# Videoconferencing in Peer Review: Exploring Differences in Efficiency and Outcomes

Elizabeth L. Pier, Cecilia E. Ford, Anna Kaatz, Molly Carnes, and Mitchell J. Nathan
epier@wisc.edu, ceford@wisc.edu, akaatz@wisc.edu, mlcarnes@wisc.edu, mnathan@wisc.edu
University of Wisconsin-Madison

Joshua Raclaw, West Chester University, jraclaw@wcupa.edu

**Abstract:** Technology-mediated communication, such as teleconference and videoconference, has been found to affect group decision-making processes compared to face-to-face settings. Scientific peer review panels offer a site of authentic, collaborative decision making among expert scientists, yet no research has examined the impact of videoconferencing on such decision-making practices. We assigned real, de-identified grant applications submitted to the National Institutes of Health (NIH) to four panels of experienced NIH reviewers, one of which met via videoconference. The videoconference panel was slightly more efficient than the face-to-face panels, but the outcomes of their decision making (i.e., the scores assigned to grant applications) did not differ. However, preliminary analyses suggest there are differences in the *nature* of the collaborative discussion among reviewers between the two meeting formats. We discuss implications for research into technology-mediated collaborative decision making, as well as for the scientific grant peer review process broadly.

As research budgets tighten, funding agencies are seeking ways to reduce the costs of conducting grant peer review meetings (Bohannon, 2011), including the use of peer review panels conducted via teleconference and videoconference (Gallo, Carpenter, & Glisson, 2013). Decades of research (e.g., Bly, 1988; Driskell, Radtke, & Salas, 2003; O'Conaill, Whittaker, & Wilbur, 1993; Walther, 1997; Whittaker, 2003) investigating how the use of such technologies alters the ways in which people interact—particularly during problem-solving or decision-making tasks—suggest that virtual teams may function categorically differently than in-person teams (Andreev, Salomon, & Pliskin, 2010; Kiesler & Cummings, 2002; Kraut, Fussel, Brennan, & Siegel, 2002; McLeod, 1992; Straus & McGrath, 1994). However, there is not a consensus regarding whether technology-mediated communication is fundamentally different from in-person communication (Doherty-Sneddon, et al., 1997; Olson, Olson, & Meader, 1995). Given these discrepant findings, examining whether videoconferencing impacts the grant peer review process and whether it can serve as a viable alternative to traditional face-to-face peer review is of crucial importance, since grant peer review is the key mechanism by which precious research funds are allocated to scientists to conduct their research. This study stands to make an original contribution to our understanding of the mediating effect of technology not only on how expert scientists engage in collaborative decision making, but also on the outcomes of the scientific peer review process itself.

## Theoretical framework

Technology-mediated communication (TMC), including the use of teleconferencing or videoconferencing, has been found to affect group decision making. For example, compared to face-to-face (FTF) settings, researchers have found the use of TMC to increase difficulty in achieving consensus (Sellen, 1995), in managing turn taking (Anderson et al., 1999; O'Conaill et al., 1993; Tang & Isaacs, 1993), and in establishing mutual understanding (Clark & Brennan, 1991; Thompson & Coovert, 2003). In particular, Cramton (2001) found that TMC negatively impacts the ability of groups to establish common ground during tasks in which team members possess unique information (i.e., when knowledge or expertise is distributed, as it is during grant peer review). This negative effect is further exacerbated when the task is more complex, involves a higher workload, and requires group interdependence—all of which are features of grant peer review.

Given the time-consuming nature of peer review meetings, questions of efficiency tradeoffs are particularly acute. Although some researchers have found that computer-mediated meetings are shorter and more efficient than FTF meetings (e.g., Denstadli, Julsrud, & Hjorthol, 2012; O'Connaill et al., 1993; Tang & Isaacs, 1993), other scholars have found that TMC decreases task efficiency and increases the time to reach consensus (e.g., Doherty-Sneddon, et al., 1997; Straus & McGrath, 1994; Whittaker, 2003). Beyond mere efficiency, prior research suggests that TMC can reduce productivity and effectiveness in accomplishing tasks (Andreev et al., 2010; Kiesler & Cummings, 2002; Kraut et al., 2002; McLeod, 1992). Thus, it is an open empirical question as to whether TMC increases or decreases efficiency in peer review panel meetings, as well as whether it affects the outcomes of the decision-making process itself.

Much of the work examining the effect of TMC on group decision making has been done in lab settings (Anderson, McEwan, Bal, & Carletta, 2007; Whittaker, 2003), with "relatively little detailed empirical evidence on the impact of different forms of multimedia communication on patterns of communication in the workplace" (Anderson et al., 2007, p. 2560) and "few studies [that] have explicitly compared the way videoconferencing and face-to-face meetings are used in modern organizations" (Denstadli et al., 2012, p. 86). Only one study to date (Gallo et al., 2013) has investigated the role of TMC in peer review specifically; the authors found few differences between face-to-face and teleconference grant peer review except for a small difference in overall discussion time. Yet, no studies to date have examined the role of videoconferencing in peer review, despite the fact that it is a format increasingly used by many funding agencies (Bohannon, 2011). This study aims to fill this gap in knowledge by posing three questions that make a preliminary attempt to explore the role of TMC in grant peer review: (RQ1) Do FTF and VC peer review meetings differ in efficiency? (RQ2) Do FTF and VC meetings differ in their outcomes (i.e., the scores they assign to grant applications)? (RQ3) Do FTF and VC meetings differ in their collaborative scoring processes?

## Methods

The research team recruited biomedical scientists with experience reviewing for the National Institutes of Health (NIH) to participate in one of four peer review panel meetings—three conducted in person (FTF), and one conducted via videoconference (VC). Figure 1 is an anonymized screenshot from video of (a) one of our FTF meetings and (b) our VC meeting. Reviewers evaluated de-identified applications previously reviewed by real panels within NIH's National Cancer Institute between 2012 and 2015. We solicited Principal Investigators (PIs) using NIH's public access database, *RePORTER*, to donate applications that were either funded or not funded on the first submission. Each panel had between eight and 12 reviewers who evaluated six applications apiece, with three reviewers assigned to a given application in each meeting. Based on the three reviewers' preliminary scores, the top 50% of applications were discussed in a given meeting, with the bottom 50% triaged out from discussion (as is typical in NIH peer review), so that each panel discussed between eight and 11 applications depending on the number of participating reviewers.

Our meetings were designed to follow the norms and practices of actual NIH peer review in all aspects of study design, and all methodological decisions were made in consultation with staff from NIH's Center for Scientific Review and with a retired Scientific Review Officer (SRO), who assisted with recruiting reviewers and chairpersons, assigning reviewers to applications, and overseeing each meeting. For a detailed description of the methods used to design the meetings, see Pier et al. (2017).
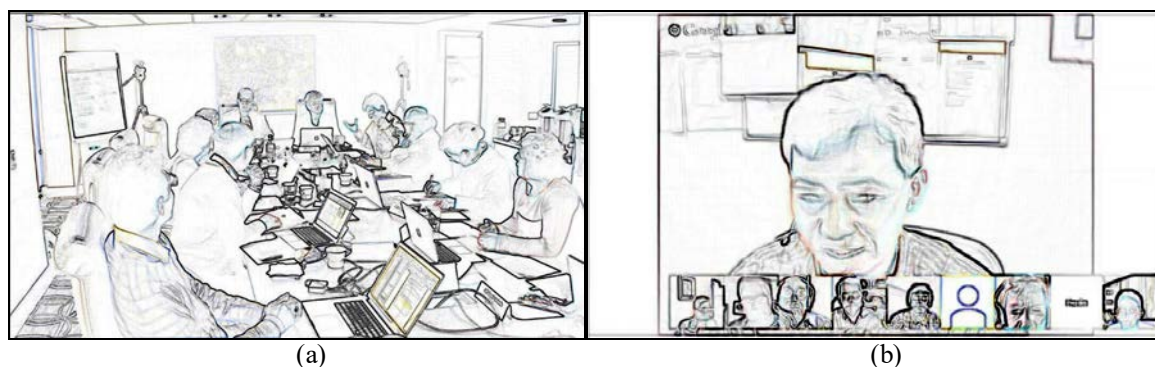


| (a) | (b) |

Figure 1. De-identified screenshots of one FTF meeting (a) and the VC meeting (b).

To answer RQ1, we measured the amount of time spent discussing each grant application in each meeting, beginning at the moment the chairperson introduced the grant to be discussed, and ending when the chairperson introduced the subsequent grant to be discussed. To answer RQ2, we compared the panels' final scores for each application. To answer RQ3, we examined the degree to which reviewers' scores changed as a function of collaborative discussion.

## Results

For RQ1, we found that on average, the videoconference meeting was the most efficient meeting in terms of average time spent per application (Table 1). Panelists in the VC meeting spent 2 minutes and 18 seconds less, per application, on average compared to the three FTF meetings. However, for RQ2, we found that the VC meeting did not perform much differently from the FTF meetings in terms of the scores they assigned to

applications (Table 2). At NIH, panel scores range from 10 (best) to 90 (worst) and constitute the average of all panelists' scores following collaborative discussion. On average, the videoconference panel assigned similar scores to their pool of applications as the FTF panels (although the first FTF panel stands apart as slightly harsher overall, since the average score was higher, i.e., worse). Importantly, for RQ3, we found that reviewers in the VC panel changed their scores during the meeting as a function of collaborative discussion less frequently (no change 55% of the time) than the FTF panels on average (no change 37.5% of the time), and they worsened their scores less frequently (30% of the time) than the FTF panels on average (52.1% of the time). Therefore, although the scores themselves do not appear to differ between the formats overall, the patterns of score changes suggest there are differences in the nature of the collaboration in FTF versus VC panels.

Table 1: Average time spent (minutes:seconds) discussing applications

| FTF 1 | FTF 2 | FTF 3 | FTF Average | VC | Total Average |
|---|---|---|---|---|---|
| $M = 14:52$ $SD = 1:54$ | $M = 16:17$ $SD = 6:07$ | $M = 17:18$ $SD = 3:40$ | $M = 16:09$ $SD = 4:15$ | $M = 13:51$ $SD = 3:26$ | $M = 15:48$ $SD = 3:59$ |
| *Note*. FTF 1 = Face-to-Face Panel #1, and so forth. VC = Videoconference Panel. | | | | | |

Table 2: Average final panel scores in each meeting

| FTF 1 | FTF 2 | FTF 3 | FTF Average | VC | Total Average |
|---|---|---|---|---|---|
| $M = 38.3$ $SD = 10.1$ | $M = 32.3$ $SD = 5.5$ | $M = 31.5$ $SD = 8.7$ | $M = 34.0$ $SD = 8.6$ | $M = 31.6$ $SD = 7.1$ | $M = 33.4$ $SD = 7.5$ |

Table 3: Number (and percentage) of times reviewers changed their scores during the meeting

| | FTF 1 | FTF 2 | FTF 3 | FTF Average | VC | Total Sum |
|---|---|---|---|---|---|---|
| Improved score | 2 (6.3%) | 3 (9.7%) | 5 (15.2%) | 3.33 (10.4%) | 3 (15.0%) | 13 (11.2%) |
| No change | 7 (21.9%) | 16 (51.6%) | 13 (39.4%) | 12.0 (37.5%) | 11 (55.0%) | 47 (40.5%) |
| Worsened score | 23 (71.9%) | 12 (38.7%) | 15 (45.5%) | 16.67 (52.1%) | 6 (30.0%) | 56 (48.3%) |
| Sum | 32 (100%) | 31 (100%) | 33 (100%) | 32 (100%) | 22 (100%) | 116 (100%) |

## Discussion and conclusion

We found that there was an efficiency gain for the videoconference peer review meeting over the three face-to-face meetings, which aligns with prior research finding that TMC meetings are shorter and more efficient (Denstadli et al., 2012; O'Connaill et al., 1993; Tang & Isaacs, 1993). This may be due to less discussion time in VC formats stemming from fewer turns of talk, in part due to the heightened barrier to entry into conversation that videoconferencing introduces. Importantly, we found that the efficiency gain of the VC panel was not accompanied by a noticeable difference in the average final scores that the panels assigned to applications. Thus, despite the complexity of the task and the distributed nature of reviewers' expertise, TMC may not hinder a group of expert scientists as they engage in grant peer review, echoing Gallo and colleagues' (2013) finding regarding the use of teleconference peer review meetings (cf. Cramton, 2001). Utilizing videoconferencing to conduct peer review meetings may thus offer a reasonable solution to funding agencies' tightening budgets.

However, our preliminary investigations into the *process* by which reviewers arrive at the final scores suggest that there are some differences in the VC format necessitating further examination. We found that reviewers in the VC meeting changed their scores less frequently than reviewers in the FTF format, and that they worsened their scores less frequently than the FTF panels; this implies there may be differences in the collaboration among panelists in this format resulting in less frequent score change, and less score change of a critical nature. Our future work plans to examine how turn taking unfolds in each panel meeting, to quantify the number of unique contributors to each discussion, and to explore the decision-making strategies each panel employs to achieve consensus.

This short paper offers descriptive insights into our expanding and evolving understanding of how technology-mediated communication affects collaborative decision making in various contexts. Given that this is an exploratory pilot study restricted to a single VC panel, it is limited in its generalizability beyond our sample. Furthermore, lack of random assignment of applications and of reviewers precludes any causal claims. Nevertheless, this work presents preliminary findings from our data that will guide our future research. Given

the importance of grant peer review for the enterprise of science as a whole, and that many funding agencies are increasingly conducting videoconference peer review meetings, understanding how the use of videoconference may change the process and outcomes of peer review is of paramount importance.

## References

Anderson, A. H., McEwan, R., Bal, J., & Carletta, J. (2007). Virtual team meetings: An analysis of communication and context. *Computers in Human Behavior, 23*, 2558–2580.

Anderson, A. H., Mullin, J., Katsavras, E., McEwan, R., Grattam, E., Brundell, P., & O'Malley, C. (1999). Multi-mediating multiparty interactions. In M. A. Sasse & C. Johnson (Eds.), *Human–computer interaction – INTERACT'99* (pp. 313–320). Amsterdam: IOS Press.

Andreev, P., Salomon, I., & Pliskin, N. (2010). Review: State of teleactivities. *Transportation Research Part C: Emerging Technologies, 18*(1), 3–20.

Bly, S. (1988). A use of drawing surfaces in collaborative settings. In *Proceedings of Conference on Computer Supported Cooperative Work* (pp. 250–256). New York: ACM Press.

Bohannon, J. (2011). Meeting for peer review at a resort that's virtually free. Science, 331, 27.

Clark, H. H., & Brennan, S. E. (1991). Grounding in communication. In L. B. Resnick, J. M. Levine, & S. D. Teasley (Eds.), *Perspectives on socially shared cognition* (pp. 127–149). Washington, DC: American Psychological Association.

Cramton, C. D. (2001). The mutual knowledge problem and its consequences for dispersed collaboration. *Organizational Science, 12,* 346–371.

Denstadli, J. M., Julsrud, T. E., & Hjorthol, R. J. (2012). Videoconferencing as a mode of communication: A comparative study of the use of videoconferencing and face-to-face meetings. *Journal of Business and Technical Communication, 26*(1), 65–91.

Doherty-Sneddon, G., Anderson, A., O'Malley, C., Langton, S., Garrod, S., & Bruce, V. (1997). Face-to-face and video-mediated communication: A comparison of dialogue structure and task performance. *Journal of Experimental Psychology: Applied, 3,* 105–125.

Driskell, J. E., Radtke, P. H., & Salas, E. (2003). Virtual teams: Effects of technological mediation on team performance. *Group Dynamics: Theory, Research, and Practice, 7*(4), 297–323.

Gallo, S. A., Carpenter, A. S., & Glisson, S. R. (2013). Teleconference versus face-to-face scientific peer review of grant application: Effects on review outcomes. *PLOS ONE, 8*(8), 1–9.

Kiesler, S., & Cummings, J. N. (2002). What do we know about proximity and distance in work groups? A legacy of research. In P. Hinds & S. Kiesler (Eds.), *Distributed work* (pp. 57–82). Cambridge, MA: MIT Press.

Kraut, R., Fussel, S. R, Brennan, S. E., & Siegel, J. (2002). Understanding effects of proximity on collaboration: Implications for technology to support remote collaborative work. In P. Hinds & S. Kiesler (Eds.), *Distributed work* (pp. 137–163). Cambridge, MA: MIT Press.

McLeod, P. L. (1992). An assessment of the experimental literature on electronic support of group work: Results of a meta-analysis. *Human-Computer Interaction, 7,* 257–280.

O'Conaill, B., Whittaker, S., & Wilbur, S. (1993). Conversations over video conferences: An evaluation of spoken aspects of video-mediated communication. *Human-Computer Interaction, 8,* 389–428.

Olson, J. S., Olson, G. M., & Meader, D. K. (1995). What mix of video and audio is useful for remote real-time work? In *Proceedings of CHI '95* (pp. 362–368). New York: ACM Press.

Pier, E. L., Raclaw, J., Kaatz, A., Brauer, M., Carnes, M., Nathan, M. J. & Ford, C. E. (2017). "Your comments are meaner than your score:" Score Calibration Talk influences intra- and inter-panel variability during scientific grant peer review. *Research Evaluation, 26*(1), 1-14.

Sellen, A. (1995). Remote conversations: The effects of mediating talk with technology. *Human–Computer Interaction, 10,* 401–441.

Straus, S. G., & McGrath, J. E. (1994). Does the medium matter? The interaction of task type and technology on group performance and member reactions. *Journal of Applied Psychology, 79,* 87–97.

Tang, J. C., & Isaacs, E. (1993). Why do users like video? Studies of multimedia-supported collaboration. *Computer Supported Collaborative Work (CSCW), 1,* 163–196.

Thompson, L. F., & Coovert, M. D. (2003). Teamwork online: The effects of computer conferencing on perceived confusion, satisfaction, and postdiscussion accuracy. *Group Dynamics, 7,* 135–151.

Walther, J. B. (1997). Group and interpersonal effects in international computer-mediated collaboration. *Human Communication Research, 23,* 342–369.

Whittaker, S. (2003). Theories and methods in mediated communication. In A. Graesser, M. Gernsbacher, & S. Goldman (Eds.), *Handbook of discourse processes* (pp. 243–286). Mahwah, NJ: Lawrence Erlbaum.