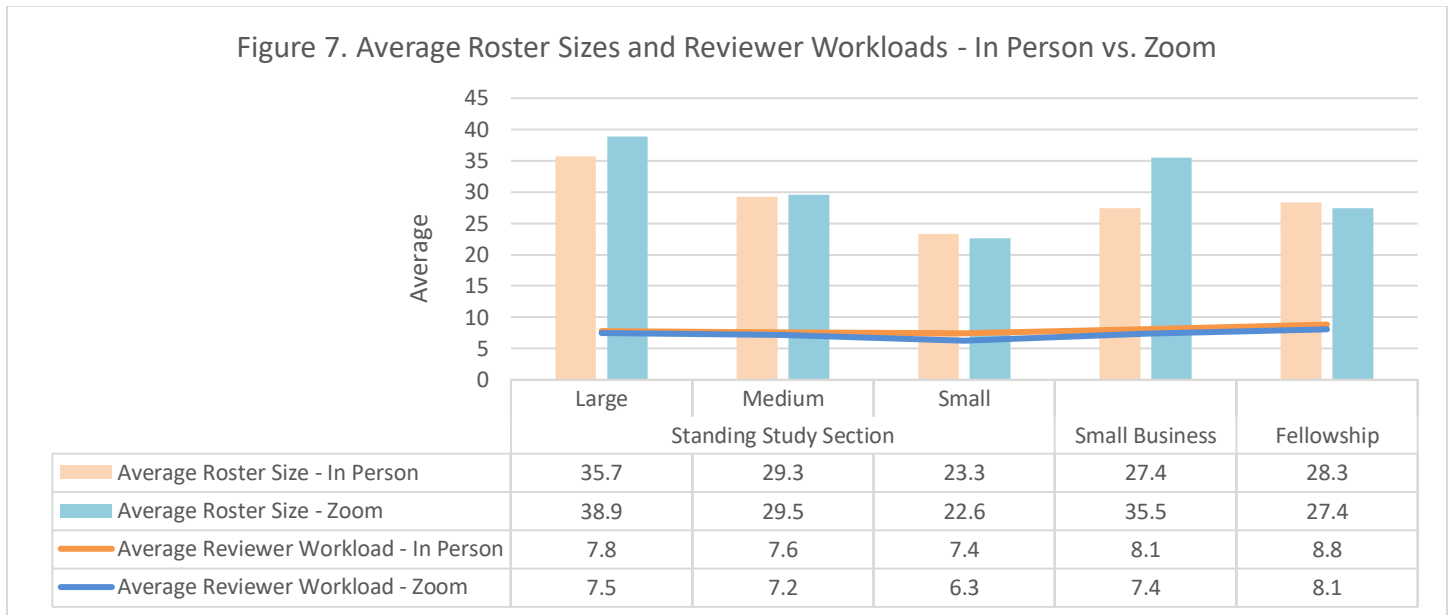
Roster Size and Reviewer Workloads

Average roster sizes and reviewer workloads were compared for meeting types and sizes between the in-person and Zoom rounds (see Figure 7). Fluctuations in roster sizes in small business panels (increased roster sizes in all meeting sizes) can mostly be accounted for by fluctuations in meeting application counts described in the previous section. Fluctuations in roster size may also, to a degree, be reflective of reviewer availability related to the circumstances around the pandemic or the transition to the zoom format. For example, reviewers and researchers may have been directly or indirectly affected by the pandemic which may have affected recruitment (those directly involved in the front lines/research, those who experienced disruptions to research, or those who experienced significant disruptions to work life balance – or, on the contrary, those who find that they have increased availability), and the transition to Zoom

format may have attracted reviewers who experience barriers to attending in-person meetings, or turned away other reviewers who prefer in-person meetings.

Reviewer workloads, defined by the average critiques submitted per reviewer in a meeting, excluding mail reviewers, slightly decreased from in-person to Zoom rounds across all meeting types and sizes, although by a negligible amount. Reviewer workloads generally remained in the appropriate range, although small study sections (n=14) had reviewer workloads close to 6 applications per reviewer in the recent Zoom round.

Figure 7. Average Roster Sizes and Reviewer Workloads - In Person vs. Zoom

	Large	Medium	Small	Small Business	Fellowship
Average Roster Size - In Person	35.7	29.3	23.3	27.4	28.3
Average Roster Size - Zoom	38.9	29.5	22.6	35.5	27.4
Average Reviewer Workload - In Person	7.8	7.6	7.4	8.1	8.8
Average Reviewer Workload - Zoom	7.5	7.2	6.3	7.4	8.1

Roster Composition Trends from In-Person to Zoom Rounds

Rosters were compared for the Zoom round and their in-person equivalents and examined for professional title distribution, demographic distribution, proportion of standing study section members and ad hoc members, and extent of prior review service (ad hoc members only). Mail reviewers were excluded, and ECRs are included for all calculations.

Professional Title Distribution

Figure 8 displays the distribution of title between in-person and Zoom rounds, by meeting type. Overall, limited changes were observed between in-person and zoom rounds in the composition of rosters by title. As expected, small business review meetings had a greater proportion of reviewers with titles other than standard academic titles, and fellowships had limited numbers of assistant professors. Both small business panels and fellowships had a slight increase in assistant professor participation from the in-person to Zoom rounds.

Figure 8. Distribution of Professional Title at In-Person and Zoom Rounds, by Meeting Type (n=119)
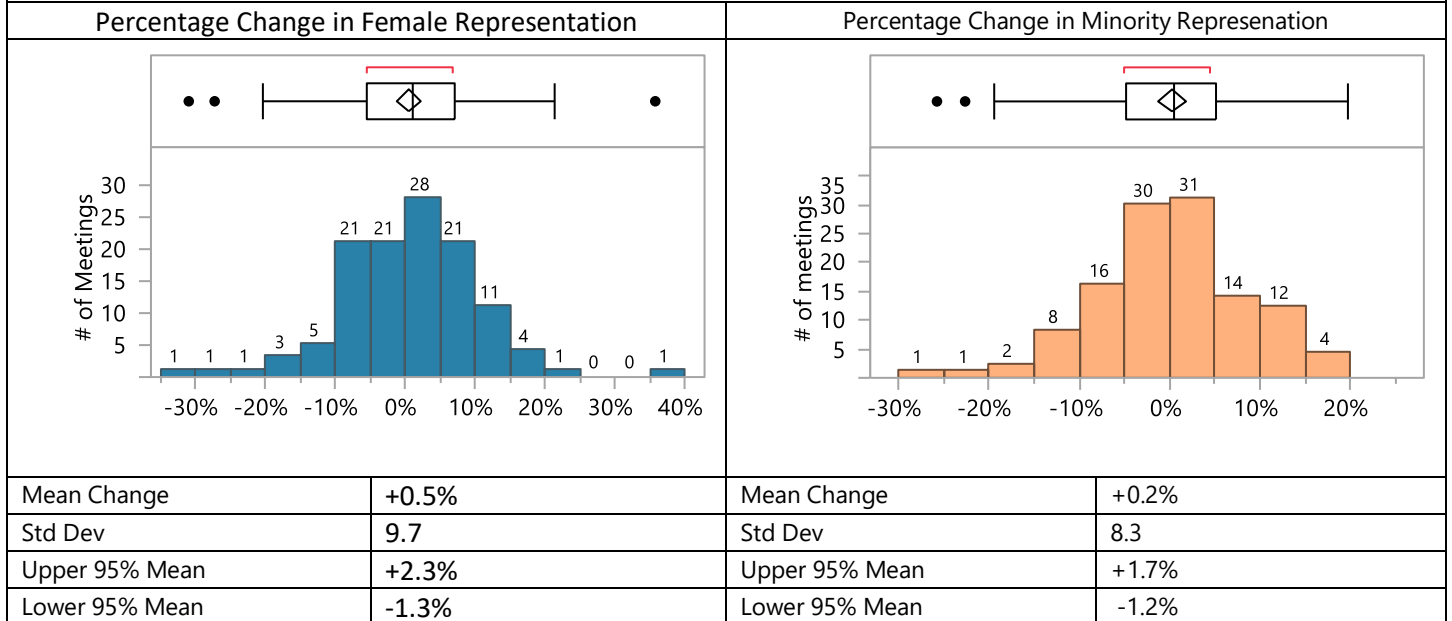
	Standing Study Section		Small Business		Fellowship	
	In Person	Zoom	In Person	Zoom	In Person	Zoom
Other	3.9%	3.4%	35.6%	35.2%	4.1%	2.7%
Assistant Professors	12.6%	12.8%	8.2%	9.6%	0.6%	2.7%
Associate Professors	30.3%	30.6%	25.8%	25.4%	45.3%	45.0%
Professors	53.1%	53.1%	30.4%	29.8%	50.0%	49.5%
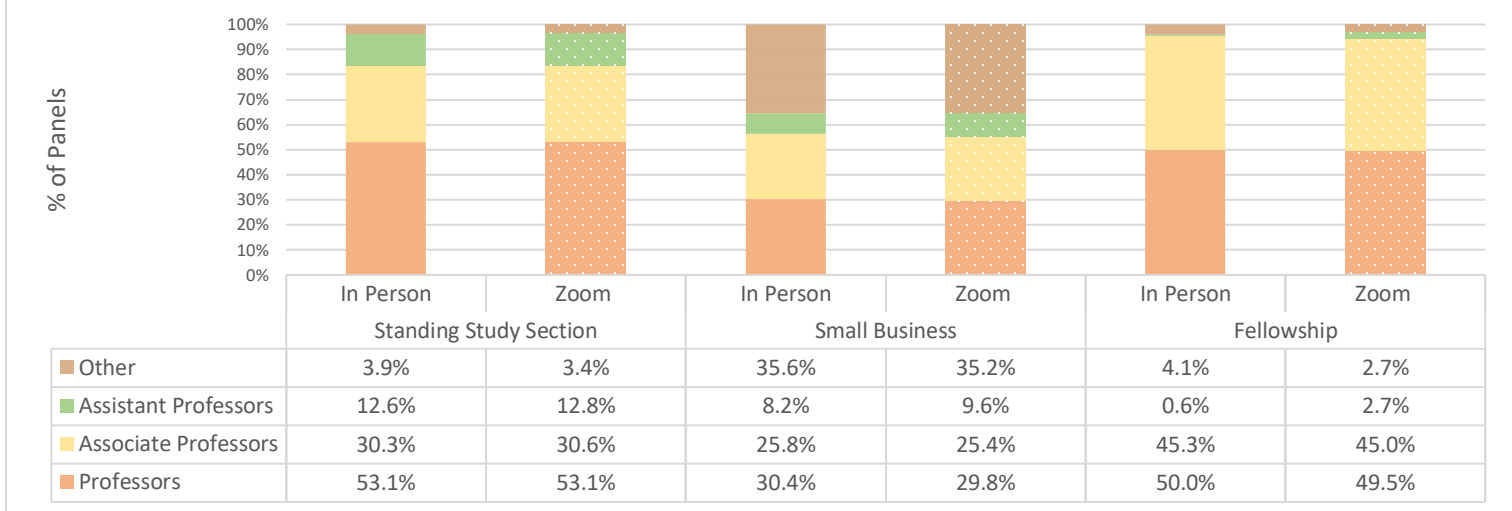
Demographic Distribution

Figure 9 displays the distribution of percent change in both female and minority reviewers represented in the panel compared to their male and non-minority counterparts from the in-person to the Zoom round at the meeting level. 76.5% of panels in the sample did not have a net gain or loss of female representation greater than 10% (mean change +0.5%). Similarly, 76.5% of panels in the sample did not have a net gain or loss of minority representation greater than 10% (mean change +0.2%). Percent change in female representation had a slightly larger spread in distribution than minority representation which displayed a tighter distribution.

Appendix J and K display average female and minority representation in panels across in-person and Zoom rounds, by meeting type as well as study section size. Average percent change at the aggregate level did not increase or decrease greater than 3.2% between the in-person and Zoom rounds.

Figure 9. Meeting-Level Percent Change in Female and Minority Representation from In Person to Zoom Round (n=119)

Percentage Change in Female Representation		Percentage Change in Minority Represenation	
Mean Change	+0.5%	Mean Change	+0.2%
Std Dev	9.7	Std Dev	8.3
Upper 95% Mean	+2.3%	Upper 95% Mean	+1.7%
Lower 95% Mean	-1.3%	Lower 95% Mean	-1.2%

Standing Vs. Ad hoc Reviewer Participation

For standing study sections, standing member participation did not fluctuate much between in-person rounds and Zoom rounds; 88% of standing study section meetings in the sample did not have a net gain or loss greater than three standing members (see Figure 10). Typically, major fluctuations in ad hoc participation on the meeting level accommodate fluctuations in applications loads. When examining proportions of standing member to ad hoc participation amongst divisions, DPPS experienced a 4-percentage point increase in the proportion of ad hoc reviewers in the Zoom round, while the other divisions remained relatively stable (see Figure 11). On average, DPPS had on average 3 more ad hoc participants invited to the panel during the Zoom round and had a relatively consistent standing member participation, although this may be reflective of normal fluctuations in meeting application counts between rounds.

Figure 10. Number of Standing Members Gained or Lost from the In-Person Round to the Zoom Round (n=119)

Number of Standing Members Gained or Lost (Mean Change = +0.09)

Figure 11. Proportion of Standing Members to Ad hoc Participants in Standing Study Section Meetings, In-Person vs. Zoom (n=119)
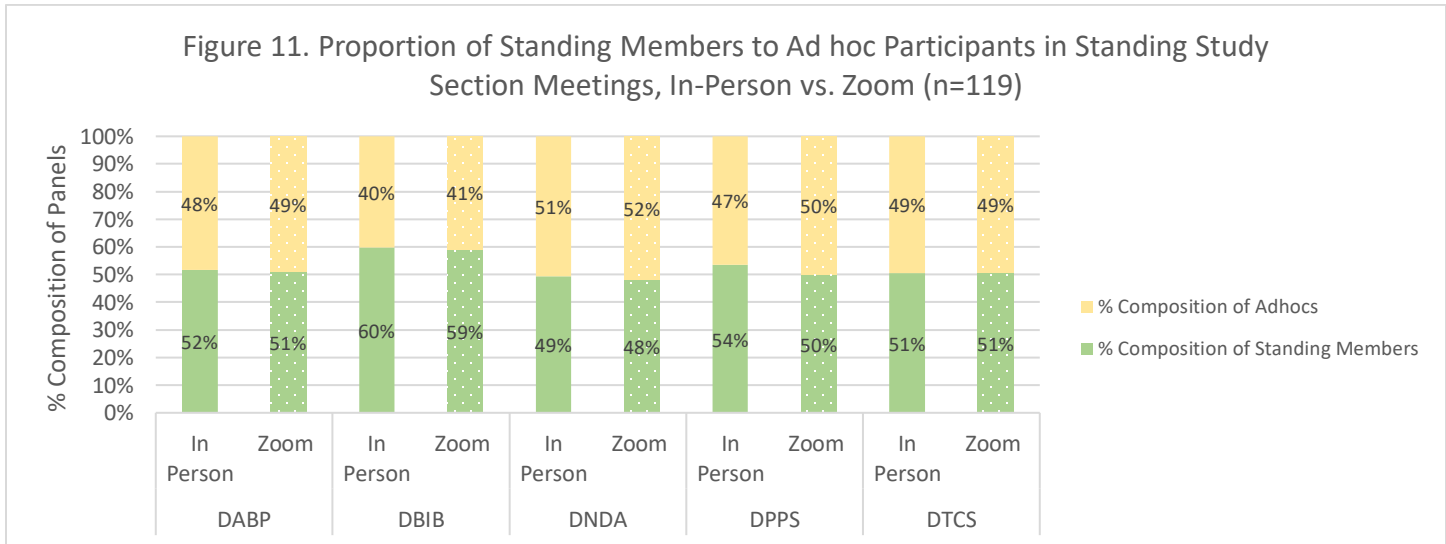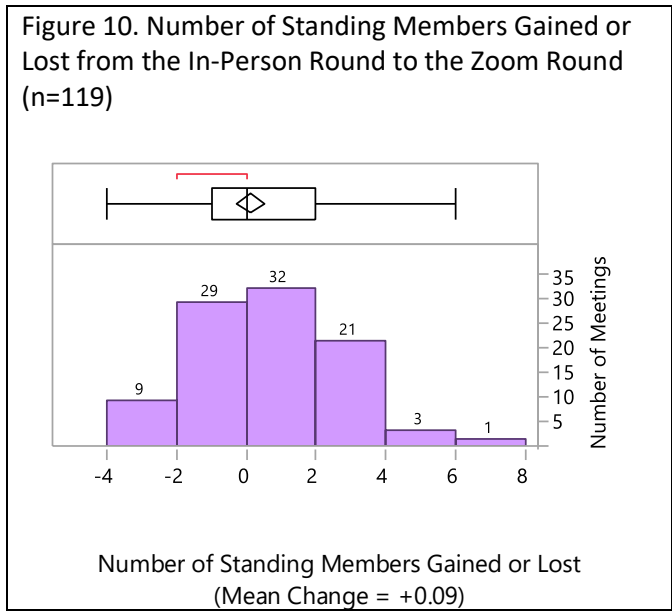
Extent of prior review service in ad hoc reviewers

New Reviewer Recruitment

Figure 12 shows the proportion of ad hoc reviewers invited to the panel by prior review service in the last 12 years, by meeting type between the in-person and Zoom rounds. It should be noted that while standing study sections have standing members that are excluded from this analysis, small business panels and fellowships do not have standing members, so the entire panel of ad hoc reviewers is accounted for in this analysis. ECRs were included in this analysis and should be considered when interpreting data on reviewers with 0 reviews.

Figure 12. Proportion of Ad hoc Reviewers Serving on Panels, by Prior Review Service and Meeting Type
*data accounts for ad hoc only in Standing Study Sections
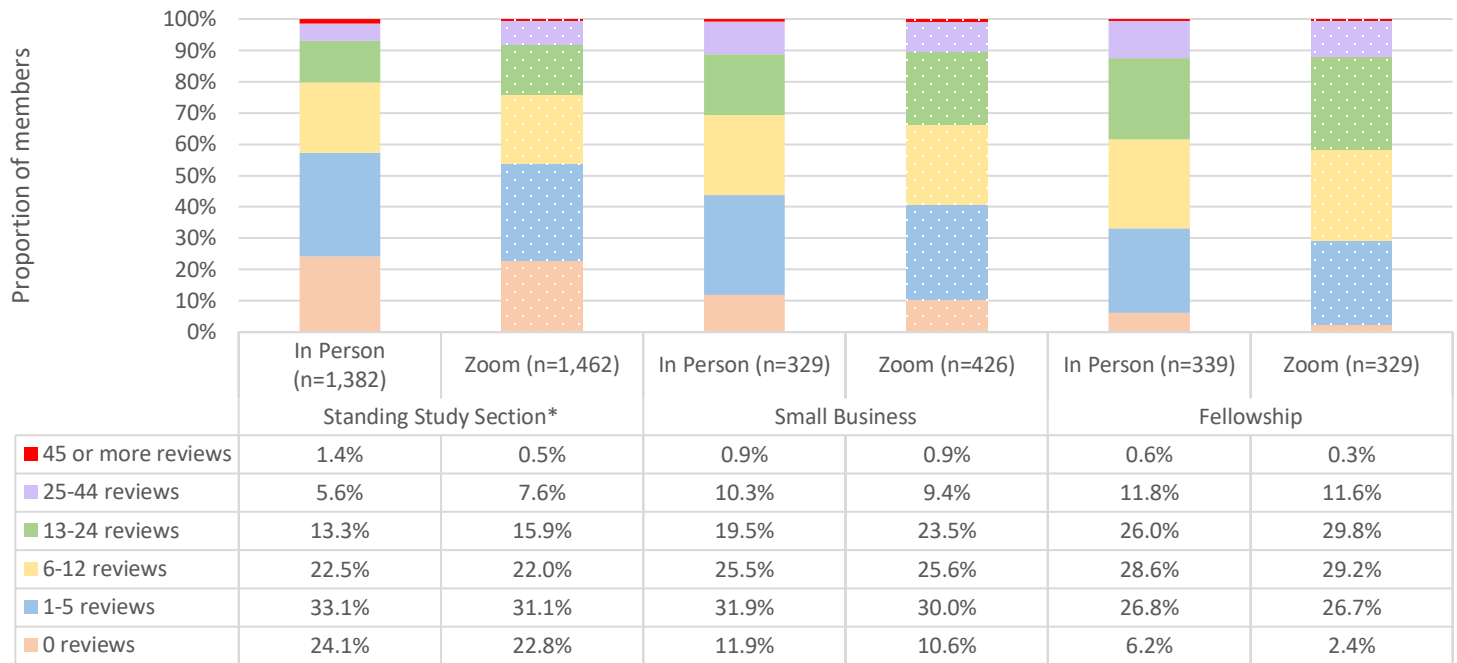
	In Person (n=1,382)	Zoom (n=1,462)	In Person (n=329)	Zoom (n=426)	In Person (n=339)	Zoom (n=329)
	Standing Study Section*		Small Business		Fellowship	
45 or more reviews	1.4%	0.5%	0.9%	0.9%	0.6%	0.3%
25-44 reviews	5.6%	7.6%	10.3%	9.4%	11.8%	11.6%
13-24 reviews	13.3%	15.9%	19.5%	23.5%	26.0%	29.8%
6-12 reviews	22.5%	22.0%	25.5%	25.6%	28.6%	29.2%
1-5 reviews	33.1%	31.1%	31.9%	30.0%	26.8%	26.7%
0 reviews	24.1%	22.8%	11.9%	10.6%	6.2%	2.4%

There were slight increases in the average representation of ad hoc reviewers who reviewed regularly (13-24 reviewers) on panels and slight decreases in the average representation of ad hoc reviewers who were newly reviewing (0 reviews or 1-5 reviews) in relation to the entire panel. However, if you look at the average number of new reviewers invited to the panels, new reviewer recruitment sustained similar levels for standing study section, slightly increased in small business panels, and slightly decreased in fellowship panels (see Table 4).

Table 4. New Reviewer Recruitment (n=119)	Reviewers with 0 prior meetings		Reviewers with 1-5 prior meetings	
	In Person	Zoom	In Person	Zoom
Standing Study Sections (ad hoc only)	3.5	3.5	5.0	5.1
Small Business	3.3	3.8	8.8	10.7
Fellowships	1.8	0.7	7.6	7.3

Reviewer overuse

In the sample, 85.7% of panels in the in-person round and 89.9% of panels in the Zoom round had no instances of reviewer overuse. The use of ad hoc reviewers with over 45 meetings in the last 12 years decreased from the in-person rounds to the most recent Zoom round, from 24 instances to 13 instances. For example, in standing study sections, ad hoc reviewers with over 45 prior meetings represented 1.4% of panels during the in-person rounds (19 reviewers), and only 0.5% of panels in the most recent Zoom round (8 reviewers). The divisions with the most frequent use of reviewers with over 45 prior meetings are DNDA (5 in Zoom round), DTCS (4 in Zoom round), and DPPS (3 in Zoom round). The more egregious instances of reviewer overuse occurred in the previous in-person rounds (i.e. a reviewer with service at 113 prior meetings). Thirteen reviewers in the most recent Zoom round had very high service levels of 46-70 meetings in the last 12 years.

Conclusions

We believe this analysis indicates that the quality of review has been sustained in the transition from in-person review to Zoom, although there may be some concern over the longer length of meetings and shorter break times over Zoom.

Meeting Duration

- Length of introductions fluctuated slightly when assessed by meeting type and size. Some of the longer introduction times can be attributed to larger panels where it takes longer for members to make initial introductions. High variability was evident in introduction times when assessed solely by meeting size; medium-sized meetings had the largest amount of variability which could in part be due to a larger number of meetings in this category.
- Overall, average active discussion times for time per application did not vary by meeting size and meeting type (19-21 minutes), although slightly longer than the target 15 minutes per application. Generally, it seems like meeting size does not highly impact the average discussion time per application as all meetings sizes had extremely close discussion times.
    - There was low variability, standard deviations between 2.3-3.3, for discussion times for all meeting sizes (large, medium, and small.)
    - Average active discussion times for 3-day meetings (n=3), 23.5 minutes, varied slightly from the standard 1- and 2-day meetings.
- Of the sample of small business panels, fellowship, and standing study sections, the majority of meetings started at 9:00-9:30 AM EST (67 meetings) and at 10:00-10:30 AM EST (33 meetings); on day 1 the majority of these meetings ended between 6-6:59 PM EST with three meetings ending between 8-8:30 PM EST.
    - On average for 2-day meetings, meeting duration for the first day for large meetings ran 11.64 hours, for medium meetings ran 11.02 hours, and for small meetings ran 10.3 hours. On average, meeting duration for the second day ran 8.14 hours for large meetings, 6.01 hours for medium-sized meetings and 3.7 hours for small meetings.
    - Total meeting duration across all days did not vary significantly between large and medium-sized meetings but did vary slightly for small meetings; there were slight variations of meeting duration when analyzed by both meeting type and meeting size.
    - The majority of 2-day meetings were medium-sized meetings (71) with 1 medium-sized meeting being a 3-day meeting.
- Average total break times by meeting size and type varied significantly between large/medium and small meetings and varied slightly between large and medium meetings; break times for small meeting member conflict SEPs and PAR and RFA SEPs are significantly lower due to smaller meeting application counts.
    - There were 5 meetings that did not take any breaks during the meeting – 2 medium-sized study sections, 2 fellowships (one medium and one small), and 1 small PAR/RFA.
    - Through the data it can be seen that break times for Zoom meetings are shorter, in rare cases non-existent, than the perceived standard total length of breaks provided at in-person meetings; it seems as if meetings tend to be running longer with potentially not enough breaks and increased screen time.

Meeting Application Counts, Roster Sizes, and Reviewer Workloads

- While negligible trends in meeting application counts are accounted for across the entire sample (as well as the subsample of standing study sections), there were a few notable trends in small business and fellowship panels, albeit using a small sample size. Meeting application counts generally increased amongst small business panels in the sample and decreased in fellowship panels in the sample.
    - Fluctuations in meeting application counts may affect trends seen in roster composition.
- Fluctuations in roster size between in-person rounds and Zoom rounds can mostly be accounted for by fluctuations in meeting application counts. Fluctuations in roster size may also, to a degree, be reflective of reviewer availability related to the circumstances around the pandemic or the transition to the Zoom format.

o   Reviewer workloads slightly decreased from in-person to Zoom rounds across all meeting types and sizes, although by a very negligible amount and generally remained in the appropriate range of reviewer workloads.

Roster Composition

Demographics

- Limited changes were seen in the academic rank of reviewers between in-person and Zoom rounds beyond normal round to round fluctuations, albeit there was a slight increase in assistant professor participation from the in-person to Zoom rounds in small business panels and fellowship panels.
- There were no notable trends in female and minority representation between in-person and Zoom rounds.
  - o   76.5% of panels in the sample did not have a net gain or loss of minority representation greater than 10%.

Standing Members and Ad hoc Participation

- For standing study sections, standing member participation did not fluctuate much between in-person rounds and Zoom rounds – 88% of standing study section meetings in the sample did not have a net gain or loss greater than 3 standing members.
- Typically, major fluctuations in ad hoc reviewers on the meeting level reflect fluctuations in applications loads in each meeting.

New Reviewer Recruitment/Reviewer Overuse

- There were slight increases in the average representation of ad hoc reviewers who reviewed regularly (13-24 reviewers) on panels and slight decreases in the average representation of ad hoc reviewers who were newly reviewing (0 reviews or 1-5 reviews) in relation to the entire panel.
- In the sample, 85.7% of panels in the in-person round and 89.9% of panels in the Zoom round had no instances of reviewer overuse.
  - o   Instances of reviewer overuse (reviewers with over 45 meetings in the last 12 years) have decreased from 24 instances in the in-person round to 13 instances in the most recent Zoom round.
  - o   The range of reviewer overuse for the most recent Zoom round was 46-70 reviews amongst 13 reviewers.

Limitations

The following limitations should be considered when interpreting the results of this analysis:

- Fluctuations in meeting application counts between rounds were not always explicitly weighted in the comparative analyses. Meeting size categorizations for each meeting were based on the size of the meeting in the Zoom rounds (2020/08 or 2020/10) and applied to the matched in-person round, regardless of whether the individual meeting application counts fluctuated beyond the meeting size parameters.
- For the comparative analyses, fellowships and small business panels each only represented 12 panels each, as opposed to 95 panels in the standing study section sample. Trends may be amplified for these meetings as individual changes in meetings may deceptively report a greater magnitude of change.
- Analyses are only as good as the data quality in QVR/CMM – for example, there were two standing study section meetings where the entire panel was designated as ad hoc members on the roster during the Zoom rounds. For these two meetings, the roster was compared to committee information in CMM to correct for standing members who were miscoded as ad hoc members. There may be other rosters that were not coded correctly.
- There may be data quality differences in meeting duration, depending on how detailed the ESA was at filling out the data collection forms. Some ESAs may have reported down to the exact minute, while others may have

rounded up. Data collection forms that were filled out in real time are probably more accurate than data collection forms that were filled out days later by memory or by referencing RTMS. ESAs who had not turned in their data collection forms two days following the meeting were prompted to submit them to help limit effects of data recall.

- ECRs were not separated out, so all roster analyses metrics include their metrics as well, including prior review service, where they will disproportionally represent reviewers who have had 0 prior meetings and are "newly engaged in the review system".

Scientific Review Officer Survey - Experience with Zoom for Review

Introduction

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 of 2020 CSR shifted most review meetings online using the Zoom.gov video meeting platform. The purpose of this survey is to assess SRO's observations on the quality of the review, meeting management, and on recruiting reviewers, compared to the normal face-to-face (F2F) review meetings.

Methods

Participants

The survey was administered to 237 scientific review officers (SROs) who led 433 CSR study section meetings between May 14th to September 14th, 2020; surveys were sent for each review meeting and so some SROs had the opportunity to respond more than one time. One hundred and sixty-two SROs responded, with 67 SROs responding to more than one survey. Fifty-seven percent of the surveys (n = 247 of 433) were returned.

Survey Administration

SROs were asked for their participation in a survey via email on August 14th, 2020 with a reminder email sent on August 20th. The email contained a weblink to the survey. SROs were told in the email that their responses would be kept confidential. All surveys returned by August 31st were included for analysis.

Measures

Peer Review Quality

Four survey items asked participants to rate on a scale from 1 (much better) to 5 (much worse) how their Summer 2020 review meeting compared to their normal F2F review meetings for the following items: 1) quality of the discussions ability to identify scientific merit, 2) level of reviewer engagement, 3) meeting management, and 4) overall quality of review.

Discussions and Engagement

Two survey items asked participants to list 1) strategies or techniques used to help manage the discussion at their review meeting and 2) strategies used to increase reviewer engagement during their meeting.

Reviewer Visual Presence

One survey item asked SROs approximately what percent of reviewers had their video turned off throughout the meeting. Response options included: a) 0, b) 1-10, c) 11-20, d) 21-30, e) 31-50, and f) more than 50%

Length of Meeting

One survey item asked SROs to rate how long their Summer 2020 review meeting lasted compared to their normal review meetings. Response options were on a 5-pt scale ranging from 1) much longer to 5) much shorter.

Reviewer Diversity

Four survey items asked SROs to rate on a scale from 1 (much more) to 5 (much less) how their Summer 2020 review meeting compared to their normal F2F review meetings in the following items related to diversity: 1) geographic, 2) gender, 3) race, and 4) career stage.

Recruitment

SROs were asked five survey items related to recruitment. Three survey items asked them to rate on a scale from 1 (much better) to 5 (much worse) how their Summer 2020 review meeting compared to their normal F2F review meetings for the following items related to recruitment: 1) ease in recruiting women, 2) ease in recruiting clinicians, and 3) ease in recruiting minorities. Two survey items asked SROs about their recruitment strategies and experience: 1) list any new strategies or flexibilities used to help recruit reviewers for their Summer 2020 or Fall review meeting, and 2) share any comments (positive or negative) about their experience or general thoughts on recruitment for their Summer 2020 review meeting.

Format Preference

One survey item asked SROs if there were no or minimal health risks from COVID-19, which meeting format would they prefer. Response options included: a) face-to face, b) Zoom, c) hybrid (some people in the room, some on video), d) other, and e) no preference.

Experience with Video Assisted Meetings (VAM)

One survey item asked SROs how many video assisted review meetings they conducted between May 2019 and March 2020 (i.e. pre COVID-19). Response options included: a) 0, b) 1-2, c) 3-5, and d) 6 or more.

Technical Difficulties

One survey item asked SROs if they experienced any technical difficulties conducting their review meeting. Response options included: a) yes, b) no, and c) don't know

Meeting Logistics

The survey asked for meeting format used to hold Summer 2020 review meetings (Zoom, video-assisted, virtual meeting, telephone assisted, or other). The survey also asked for the number of days the meeting lasted and had a blank to enter the number of applications discussed.

Lessons Learned

In an open text box, participants were asked what they would do differently for their next review meeting.

Reviewers

The number of reviewers present in each study section was gathered from IMPAC II.

Results

- Sent 433 surveys via email, 57% completed surveys received (n = 247)
- 230 used Zoom (17 did not use Zoom; i.e. telephone assisted and virtual meetings)
- 24% of study sections experienced technical difficulties (n = 56)

Table 1. Number and Type of Study Section for Survey Respondents

Study Section Type	Number of Study Sections	Number of Non-Zoom Study Sections	Number with Technical Difficulties
Chartered	115	1	31
Member Conflict	33	7	7
Recurring Small Business	32	3	7
Recurring Fellowships	15	2	3
Other SEPs	52	4	8
Total	247	17	56

The following results were derived from surveys concerning Zoom meetings only (n = 230).

Quality of Review

- For all types of study sections and for all measures of quality, the study section meetings were rated about the same as F2F meetings, usually by 80% or more of SROs.

- A small group, ~15%, rated Zoom meetings as better than their normal meetings overall, and said discussions were better and reviewers more engaged.

Figure 1. SROs' Assessments of Peer Review Quality for All Meetings (compared to normal meetings; including special emphasis panels (SEP))
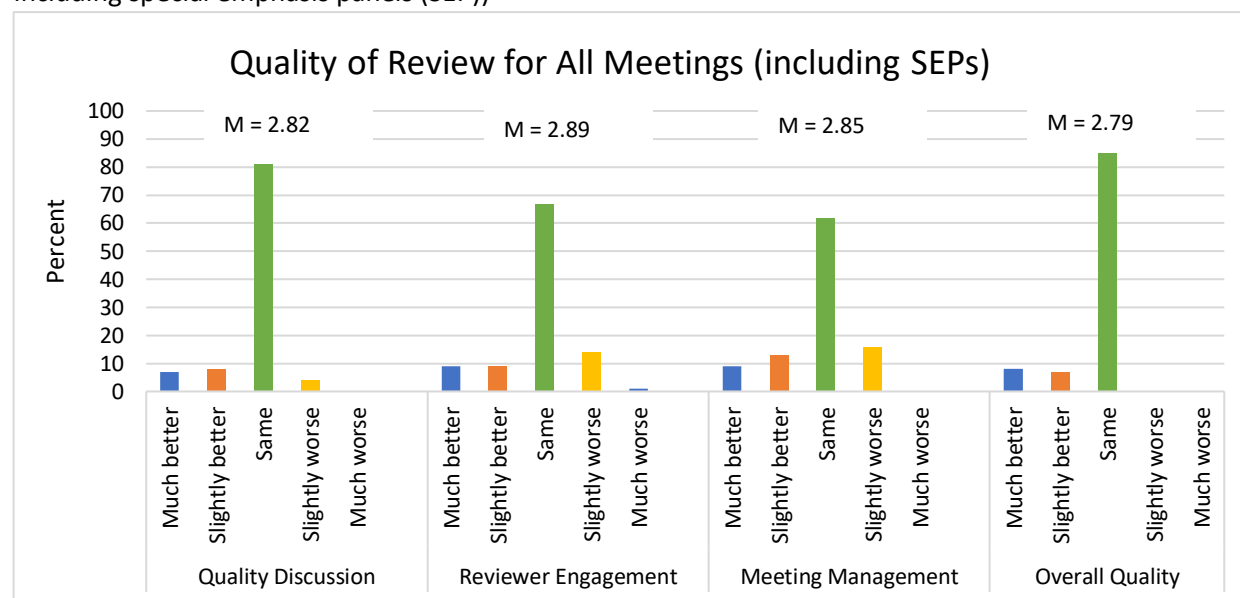
Figure 2. SROs' Assessments of Peer Review Quality
for Chartered Meetings
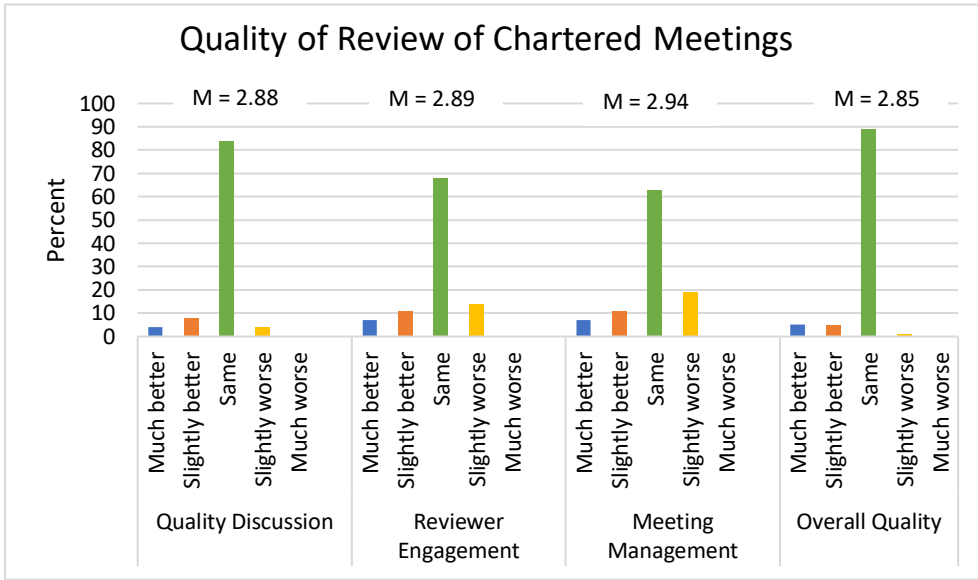
Figure 3. SROs' Assessments of Peer Review Quality
for Member Conflict Meetings
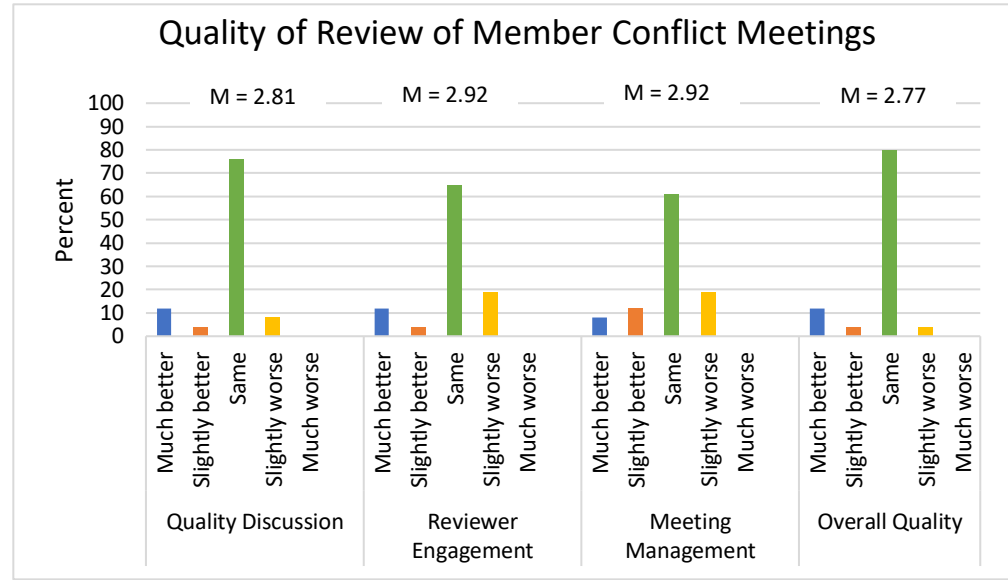
Figure 4. SROs' Assessments of Peer Review Quality
for Recurring Small Business Meetings
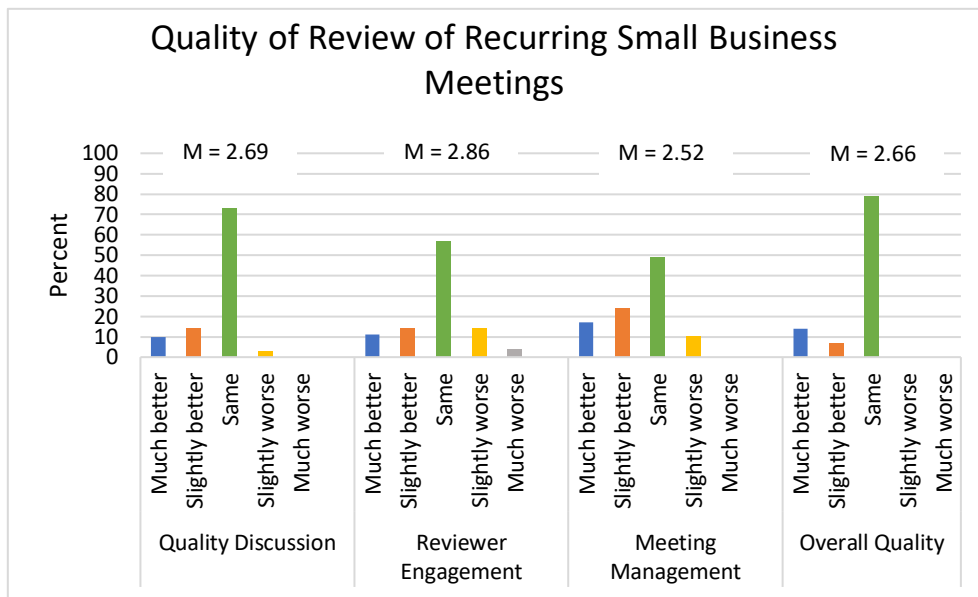
Figure 5. SROs' Assessments of Peer Review Quality
for Recurring Fellowships Meetings
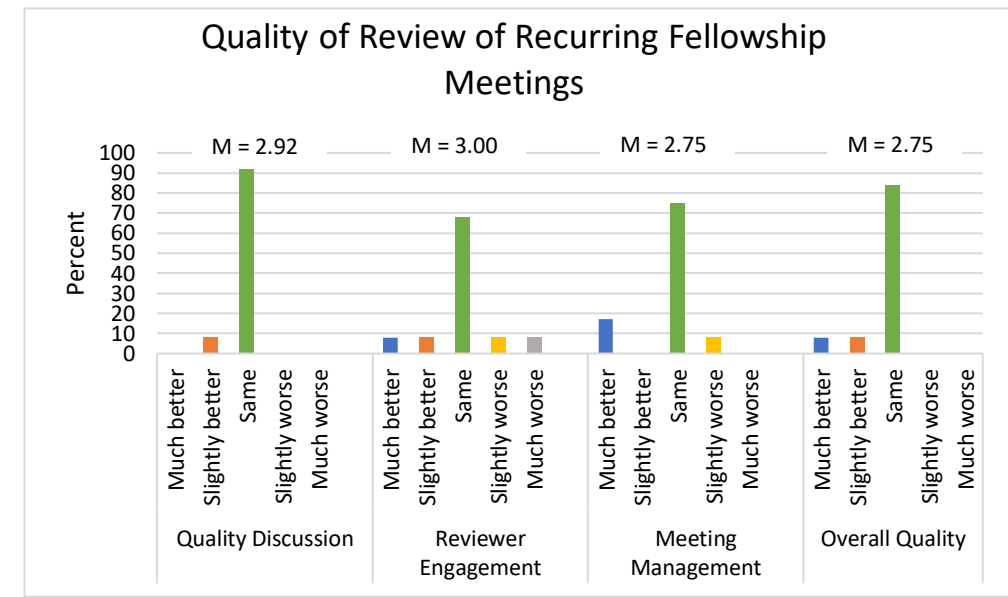
Meeting Format Preference Among SROs

- Figure 6 displays data for SROs meeting format preferences.
- SROs overall prefer F2F meetings, but there are significant differences according to meeting type:
  - SROs of chartered study sections prefer F2F 52% to 26%,
  - SROs who lead fellowship review meetings split 46%-46%,
  - SROs for small business review meetings preferred Zoom 38% to 34% (and 28% preferred a hybrid meeting)
- Note that the sample size for non-chartered meetings is modest.

Figure 6. Meeting Format Preference by Study Section Type
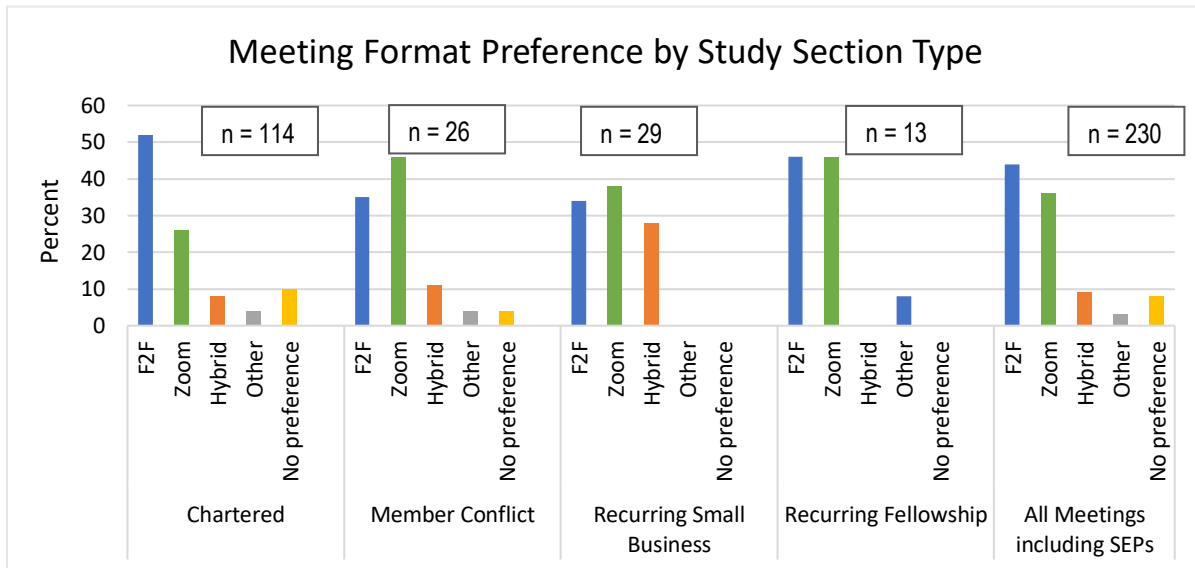
Length of Meeting

Across all study section types, the study section meetings were rated as lasting longer than their normal meetings 45% of the time, and as much longer in 12%.

Figure 7. Length of Meeting by Study Section Type (compared to normal meetings)
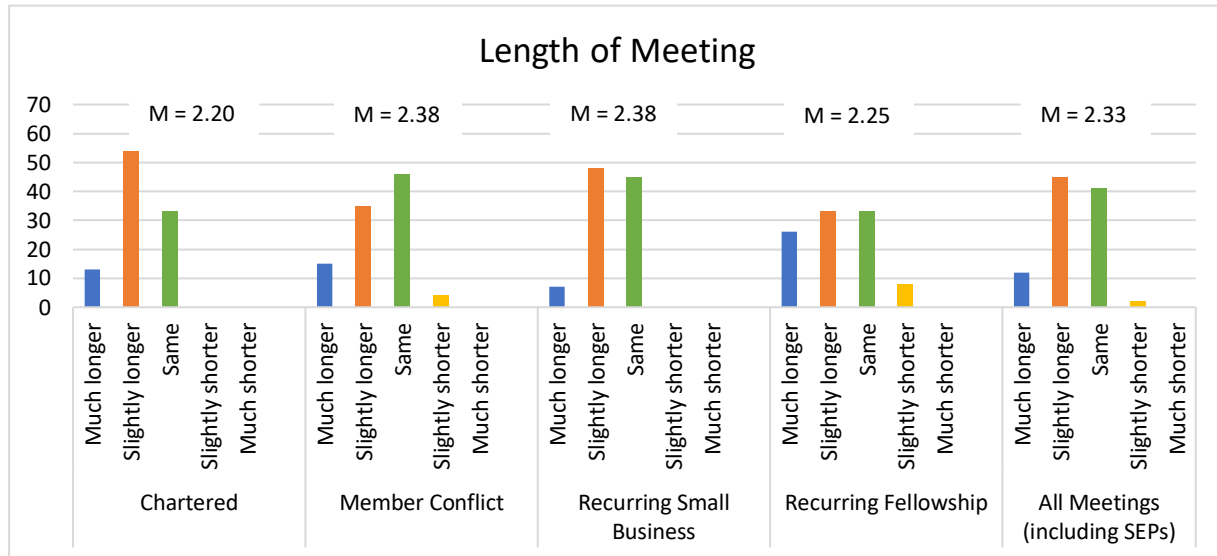
Number of reviewers and perceived Quality of Review

As shown in Table 2, small business meetings had the most reviewers, followed by chartered meetings. Member conflict meetings had the least number of reviewers, followed by other special emphasis panels (SEP).

Table 2. Mean Number of Reviewers by Study Section Type and All Study Sections

Study Section Type	M (SD)	Range
Chartered	31.43 (7.27)	15 -47
Member Conflict	15.62 (5.74)	7 – 30
Small Business	33.79 (8.95)	16 - 49
Fellowship	26.69 (10.34)	13 - 41
Other SEPs	17.02 (8.94)	7 - 46
All Study Sections	26.67 (10.61)	7- 49

- There were no significant relationships between the number of reviewers and SRO ratings of a) quality of the discussion, $r(228) = .113$, $p = .090$; b) reviewer engagement, $r(227) = .121$, $p = .070$; c) meeting management, $r(228) = .025$, $p = .709$ or d) overall quality of the review, $r(228) = .111$, $p = .093$.

Reviewers' visual presence and perceived Quality of Review

Table 3 shows that the percent of meeting participants who had their cameras off was generally small; 77% of meetings were rated as having less than 10% cameras off. Reviewers were more likely to turn their cameras off at small business and fellowship meetings.

Table 3. Percent of Reviewers with Video Off Throughout the Meeting by Study Section Type

Percent of reviewers	Chartered		Member Conflict		Small Business		Fellowship		Other SEPs		All Study Sections	
	n	%	n	%	n	%	n	%	n	%	n	%
0	48	42	15	58	6	21	5	39	34	71	108	47
1-10	46	40	7	27	6	21	3	23	8	17	70	30
11-20	7	6	0	0	7	24	3	23	1	2	18	8
21-30	7	6	0	0	8	28	1	8	4	8	20	9
31-50	3	3	0	0	1	3	0	0	1	2	5	2
More than 50	3	3	4	15	1	3	1	8	0	0	9	4

There were weak relationships between reviewers' visual presence at the meeting and SRO ratings of a) quality of the discussion (compared to F2F), $r(228) = .15$, $p = .024$; b) reviewer engagement (compared to F2F), $r(227) = .19$, $p = .004$, and c) overall quality of the review (compared to F2F), $r(228) = .17$, $p = .011$. The more reviewers with their videos off, the lower the relative quality of the discussion, relative reviewer engagement, and relative overall quality of the review.

Reviewer Diversity

Overwhelmingly (~80%), across the board—for all types of study sections and for all areas of diversity, the demographic diversity of study section meetings was rated as being the same as F2F (See Figures 8-12). There was no item that raised concerns about loss of demographic diversity.

Figure 8. Reviewer Diversity at All Study Section Meetings (compared to normal meetings)
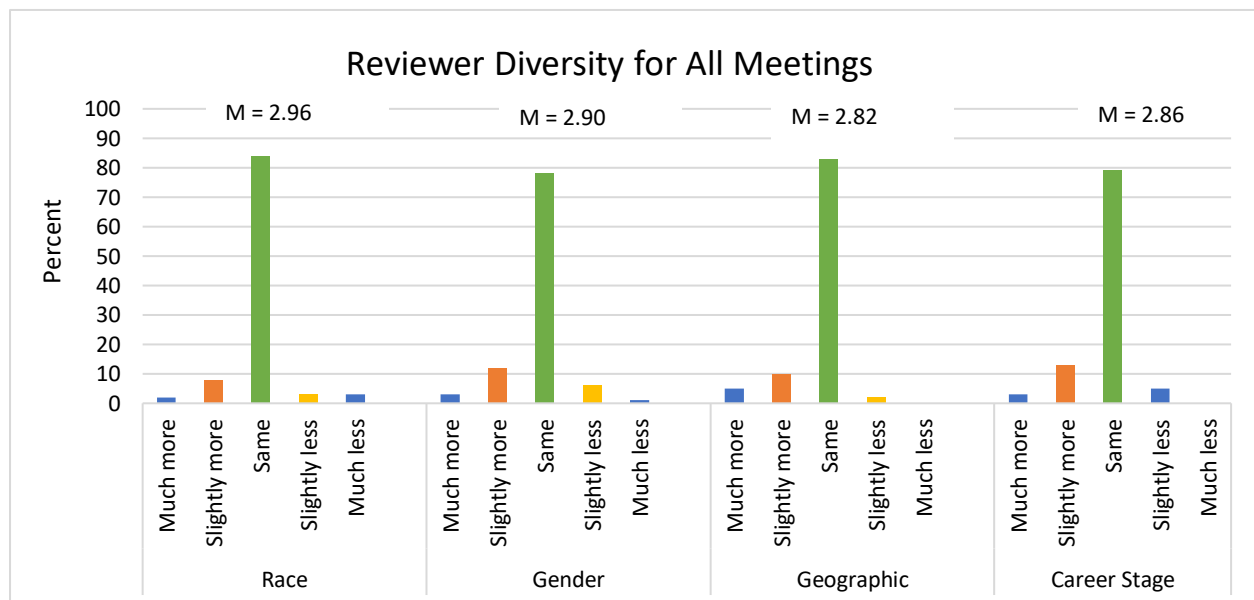
Figure 9. Reviewer Diversity at Chartered Meetings
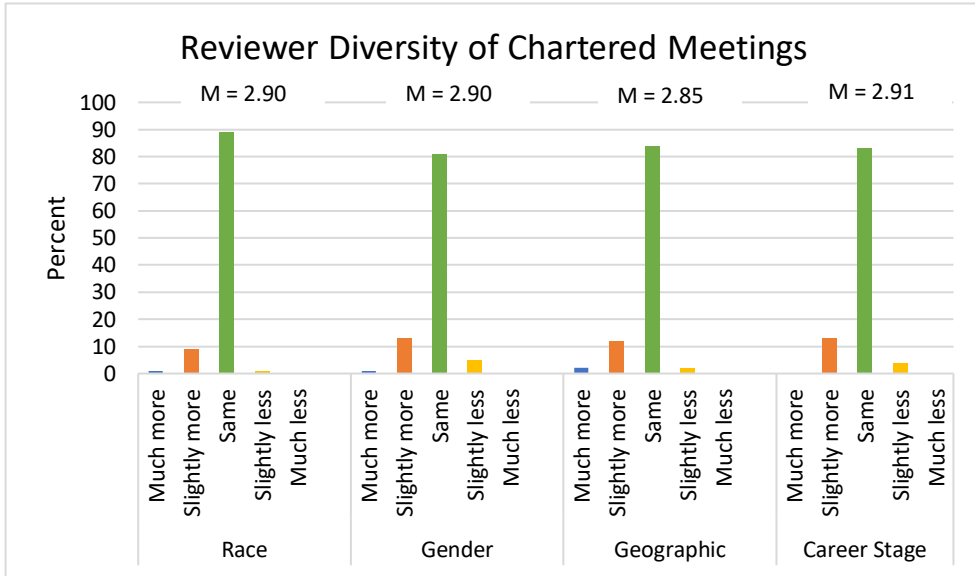
Figure 10. Reviewer Diversity at Member Conflict Meetings
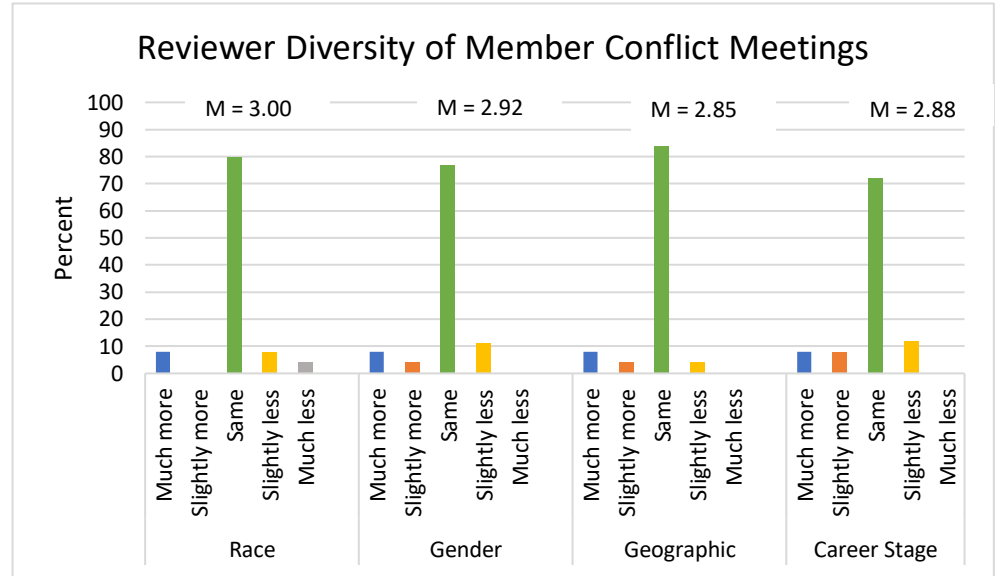
Figure 11. Reviewer Diversity at Recurring Small Business Meetings
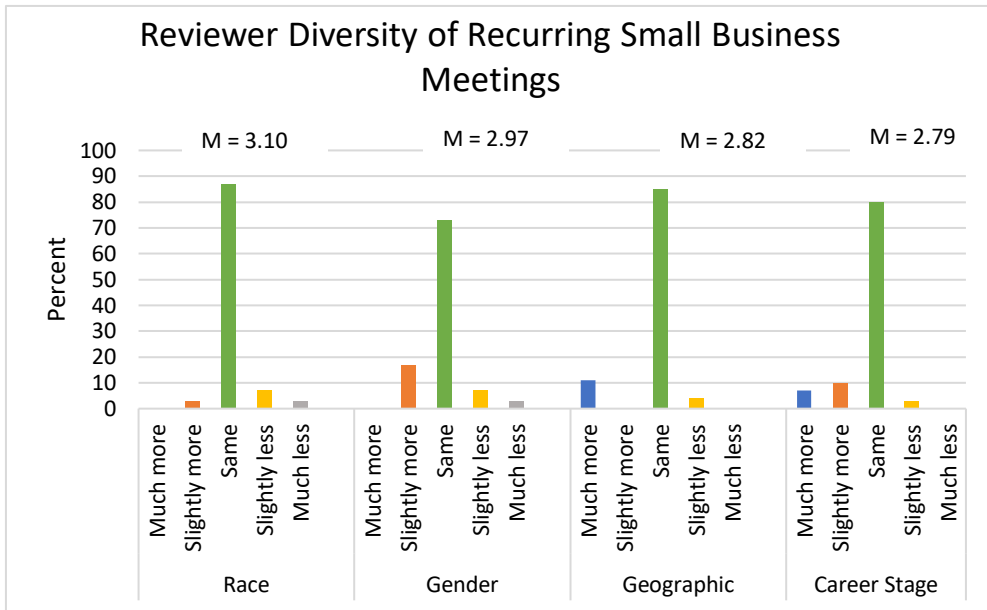
Figure 12. Reviewer Diversity at Recurring Fellowship Meetings
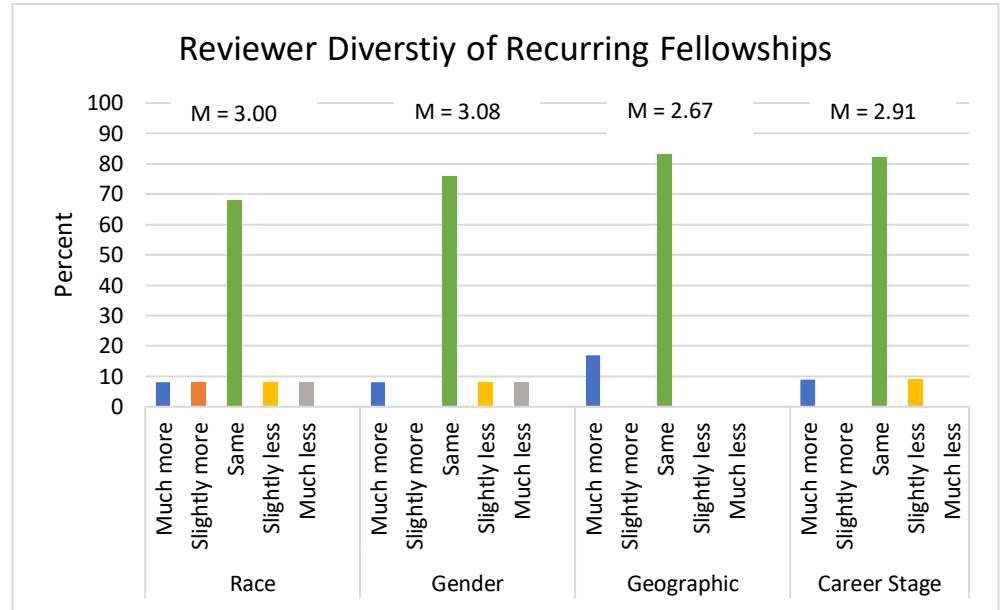
Ease in Recruiting

SROs responses regarding ease of recruiting specific types of reviewers are tabulated in Figures 13-17. A plurality of SROs rated ease of recruiting women, minorities, and clinicians the same as F2F. Among SROs who said recruitment was different for this meeting, slightly more said it was easier rather than harder to recruit minorities and women; slightly more said it was harder rather than easier to recruit clinicians.

Figure 13. Ease in Recruiting Diverse Reviewers for All Study Section Meetings (compared to normal meetings)
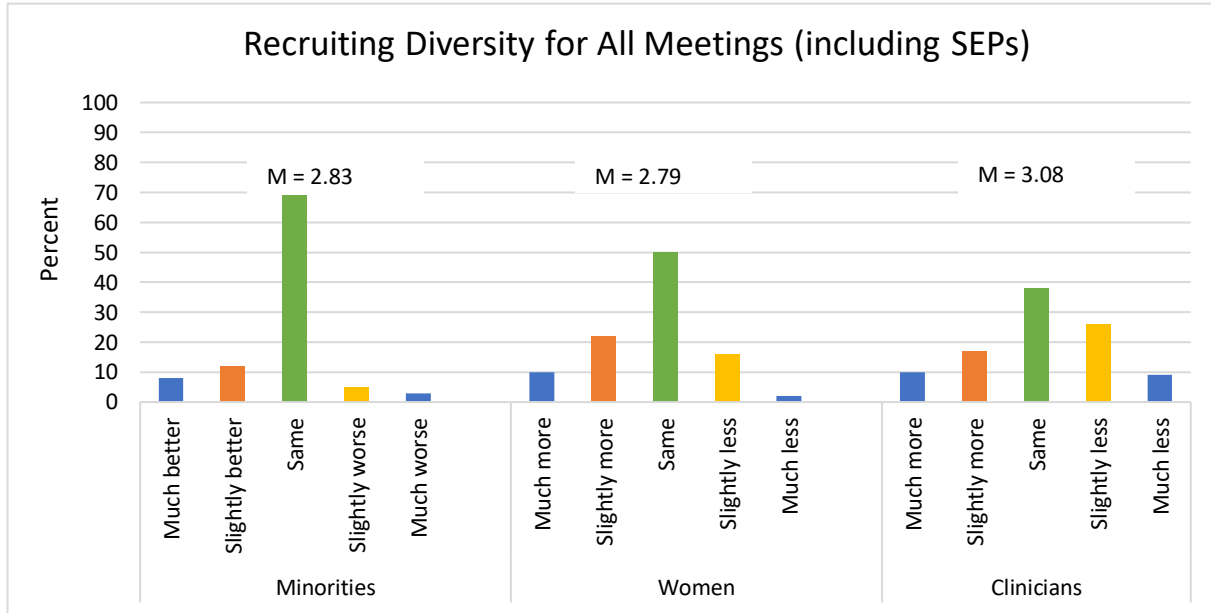
Figure 14. Ease in Recruiting Diverse Reviewers
for Chartered Meetings

Recruiting Diversity at Chartered Meetings
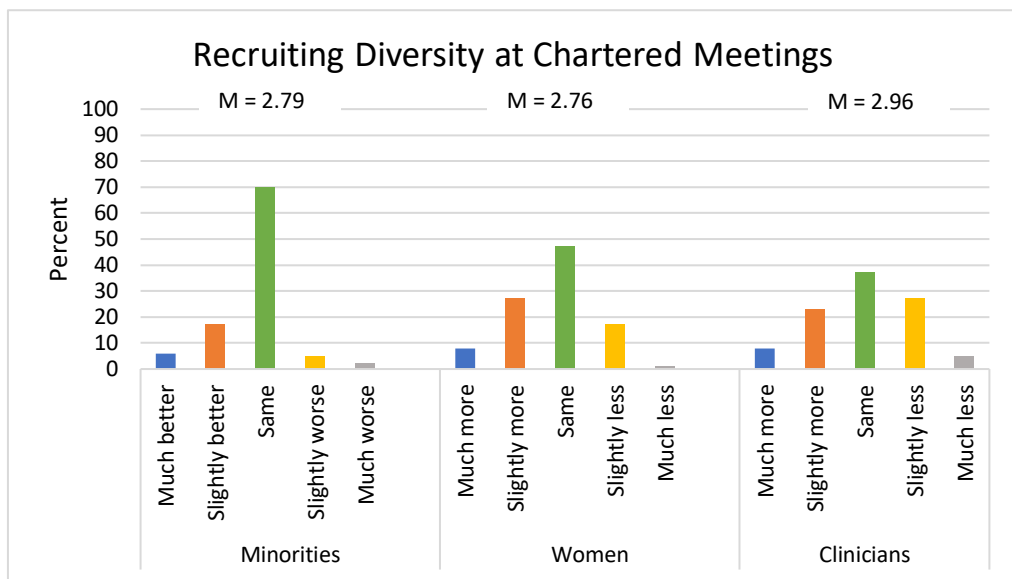
Figure 15. Ease in Recruiting Diverse Reviewers
for Member Conflict Meetings

Recruiting Diversity at Member Conflict Meetings
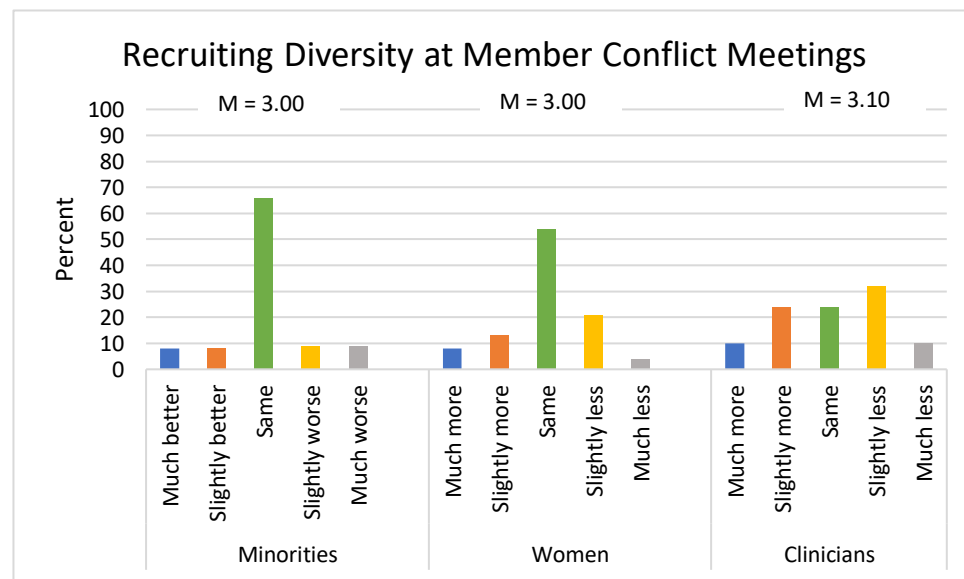
Figure 16. Ease in Recruiting Diverse Reviewers
for Recurring Small Business Meetings

Recruiting Diversity at Recurring Small Business Meetings
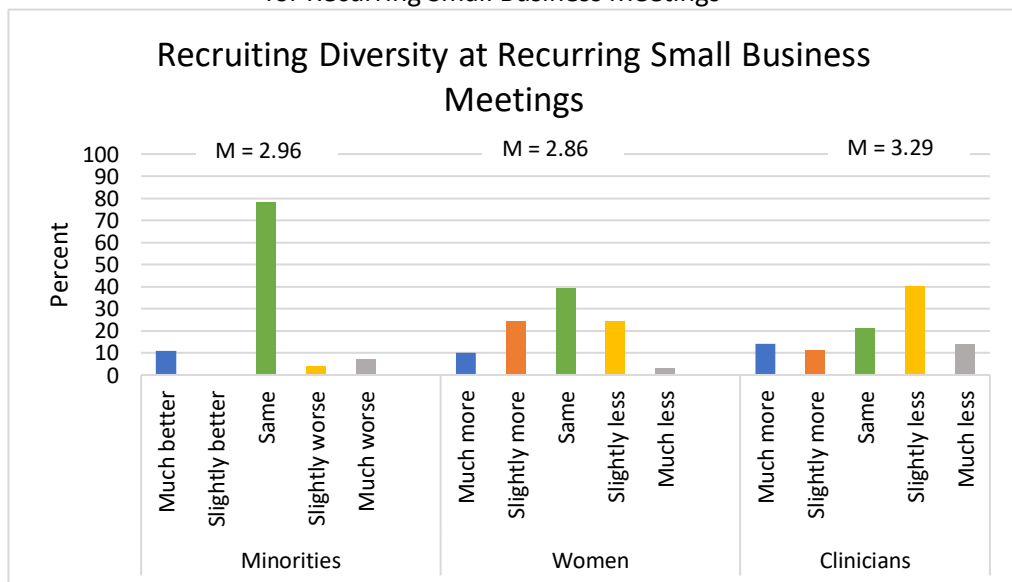
Figure 17. Ease in Recruiting Diverse Reviewers
for Recurring Fellowship Meetings

Recruiting Diverstiy at Recurring Fellowship Meetings
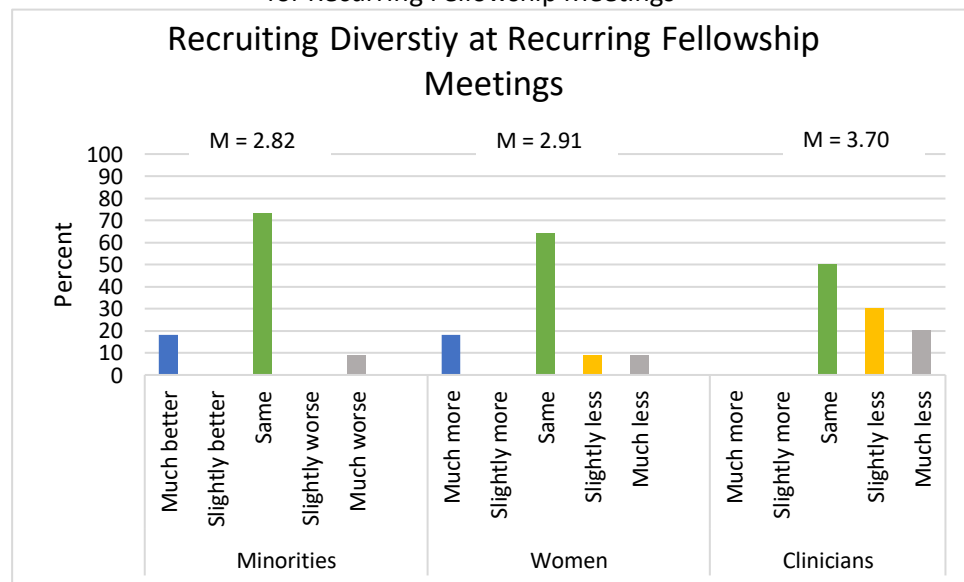
Technical Difficulties

The following results are for those study sections who used Zoom and experienced technical difficulties (n = 56).

Table 4. Technical Difficulties by Study Section Type

Study Section Type	Number of Study Sections
Chartered	31 (55%)
Member Conflict	7 (13%)
Recurring Small Business	7 (13%)
Recurring Fellowships	3 (5%)
Other SEPs	8 (14%)
Total	56

Quality of the Review Among Study Sections with Technical Difficulties

The quality of the review for study sections that experienced technical difficulties was slightly worse than those that did not experience technical difficulties—for all measures of quality (for all measures averaged: M = 2.92 for technical difficulties, M = 2.77 for non-technical difficulties).

Figure 18. Quality of the Review with Technical Difficulties (compared to normal meetings)
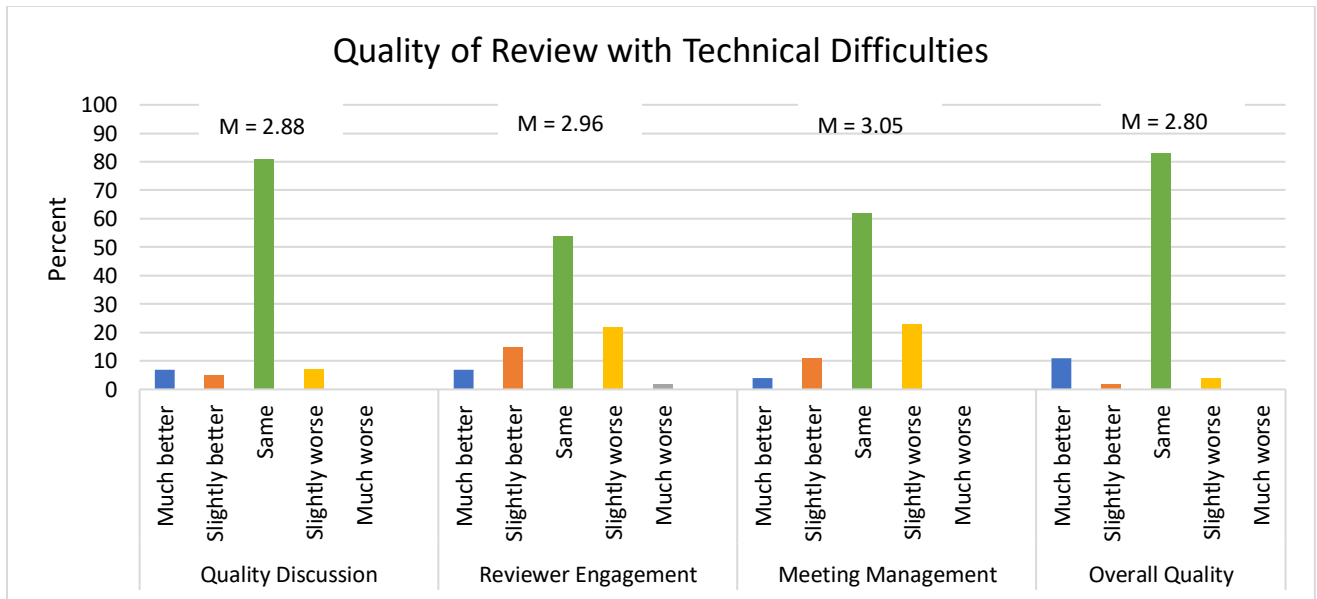
Extramural Support Assistants Survey - Experience with Zoom for Review

Introduction

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 of 2020 CSR shifted most review meetings online using the Zoom.gov video meeting platform. The purpose of this survey was to help CSR understand the impact that COVID-19 had on the work duties of staff who support CSRs review meetings and how to better prepare for the upcoming rounds to ensure support staff have all the resources they need to successfully perform their job.

Methods

Participants

Extramural Support Staff (ESAs) who participated in CSR study section meetings (n = 73) between May 14th to September 14th, 2020.

Survey Administration

ESAs were asked for their participation in a survey via email August 14th, 2020, with three reminder emails sent. The email contained a weblink to the survey. ESAs were told in the email that their responses would be kept confidential and that the survey would take about ten minutes to complete. All surveys returned by September 9th were included for analysis.

Measures

Meeting Tasks

Six survey items asked ESAs to rate on a scale from 1 (much easier) to 5 (much harder) how their Summer 2020 review meetings compared to their normal review meetings (i.e. pre-COVID) for the following items: 1) creating rosters and verifying accuracy, 2) dealing with last minute changes or requests, 3) handling conflicts of interest during the meeting, 4) trouble-shooting technical issues during the meeting, 5) ensuring post-meeting conflict of interest forms were signed, and 6) solving problems that arose.

Length of Meeting

One survey item asked ESAs to rate on average, how much longer or shorter their Summer 2020 review meetings lasted compared to their normal pre-COVID meetings. Response options were on a 5-pt scale ranging from 1) much longer to 5) much shorter.

Additional Support Time

One survey item asked ESAs to rate on average, how much more time was spent supporting each of their Summer 2020 meetings (i.e. before, during, and after the meeting) compared to the meetings they

supported pre-COVID meetings. Response options were a) 0 b) less than 1 hour, c) 1 hour, d) 2 hours, e) 3 hours, and f) more than 3 hours.

Time-consuming Activities

One open-ended survey item asked ESAs to list which activities or duties demanded more of their time, if they spent more time than usual with their Summer 2020 meetings.

Troubleshooting

One survey item asked participants approximately how much time they spent troubleshooting technical issues with reviewers during their meetings. Response options included a) <30 mins. b) 30 mins - 1 hour, c) 1-2 hours, d) 2-3 hours, e) 3-4 hours, f) more than 4 hours.

Technical Issues

Two survey items asked ESAs to list or describe 1) technical or logistical issues experienced before or during the review meeting, and 2) major technical issues reviewers faced with Zoom (if they tested Zoom with reviewers before the meeting).

Number of Video Assisted Meetings (VAM)

One survey item asked ESAs how many video assisted review meetings they supported between May 2019 and March 2020 (i.e. pre COVID-19). Response options included: a) 0, b) 1-2, c) 3-5, and d) 6 or more.

Format

One survey question asked ESAs which was the most commonly used format for the meetings they supported this summer. Response options included: Zoom, virtual meeting (VM), telephone-assisted meeting (TAM), and other.

Zoom

Two open-ended survey items asked ESAs about work efficiency using the Zoom platform: 1) What changes can VAM (Virtual Assisted Meeting) personnel make to the Zoom platform to help you prepare for and manage a more efficient and effective meeting, and 2) What can CSR do to help you manage review meetings using the Zoom platform.

Results

- Seventy-three (n = 73) ESAs completed the survey. The survey was sent to the Extramural Support Assistant distribution list (97 people) and a follow up message was sent to all of CSR.
- Zoom was the most commonly used format for all ESAs
- Overwhelmingly, ESAs thought support tasks were the same or easier to manage in Zoom.
    - 54% said last minute changes were easier
    - 35% said handling COIs was easier, 62% said it was the same as F2F
    - 46% said technical troubleshooting was easier
    - 48% said solving problems in general was easier

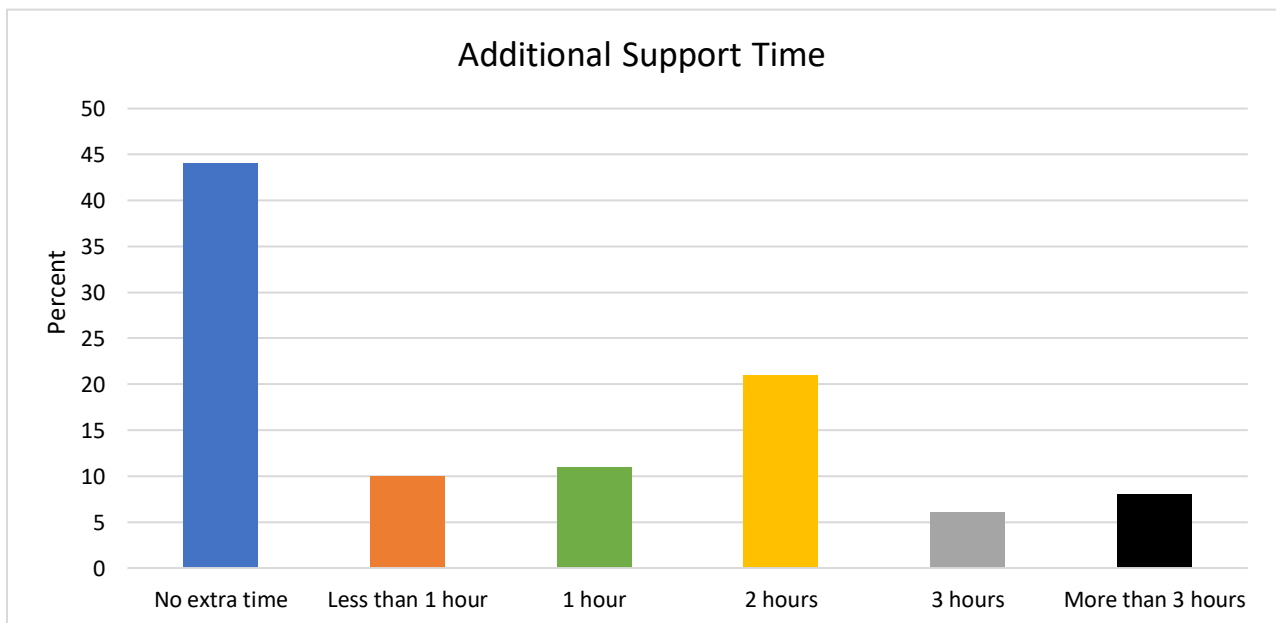**Change in Effort for Zoom Meeting Tasks Compared to Normal Meetings**

Change in Effort for Meeting Tasks
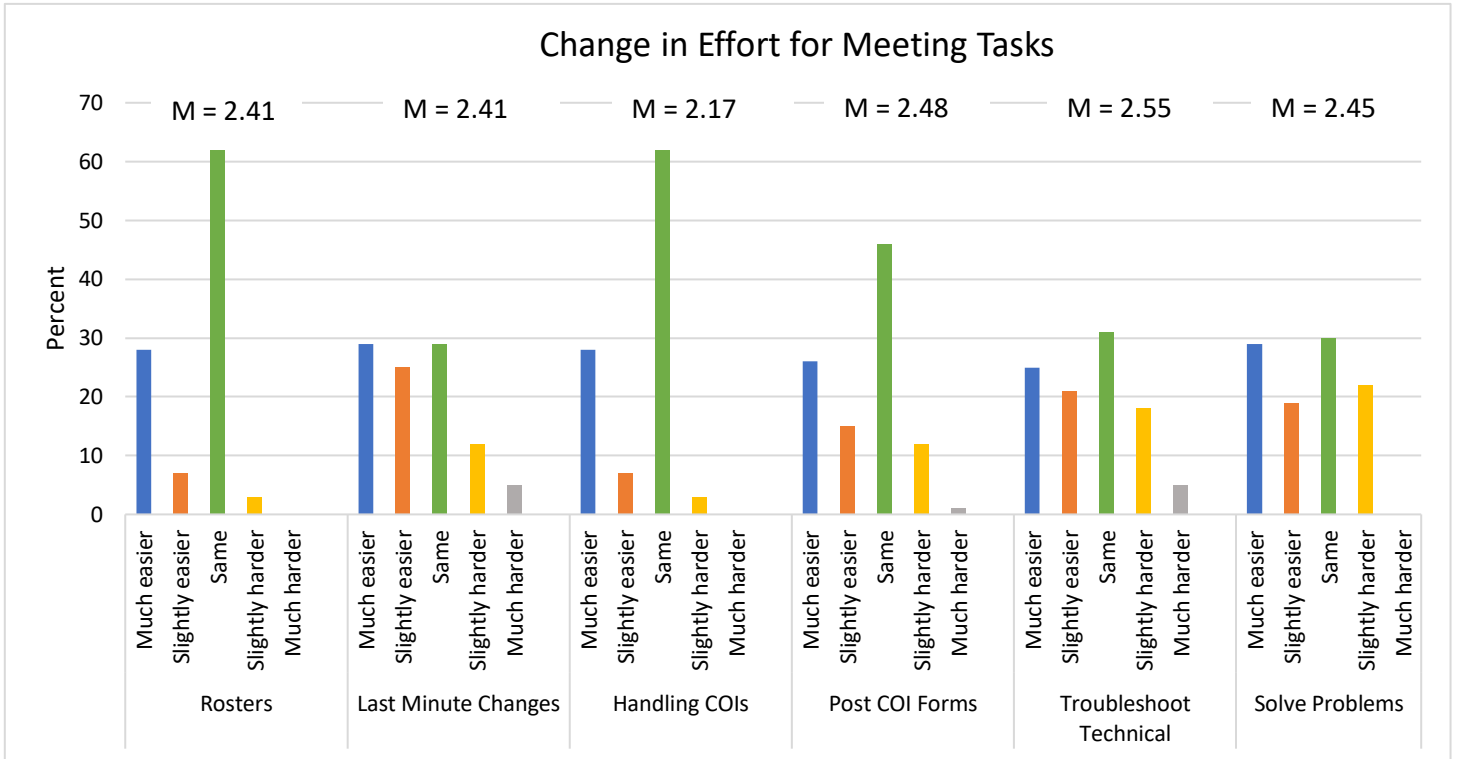
Length of Summer Meetings Compared to Normal Meetings (pre COVID-19)

Length of Meetings

M = 2.40

Response	Percent
Much longer	13
Slightly longer	43
Same	34
Slightly shorter	8
Much shorter	2
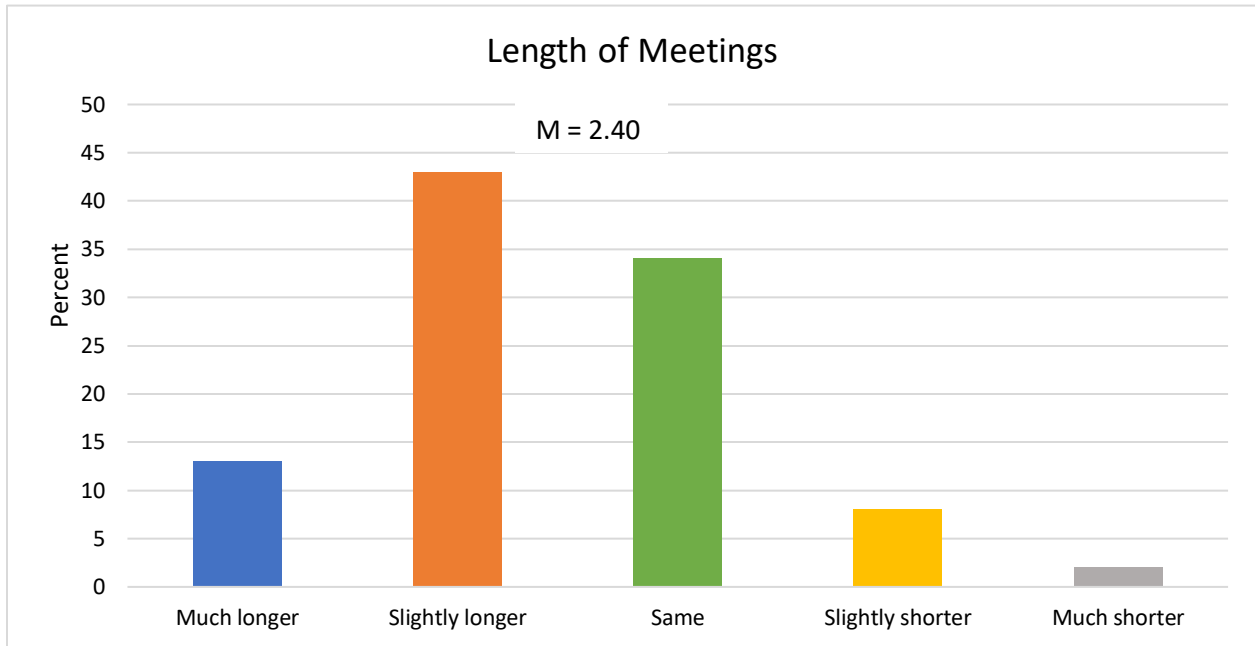
Additional Time Needed for Support Tasks

- A majority said Zoom took no extra support time or less than one hour additional
- ~15% said it took 3 or more additional hours

Additional Support Time

Response	Percent
No extra time	44
Less than 1 hour	10
1 hour	11
2 hours	21
3 hours	6
More than 3 hours	8

Time Spent Troubleshooting Technical Issues During the Meeting

Sixty-three percent reported that solving technical problems required less than 30 min

Time Spent Troubleshooting Technical Issues During Meeting

Percent	Less than 30 mins.	30 mins - 1 hour	1 - 2 hours	2 - 3 hours	3 -4 hours	More than 4 hours
	63	30	4	3	0	0
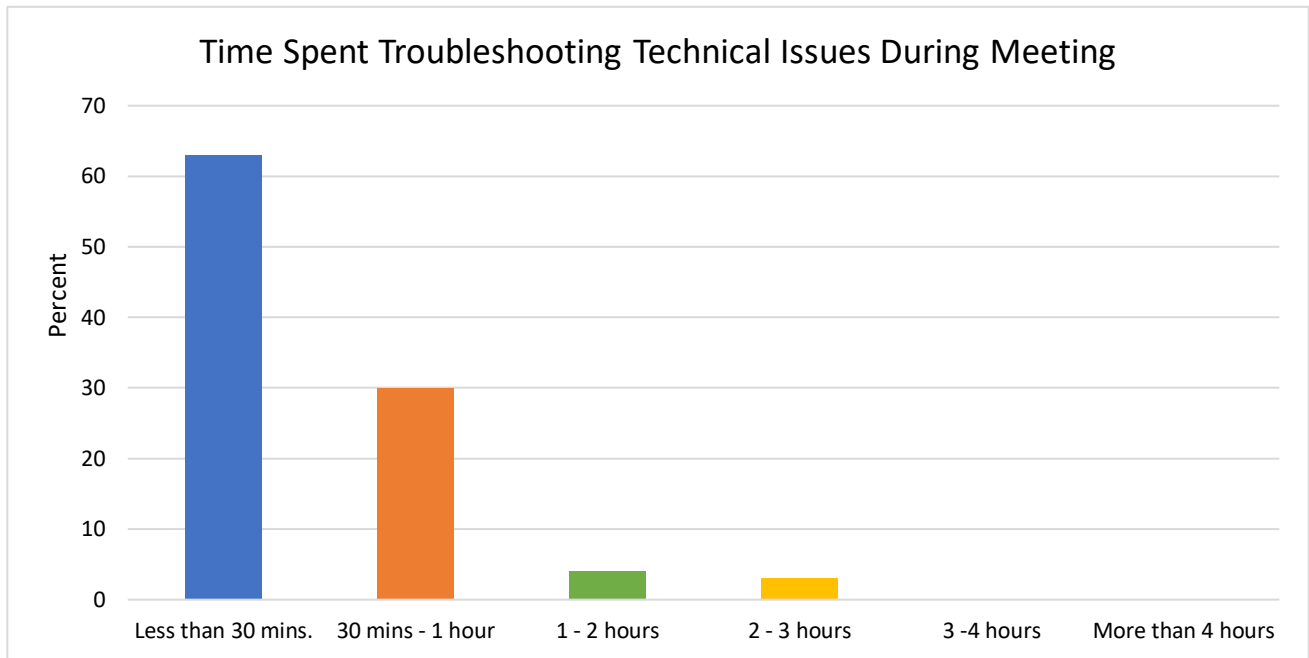
Prior Experience with Video Assisted Meetings (pre COVID-19)

Number of Video-Assisted Meetings

Percent	0	2-5	3-5	6 or more
	21	12	23	44
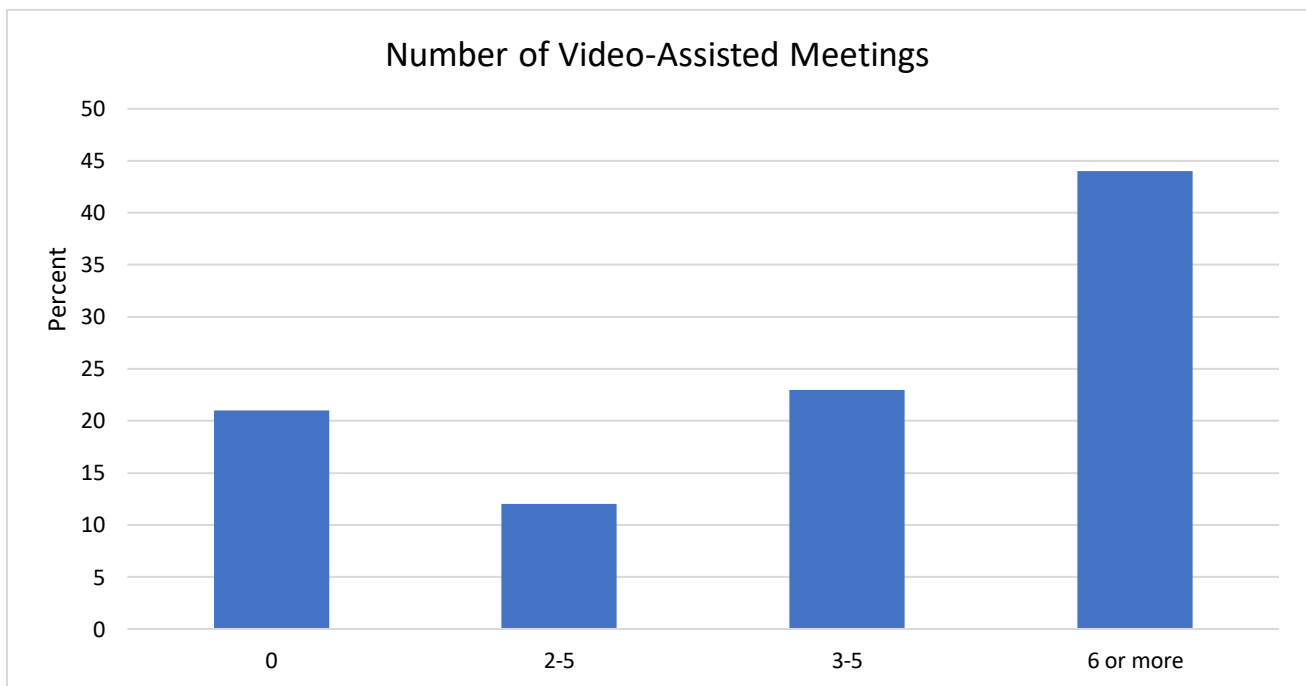
Conclusions

Moving to a virtual meeting format, using the Zoom platform, does not appear to have negatively impacted the work of CSR's extramural support assistants.

In terms of effort to support a meeting – the majority of respondents found all measures to be easier or the same as the effort normally required to support a meeting, when more review meetings are conducted in person. The majority of respondents found effort to be the same or less to create rosters, deal with last-minute changes, handle conflicts of interest, and to ensure that post-meeting conflict of interest forms were signed. Even the time needed to troubleshoot technical issues or, generally, problem solve was not increased for the majority of respondents supporting Zoom review meetings. Sixty-three percent reported that solving technical problems required less than 30 min and 30% reported that it required 30 min to 1 hour. Forty-four percent of ESAs found that support tasks took no extra time and 21% reported that support required 1 or fewer additional hours. A small percent reported a large increase in time required of more than 3 hours additional work (8%). The results are not surprising in light of staff familiarity with alternative review formats; only 21% of respondents had not supported a video-assisted meeting in the past and 44% had supported 6 or more video-assisted meetings before this summer.

Zoom review meetings do appear to be longer than in-person meetings. 66% of ESA reported that Zoom review meetings this summer were either much longer or slightly longer than typical.

As part of our continuous evaluation efforts, we would like to hear your opinions on your recent review meeting using the Zoom platform. Your identity will be kept private and only aggregate responses will be used in our reports.

1. For the following items, please select how your Zoom review meeting **compared** to your normal face-to-face meetings.

	Much better	Slightly better	About the same	Slightly worse	Much worse	Don't know/I did not use the Zoom platform
Overall quality of review	○	○	○	○	○	○
Productivity of the discussions	○	○	○	○	○	○
Level of reviewer engagement	○	○	○	○	○	○
Meeting management	○	○	○	○	○	○

2. Please select how your Zoom review meeting **compared** to your normal face-to-face meetings.

	Much more	Slightly more	About the same	Slightly less	Much less	Don't know/I did not use the Zoom platform
I contributed to the discussion.	○	○	○	○	○	○
I felt confident voicing my opinions.	○	○	○	○	○	○
I felt others were receptive and responsive to my feedback.	○	○	○	○	○	○
I was able to clearly communicate my opinions.	○	○	○	○	○	○
I felt comfortable voting outside the range.	○	○	○	○	○	○
My attention span at the meeting lasted.	○	○	○	○	○	○

3. Approximately how many face-to-face review meetings have you participated in for NIH?

- ○ 0
- ○ 1
- ○ 2-5
- ○ 6-15
- ○ 16+

4. Assuming no or minimal health risks from COVID-19, would you be more likely to participate in a review meeting if it was held face-to-face or over Zoom/video?

- ○ Face-to-face
- ○ Zoom/video
- ○ No preference

5. Did you experience any technical difficulties with your Zoom review meeting (e.g. audio, visual, connecting, etc.)?

- ○ Yes
- ○ No

○ I did not use Zoom for my review meeting

6. What stage of your career are you in?

○ Assistant Professor
○ Associate Professor
○ Professor
○ Other

7. What is your sex?

○ Male
○ Female
○ I prefer not to respond

8. Are you Hispanic or Latino?

○ Yes
○ No
○ I prefer not to respond

9. Which racial group do you primarily identify as? Select all that apply.

☐ American Indian or Alaska Native
☐ Asian
☐ Black or African-American
☐ Native Hawaiian or other Pacific Islander
☐ White
☐ I prefer not to respond

10. How was the overall quality of your virtual review meeting compared to your normal face-to-face meetings?

11. Please share any comments (positive or negative) about your experience or general thoughts on having your review meeting over Zoom?

Please click on the 'Next' button below to send your responses. Thank you for taking the survey.

Submit

**Center for
Scientific Review**

To help us understand the impact of COVID-19 on our review meetings, we would like to hear your opinions on your recent Summer 2020 review meeting. Your identity will be kept private and only aggregated responses will be used in our reports

1. For the following items, please select how your Summer 2020 review meeting **compared** to your normal review meetings (i.e. pre COVID-19).

	Much better	Slightly better	About the same	Slightly worse	Much worse	Don't know/NA
Quality of the discussions (ability to identify scientific merit)	○	○	○	○	○	○
Level of reviewer engagement	○	○	○	○	○	○
Meeting management	○	○	○	○	○	○
Overall quality of review	○	○	○	○	○	○
Ease in recruiting women	○	○	○	○	○	○
Ease in recruiting clinicians	○	○	○	○	○	○
Ease in recruiting minorities	○	○	○	○	○	○

2. Compared to your normal review meetings, how long did your Summer 2020 review meeting last?

○ Much longer
○ Slightly longer
○ About the same
○ Slightly shorter
○ Much shorter
○ Don't know/NA

3. Please list any strategies or techniques you used to help manage the discussion at your Summer 2020 review meeting (including pre-meeting preparations).

[text box]

[text box]

[text box]

4. What strategies, if any, did you use to increase reviewer engagement during your Summer 2020 review meeting (including pre-meeting preparations)?

[text box]

[text box]

[text box]

5. Please select how your Summer 2020 review meeting compared to your normal review meetings in the following areas related to diversity.

	Much more	Slightlty more	About the same	Slightly less	Much less	Don't know/NA
Geographic	○	○	○	○	○	○
Gender	○	○	○	○	○	○
Race	○	○	○	○	○	○
Career stage	○	○	○	○	○	○

6. What new strategies or flexibilities did you use to help recruit reviewers for your Summer 2020 or Fall 2020 review meeting?

[text box]

[text box]

[text box]

7. Please share any comments (positive or negative) about your experience or general thoughts on recruitment for your Summer 2020 review meeting.

[text box]

8. Which meeting format did you use to hold your Summer 2020 review meeting (i.e. post COVID-19)?

○ Zoom
○ Video Assisted (VAM; not Zoom)
○ Virtual Meeting (VM)
○ Telephone Assisted (TAM)
○ Other

9. How many days was your review meeting?

○ 1
○ 2
○ 3

10. How many applications did you discuss at your review meeting (please use numeric format)?

[text box]

11. Approximately what percent of reviewers had their video turned off throughout the meeting?

○ 0
○ 1-10%
○ 11-20%
○ 21-30%
○ 31-50%
○ More than 50%
○ I did not use video to conduct the meeting

12. Did you experience any technical difficulties conducting your review meeting (e.g. audio, visual, connecting,

etc.)?

    ○ Yes
    ○ No
    ○ Don't know/NA

13. How many video assisted (VAM) review meetings did you conduct between May 2019 and March 2020 (i.e. pre COVID-19)?

    ○ 0
    ○ 1-2
    ○ 3-5
    ○ 6 or more

14. Assuming no or minimal health risks from COVID-19, which meeting format would you prefer?

    ○ Face to Face
    ○ Zoom
    ○ Hybrid (some people in the room, some attending by video)
    ○ Other
    ○ No preference

15. What would you do differently for your next review meeting?

Submit

To help us understand the impact of COVID-19 on our review meetings, we would like to hear your opinions on your recent Summer 2020 review meetings. Your identity will be kept private and only aggregated responses will be used in our reports

1. Of the meetings you supported this summer, which was the most commonly used format?

- ○ Zoom
- ○ Virtual Meeting (VM)
- ○ Telephone Assisted (TAM)
- ○ Other

2. How many video assisted (VAM) meetings did you support between May 2019 and March 2020 (i.e. pre-COVID)?

- ○ 0
- ○ 1-2
- ○ 3-5
- ○ 6 or more

3. On average, how much longer or shorter did your summer 2020 meetings run compared to your normal pre-COVID meetings?

- ○ Much longer
- ○ Slightly longer

○ About the same
○ Slightly shorter
○ Much shorter
○ Don't know/NA

4. For the following items, please select how your Summer 2020 review meetings **compared** to your normal review meetings (i.e. pre-COVID-19)?

	Much easier	Slightly easier	About the same	Slightly harder	Much harder	Don't know/NA
Creating rosters and verifying accuracy	○	○	○	○	○	○
Dealing with last minute changes or requests	○	○	○	○	○	○
Handling COIs during the meeting	○	○	○	○	○	○
Trouble-shooting technical issues during the meeting	○	○	○	○	○	○
Ensuring post-meeting COI forms were signed	○	○	○	○	○	○
Solving problems that arose	○	○	○	○	○	○

5. Compared to the meetings you supported pre-COVID, on average, how much **more** time did you spend supporting **each** of your summer 2020 meetings (i.e. before, during and after the meeting)?

○ 0 (I did not spend any more time)
○ Less than 1 hour
○ 1 hour
○ 2 hours

○ 3 hours
○ More than 3 hours
○ Don't know/NA

6. If you spent more time than usual with your Summer 2020 meetings, please list which activities or duties demanded more of your time.

```

```

7. Approximately how much time did you spend trouble-shooting technical issues with reviewers during your meetings?

○ <30 mins.
○ 30 mins - 1 hour
○ 1-2 hours
○ 2-3 hours
○ 3-4 hours
○ More than 4 hours
○ Don't know/NA

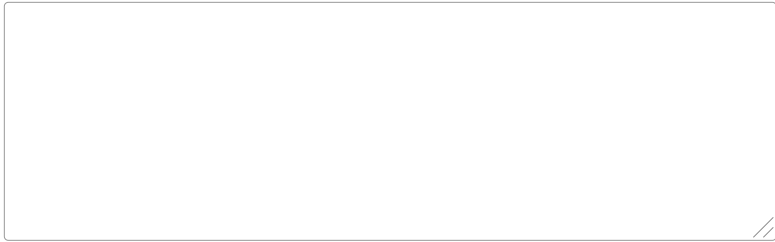8. Please list or describe any technical or logistical issues you experienced before or during the review meeting?

```

```

9. If you tested Zoom with reviewers before the meeting, what were the major technical issues they faced?

[text area]

10. What changes can VAM personnel make to the Zoom platform to help you prepare for and manage a more efficient and effective meeting?

[text area]

11. What can CSR do to help you manage review meetings using the Zoom platform?

[text area]

Submit