Can You See Me Now? How Field of View Affects Collaboration in Robotic Telepresence

Steven Johnson,[1] **Irene Rae,**[1] **Bilge Mutlu,**[1] **Leila Takayama**[2*]
(1) University of Wisconsin–Madison; (2) Willow Garage, Inc.
sjj@cs.wisc.edu; irene@cs.wisc.edu; bilge@cs.wisc.edu; takayama@google.com

ABSTRACT

Robotic telepresence systems—videoconferencing systems that allow a remote user to drive around in another location—are an emerging technology for supporting geographically-distributed teams. Thus far, many of these systems rely on affordances designed for stationary systems, such as a single, narrow-view camera to provide vision for the remote user. Teleoperation has offered some solutions to this via an augmented field-of-view, but how these solutions support task outcomes in collaborative mobile telepresence tasks has yet to be understood. To investigate this, we conducted a three condition (field-of-view: narrow (45°) vs. wide-angle (180°) vs. panoramic (360°)) between-participants controlled laboratory experiment. We asked participants ($N = 24$) to collaborate with a confederate via a robotic telepresence system while using one of these views in a redecoration task. Our results showed that wider views supported task efficiency and fewer collisions, but were perceived as more difficult to use.

ACM Classification Keywords

H.5.3 **Information Interfaces and Presentation**: Group and Organization Interfaces–*Collaborative computing, Computer-supported cooperative work, Evaluation/methodology*; H.4.3 **Information Systems Applications**: Communications Applications–*Computer conferencing, Teleconferencing, Videoconferencing*

Author Keywords

Remote collaboration, robot-mediated collaboration, telepresence, teleoperation, field-of-view, situation awareness

INTRODUCTION

Robotic telepresence systems are an emerging technology that meld videoconferencing and robotic technologies, allowing geographically distributed people to interact in a way that is similar to face-to-face [22]. Users who are remote from the system, referred to as *operators* or *teleoperators*, appear on the screen in the same way that they would in videoconferencing (e.g. Skype), but are also able to move in the local user's space. While these systems have existed in research settings for many years [28], they have only recently become mature enough to reach the consumer market and are now increasingly being adopted in business [23, 44], education [11], home care [26], and medical settings [12, 42]. A variety of systems have been designed to meet this growing consumer demand, and numerous systems are continually being developed and introduced to address user needs in these settings [7, 31, 37].

Figure 1. Participants remotely operated a telepresence robot using a desktop interface (top) to remotely collaborate with a confederate in a room-redecoration task (bottom).

Despite this proliferation of interest, these systems are still relatively new and current designs often rely on affordances designed for stationary systems, augmented with additional robotic capabilities. For example, these systems typically feature a videoconferencing screen mounted on a mobile robotic base [7, 37]. Previous research has shown support for these systems improving remote users' feelings of presence in the local user's environment [29] and engagement with local users in collaborative work settings [23, 44].

However, literature has shown that these systems are not universally beneficial and that simply adding a mobile robotic

[*] Author is currently affiliated with Google, Inc.

CHI 2015, April 18–23, 2015, Seoul, Republic of Korea.

ACM 978-1-4503-3145-6/15/04 ...$15.00.
http://dx.doi.org/10.1145/2702123.2702526

base is not sufficient to improve all collaborative outcomes, particularly in mobile tasks [29]. Past work has suggested that a reason for this outcome may have been the inherent limitations of adapting traditional video capabilities to mobile systems. Although remote users depend heavily on visual information to navigate and interact via the telepresence robot [30], current commercial systems still use traditional videoconferencing designs, relying on a single camera with a narrow field-of-view.

Research into teleoperation—control over a machine from a distance—has explored ways to integrate video to better support the remote user's mobility when operating the system. For example, prior work has shown that augmenting video (e.g. via wide-angle lenses, panoramic views, zoom/pan capabilities, etc.) improves operator efficiency and awareness [35, 45]. While these solutions hold promise, they have not been extensively studied in the context of collaborative work amongst geographically-distributed teams.

Our goal in this work is to further disentangle how the field-of-view affects collaborative outcomes in mobile tasks, contributing to the field in two ways. First, we conduct a systematic investigation of how augmented video solutions, proposed in teleoperation research, affects outcomes in collaborative settings. Second, through our study design and measures, we present a methodology for comparing fields-of-view in collaboration and demonstrate its use to highlight key differences, opening the door for further studies of this type.

Next, we provide an overview of related work in telepresence and teleoperation. We then describe our hypotheses and study design. Last, we present our results, discuss our findings, and highlight their implications for future research and design.

RELATED WORK

Prior work in telepresence and teleoperation has identified the operator's vision as playing a key role in improving user experience and task performance [30]. Much of the work on supporting this element, to date, has fallen into one of two categories: (1) discussing the limitations of providing vision through video and (2) offering ways of improving upon these limitations. In this section, we review prior work in each of these categories to highlight how it has informed our hypotheses and study design.

Limitations of Video

Unlike face-to-face communications, the limitations of providing a live video-feed for telepresence users are dictated by the technologies selected. Designers of telepresence and teleoperated systems may choose from a range of options for the video feed, from single, narrow-view cameras to "better-than-human" full, 360° panoramas. However, each of these approaches has been shown to have its own problems, leaving the question open as to how much view is "just right."

Narrow View Limitations

Operator *situation awareness*—awareness of the remote environment allowing a teleoperator to give informed instructions to the system—is needed to efficiently function in a remote environment [49]. Prior work has shown that a lack of situation awareness correlates with poor operator performance [5, 32], often stemming from the operator's lack of understanding of the remote system's surroundings, status, or location [9]. The use of standard video cameras that are designed to provide only a narrow range of view can lead to what is known as the *soda straw* or *keyhole effect*, where the limited visibility requires extra user effort to achieve awareness levels similar to having a direct view of the environment [48].

In addition to issues involving situation awareness, previous work has identified a number of other issues caused by a restricted field-of-view, such as *cognitive tunneling*—an individual failing to recognize landmarks, even those which are in view [39], errors in distance and depth judgments [47], and an impaired ability to detect targets or obstacles [3, 43].

Wider View Limitations

Other work has revealed that increasing the range of view brings with it additional costs. For example, prior literature has shown evidence that wider views may increase cognitive workload for driving and object localization tasks [2], cause motion sickness [36], and distort perceptions of velocity resulting in reduced driving speed [27, 36]. The use of multiple cameras, an alternate method for increasing the operator's view, has been discussed as an option which is subject to bandwidth limitations, potentially untenable in real world settings [19].

Improving Upon View Limitations

Literature targeted at improving upon field-of-view limitations has taken numerous technological approaches. For example, past work has found that wider fields-of-view are useful for maintaining spatial orientation with respect to terrain features [21], ease of operation in off-road driving tasks [24], and performance in unfamiliar terrains was better when using wider fields-of-view [34]. Additionally, other view explorations showed that using images from five forward-facing cameras aided operators in achieving efficient navigation [45].

Related work compared operator performance when using direct control, front camera, fish-eye camera, omni-camera, and panoramic views, and found that those using the fish-eye and omni-cameras performed better than those using direct control [35]. However, the differences in amounts of distortion and blind spot location from their particular implementations of the fish-eye and omnidirectional view make it difficult to isolate the effects that field-of-view had on their results.

Telepresence research has also identified problems associated with providing a limited field-of-view. Work in this area has suggested that a wide view may be beneficial for operators [6] and has experimented with alternate solutions, such as pan/tilt capabilities [44], foveal video to provide less-detailed but wider fields-of-view while accommodating bandwidth limitations [18], automated driving capabilities to reduce reliance on video [41], and cameras placed in the remote environment to add additional perspective [16].

While both teleoperation and telepresence work has highlighted the problems in existing systems and the need to provide the remote user with the correct amount of view to support their use of the system, a definitive solution has yet to

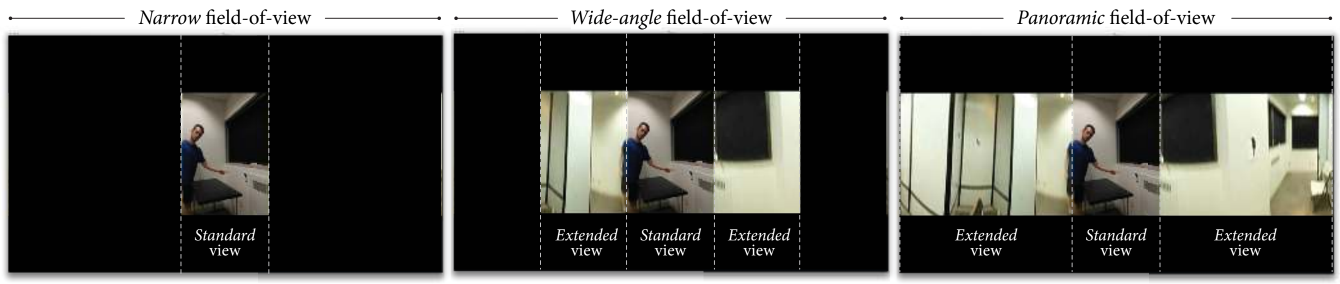

Figure 2. Participants collaborated with a partner using one of three view conditions: a 45° *narrow* condition (left), a 180° *wide-angle* condition (middle), or a 360° *panoramic* condition (right). In all conditions, the video outside of the central 45° is a lower-quality peripheral view.

emerge. However, what past literature has demonstrated is that the benefits and drawbacks of these solutions and how they support task outcomes are not clear cut. Measurement of the success of these solutions is complex and composed of many elements, such as task performance, operator situation awareness, feelings of *presence*—the sensation of being there—for both remote and local users [29], perceptions of collaborative success—the confidence in the correctness of the task and the fluidity of the interaction—for operators [17], and many others [22].

In this work, we gain a deeper understanding of how the field-of-view for operators affects collaborative outcomes in mobile tasks. We next describe how this body of literature informed our selection of measures and the formation of our hypotheses.

HYPOTHESES

To study the effects that augmented views may have on remote collaborators, we focused on four key aspects of collaborative outcomes: task performance, situation awareness, feelings of presence, and perceptions of success. Informed by prior literature, we present five hypotheses predicting the effect that differences in the field-of-view may have. We also examined another aspect contributing to collaborative outcomes, the usability and difficulty of the interface, and posit a sixth hypothesis of how it may be influenced by the field-of-view.

Hypothesis 1: The use of a wide-angle or panoramic view will improve users' task performance relative to a narrow camera view by reducing the effects of cognitive tunneling [39].

Hypothesis 2: A panoramic view will increase cognitive workload while driving [2] and reduce driving speed [27, 36] compared to a wide-angle view. While the panoramic view may further reduce the effects of cognitive tunneling, we posit that this reduction will be overshadowed by the increase in cognitive workload. Thus, the use of a wide-angle view will improve users' task performance relative to a panoramic view.

Hypothesis 3: Wider fields-of-view will lead to greater situation awareness by reducing the impact of the keyhole effect [48]. Specifically, situation awareness will be greater in the wide-angle than the narrow view and greater in the panoramic view than the wide-angle view.

Hypothesis 4: Wider fields-of-view will increase user immersion—shown to influence feelings of presence [33]—increasing user feelings of presence relative to narrower fields-of-view. Specifically, feelings of presence will be greater in the wide-angle view than the narrow view, and greater in the panoramic view than the wide-angle view.

Hypothesis 5: Wider fields-of-view will increase perceptions of collaborative success relative to narrower fields-of-view due to higher levels of situational awareness [49], resulting in increased knowledge of the state of the workspace, shown to be critical to remote collaboration [13]. Specifically, perceptions of collaborative success will be greater in the wide-angle view than the narrow view and greater in the panoramic view than the wide-angle view.

Hypothesis 6: Wider fields-of-view will increase perceived interface difficulty due to the combination of more complex viewing information and greater cognitive processing requirements [2] relative to narrower fields-of-view. Specifically, perceptions of difficulty will be greater for wide-angle view users than narrow view users, and greater for panoramic view users than wide-angle view users.

METHOD

To test our hypotheses, we conducted a between-participants controlled laboratory experiment with three field-of-view conditions: a standard *narrow* view (45°), a *wide-angle* periphery-augmented view (180°), and a *panoramic* periphery-augmented view (360°). The wide-angle and panoramic conditions used the 45° high-resolution view displayed in the narrow condition, augmented with a lower-resolution periphery, depicted in Figure 2. We selected 45° because several commercial systems use a similar field-of-view [7, 31], 180° as it corresponds to the human perceptual system [15], and 360° as it represents the maximum view possible. Participants in each condition operated a telepresence robot as a remote user to collaborate on a room-redecoration task with a confederate acting as a local user. The confederate was a 23-year-old male working in a different room in the same building. During the task, the confederate did not initiate any actions, but cooperated fully with the participant. As in prior work [20], we maintained consistency across participants by limiting the confederate's dialog using the semi-scripted response rules of: 1) answering any question asked, 2) trying to give the best answer in response to questions, and 3) asking for clarification if directives were unclear. Participants were told that the confederate was working with them from a different building on campus. The same confederate participated in all trials.

Participants

A total of 24 participants (12 male, 12 female), stratified by gender within each condition, took part in our study. All participants were native English speakers recruited from the University of Wisconsin–Madison campus. Participants were aged 19–40 ($M = 24.208$, $SD = 5.816$), and reported that they were unfamiliar with robots on a seven-point scale ($M = 2.625$, $SD = 1.498$) (1 = not at all familiar, 7 = very familiar) and moderately comfortable with videoconferencing ($M = 4.958$, $SD = 1.334$) (1 = not at all comfortable, 7 = very comfortable).

System

The interactions in our study took place via a Double telepresence robot [7] shown in Figures 1 and 3. The robot weighs 15 pounds (6.8 kg) and has an adjustable height of 47 or 60 inches (101.6 or 152.4 cm). We limited the height to the shorter setting to be consistent for all participants. The Double's screen was an iPad 2 tablet computer[1] with a diagonal screen size of 9.7 inches (24.6 cm) and a 2048 × 1536 resolution.

As our objective was to study the effects that changes to the field-of-view might have, we used the same lens for each condition to eliminate lens distortion [35]. We used a catadioptric lens (a lens employing a curved mirror), which allowed capture of up to a full 360° view around the robot, and used techniques from prior literature for unfolding the video from the lens [38]. Because this type of lens disallows use of the entire camera capture area, the periphery views were necessarily of a lower resolution than the forward-facing camera. Consequently, the peripheral views appeared slightly blurry (see Figure 2).

The interface for the remote participant was similar to a focus+context display [1] in that it had two parts: a high-resolution forward-facing view, a typical part of the robot's standard interface, and a low-resolution augmented view of the periphery displayed on the sides in the wide-angle and panoramic conditions. The controls for the robot were the uniform across conditions.

We augmented the standard view provided by the forward-facing camera with a peripheral view of the environment using two additional robot-mounted cameras. One camera provided the forward-facing view, typically provided by the iPad tablet. To have better control over the interface elements of the remote conference GUI, we used the back-facing camera of a Nexus 7 tablet[2] to provide this video to our own videoconferencing application. For the second camera, we used the back-facing camera of a second Nexus 7 tablet equipped with an off-the-shelf, inexpensive catadioptric lens to allow capture of a 360° view around the robot. This lens was a GoPano Micro lens[3]. Both cameras transmitted video at approximately 30 frames per second with 1024 × 768 resolution. Figure 3 provides a diagram of the elements of the modified telepresence system.

Setup

The experimental room simulated a complex spatial environment, such as a museum or stock room, containing a dividing wall, tables, and chairs. The furniture offered reference points for remote participants to use to orient themselves within the room [29]. Figure 4 shows the experimental room layout and a participant navigating around the dividing wall.

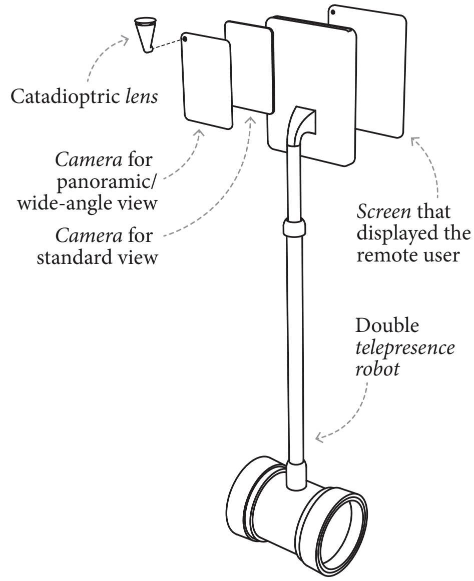

Figure 3. An overview of the elements of our system.

Tasks

Participants operated the telepresence robot in both a training and an experimental task. Both tasks took place in the same room, which was not changed between tasks. A redecoration task was selected as it involves a high degree of mobility and requires navigating an unfamiliar environment, simulating many real world tasks, such as searching in a warehouse or exploring a museum.

Training Task

Participants were given up to ten minutes to familiarize themselves with the controls and to practice driving the telepresence robot. For the training task, participants were given a blank map of the room showing the layout of the room's outer and center walls and were asked to label photo locations on the map. Provided with the map were examples and descriptions of three types of galaxies: spiral, elliptical, and irregular. Participants were asked to find, classify, and label the thirteen galaxy photos in the room using the blank map provided. In addition to the galaxy photos, the room also contained twelve pictures of tangrams for use in the experimental task. After ten minutes had elapsed or the participant indicated completion, the connection with the robot was terminated and the photos of the galaxies were removed.

Tangrams and galaxies were selected for their level of abstraction. This abstraction made the pictures difficult to describe quickly, increasing the need for participant mobility during the task and facilitating greater discussion and collaboration. Additionally, the lack of inherent orientation in these pictures encouraged participants to employ greater attention to detail and the differences between picture sets helped to differentiate the tasks for participants.

Experimental Task

Participants engaged in a room redecoration task where they were provided with three reference sheets illustrating the new

[1]http://www.apple.com/ipad/
[2]http://www.google.com/nexus/7/
[3]http://www.gopano.com/products/gopano-micro/

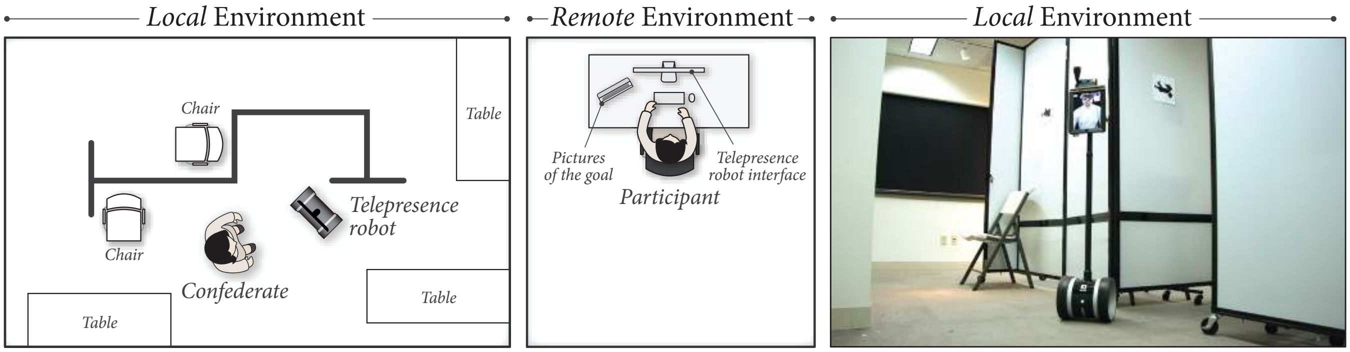

Figure 4. ***Left:*** **The physical arrangement of the local and remote environments.** ***Right:*** **A participant navigating in the remote environment.**

configuration of tangrams in the room. Participants were instructed to change the room to match the reference sheets and were given unlimited time to complete the task.

Measurement

To capture collaborative outcomes in the redecoration task, we used several objective, subjective, and behavioral measures.

Objective Measures

We measured task performance and situation awareness using completion time—timed from when the experimenter left the room to when the participant indicated task completion—and error rate—the number of incorrectly placed or oriented tangram photos. We also recorded the number of collisions—the participant running into something with the robot—as a measure of situation awareness.

Subjective Measures

We administered a post-task questionnaire consisting of 28 questions, designed to capture situation awareness, perceived success of the collaboration, perceived difficulty of using the telepresence interface, and feelings of presence in the remote environment. We used two measures of situation awareness, a modified version of the NASA Situation Awareness Rating Technique (SART) [10] and a blank map of the room where participants were asked to draw the locations and orientations of the furniture. Position and orientation errors were scored by two experimenters who were blind to condition.

Perceived collaborative success and perceptions of the difficulty of using the telepresence interface were measured via participant agreement with six items on a five-point scale (1 = Strongly Disagree to 5 = Strongly Agree), e.g., "I am confident that our final layout had fewer errors than other teams." and "I was able to control how I moved in the remote environment."

Feelings of presence were captured using a map, similar to previous work [29]. Participants were asked to mark where they felt they and their partner had worked during the task on a printed map of the rooms. These maps were coded as "co-located" if they marked that they worked in the room with the confederate, or "remote" if they marked that they worked in the room where the computer they were using was.

Behavioral Measures

Participants wore head-worn eye-trackers (SMI eye-tracking glasses[4]) during the experimental task. While we had no specific hypotheses related to participants' gaze, we used the eye-tracking data to explore trends, common occurrences, and unique events that occurred. We used this data to aid in the explanation and discussion of our results and to identify what parts of the field-of-view were directly used during the task.

Procedure

Participants were greeted, asked for consent, and were then given an overview of the tasks. The experimenter then seated the participant in front of a computer, explained the controls for the robot, and provided the participant with a reference sheet for the controls. At this time, the participant was asked to wear the head-worn eye-tracker to allow them to acclimate to wearing the device. Next, the participant was given up to ten minutes to complete the training task. Once the time had elapsed or the participant indicated task completion, the experimenter re-entered the room and disconnected from the telepresence robot. The experimenter then instructed the participant on the experimental task, provided the reference sheets depicting the goal layout, and calibrated the eye-tracker. Once the eye-tracking system was online, the experimenter re-established the connection with the telepresence robot and introduced the confederate as another participant in the study. The experimenter reiterated that the fastest team to correctly redecorate the room would receive an extra dollar, told them to begin when the timer was started, and asked them to open the door when they finished. After answering any questions, the experimenter started the timer and exited the room.

After the participant indicated completion of the task, the experimenter stopped the timer, disconnected the robot, and administered the post-task questionnaire. Upon completion of the questionnaire, the participant was debriefed. Each session took between 45 – 60 minutes and all participants were paid $11 USD, which included the incentive dollar.

Analyses

A one-way fixed-effects analysis of variance (ANOVA) was conducted with the field-of-view as the input variable and the objective and subjective measures as response variables. Tangram placement errors were not analyzed due to a ceiling effect, as only one participant made errors in placement. Following Wickens and Keppel [46], we calculated *a priori* contrast tests to make comparisons indicated by our hypotheses. A Pearson's Chi-squared test was used to determine the effects of field-of-view on participants' feelings of presence

[4] http://www.eyetracking-glasses.com/

in the drawn map measure. To address the non-normal distributions of the counted data (collisions and map errors), a generalized linear model with a Poisson distribution was used. Age and gender were considered as potential covariates in the analyses and were included in the models when they served as a significant predictor of the response variable ($p < 0.25$ cutoff was used). Video game experience was tested as a confound for driving ability, but was insignificant as a covariate.

To construct scales from the questionnaire items, we conducted an exploratory factor analysis, resulting in two factors: perceived collaborative success (two items, Cronbach's $\alpha = .772$) and perceived interface difficulty (three items, $\alpha = .698$).

Recordings of participants' gaze from the eye-tracking system were reviewed for patterns and unique events.

RESULTS

Our first hypothesis predicted that participants in the wide-angle and panoramic view conditions would outperform those in the narrow condition in measures of task performance. Our analysis provides support for this hypothesis; the average task completion times in seconds were 1428.5 (SD = 304.36), 952.88 (SD = 256.53), and 978.38 (SD = 313.84), in the narrow, wide-angle, and panoramic conditions, respectively. We found a main effect of field-of-view on task completion time, $F(2,23) = 6.686$, $p = .006$, $\eta^2 = .389$. Participants in the wide-angle condition completed the task faster than those in the narrow condition, $F(1,23) = 10.565$, $p = .004$, $\eta^2 = .307$. Participants in the panoramic condition were also significantly faster than those in the narrow condition, $F(1,23) = 9.463$, $p = .006$, $\eta^2 = .275$.

Our second hypothesis predicted that users in the wide-angle condition would outperform those in the panoramic condition in measures of task performance. We did not find support for this hypothesis. A contrast test comparing the average task completion time of the wide-angle and panoramic conditions revealed no significant differences, $F(1,23) = .030$, $p = .863$.

We found partial support for our third hypothesis, which posited that wider fields-of-view would improve situation awareness. The average number of collisions was 1.50 ($SD = .926$), .375 ($SD = .518$), and .625 ($SD = .916$), in the narrow, wide-angle, and panoramic conditions, respectively. We found a main effect of field-of-view on the number of collisions, $\chi^2(2, N = 24) = 6.439$, $p = .040$. Both augmented view conditions reduced the number of collisions compared to the narrow condition. Users made fewer collisions when using the wide-angle over the narrow view, $\chi^2(1, N = 24) = 5.782$, $p = .016$, and marginally fewer collisions using the panoramic over the narrow view, $\chi^2(1, N = 24) = 2.970$, $p = .085$. There was no evidence that there were fewer collisions in the panoramic vs. wide-angle view, $\chi^2(1, N = 24) = .505$, $p = .477$.

Our analysis of the room recall maps revealed support for improved situation awareness of panoramic over wide-angle view users. The average number of errors was 1.875 ($SD = 1.553$), 2.000 ($SD = 1.309$), and .875 ($SD = .835$) in the narrow, wide-angle, and panoramic conditions, respectively. Controlling for participant age, we found a main effect of field-of-view on the number of map errors, $\chi^2(3, N = 24) = 8.999$, $p = .029$. Panoramic users made fewer errors than narrow, $\chi^2(1, N = 24) = 5.162$, $p = .023$, and wide-angle, $\chi^2(1, N = 24) = 4.579$, $p = .032$ users.

Our analysis of responses to the SART scale for measuring situation awareness revealed no effect of field-of-view on scale responses, $F(2,23) = .307$, $p = .739$.

We did not find support for our fourth hypothesis, which predicted that a wider field-of-view would increase users' feelings of presence. Our analysis revealed no significant effect of field-of-view on feelings of presence, $\chi^2(1, n = 24) = 2.4$, $p = .301$.

Responses to the perceived collaborative success scale revealed mixed support for our fifth hypothesis, which predicted that a wider field-of-view would increase perceptions of collaborative success. Average responses were 4.750 ($SD = .463$), 3.938 ($SD = .623$), and 4.563 ($SD = .563$) in the narrow, wide-angle, and panoramic conditions, respectively. We found a main effect of field-of-view on perceived success, $F(2,23) = 4.723$, $p = .020$, $\eta^2 = .310$. Participants rated their collaboration in the wide angle condition as less successful than in the narrow condition, $F(1,23) = 8.614$, $p = .008$,

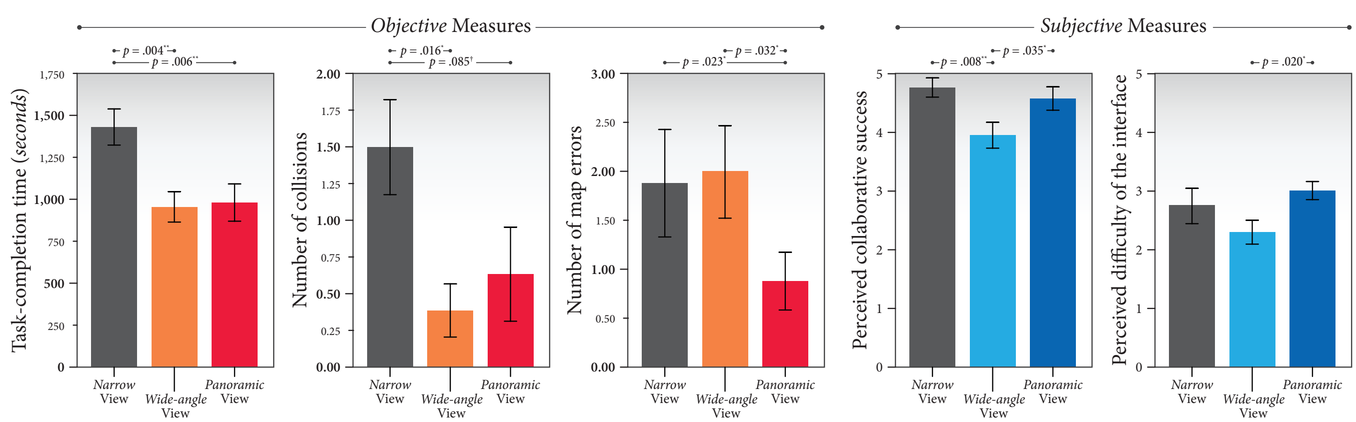

Figure 5. ***Left:*** **Mean task completion time, task errors, and remote map recall errors.** ***Right:*** **Mean scale responses of participants' perceptions of collaborative success and interface difficulty. (†), (*), and (**) denote p < .10, p < .05, and p < .01, respectively. The augmented view conditions significantly improved task completion time and reduced collisions over the narrow condition. Panoramic users made significantly fewer recall errors drawing the map of the remote environment compared to narrow and wide-angle users. Participants perceived the panoramic interface to be significantly more difficult to use than the wide-angle interface, but also perceived their collaboration to be significantly more successful.**

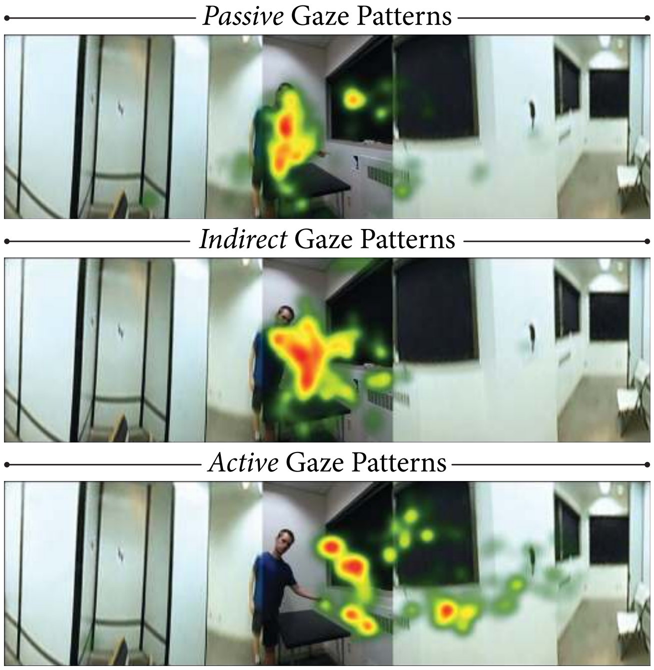

Figure 6. Sample heat maps extracted from gaze recordings for the three user classes. *Passive* users infrequently fixate on the view of the periphery, and *indirect* users never do. *Active* users spent a large portion of the task fixated on objects in the peripheral views.

$\eta^2 = .283$, and lower in the wide-angle condition than the panoramic condition, $F(1,23) = 5.097$, $p = .035$, $\eta^2 = .167$. We found no differences in perceived success between the narrow and panoramic conditions, $F(1,23) = .459$, $p = .506$.

We found partial support for our sixth hypothesis, which posited that a wider field-of-view would increase users' perceived difficulty in using the interface. Average scale responses were 2.750 (SD = .886), 2.292 (SD = .576), and 3.000 (SD = .436), in the narrow, wide-angle, and panoramic conditions, respectively. Controlling for gender, we found a marginal main effect of field-of-view on scale responses, $F(2,23) = 3.299$, $p = .058$, $\eta^2 = .184$. Users reported that the panoramic view was significantly more difficult to use than the wide-angle view, $F(1,23) = 6.413$, $p = .020$, $\eta^2 = .179$. A comparison of our results to prior findings is shown in Table 1.

Behavioral Observations

An examination of the gaze data revealed that participants used the periphery to improve their task efficiency. In both augmented view conditions, 12 of 16 of the participants used the peripheral view to scan the environment when transitioning between photograph placements and orienting themselves relative to prominent landmarks. These participants employed a "passive" use of the periphery, where usage was brief and rarely the primary focus of their attention. We therefore classified these participants as *passive* users.

We also identified two other classes of users from the gaze recordings. We categorized two of our participants as *indirect* users, as they never fixated directly on the periphery, but formed accurate mental models of the room and completed the task in less time than average. We classified the remaining two augmented view users as *active* users. These participants spent a large portion of the task fixated on objects in the periphery and actively browsed them while moving. One of these users completed the task slightly faster than average and the other was the second fastest to complete the task overall. While these active users were taking advantage of the full extent of the view offered by the interface, it was not always to their advantage. One of the active users had two collisions, caused when the participant traveled forward with their focus on the peripheral view. Figure 6 shows sample heat maps extracted from gaze recordings demonstrating each class of user.

DISCUSSION

Our results showed that both a wide-angle and a panoramic periphery improved task performance over a non-augmented forward-facing view. From the gaze data, we observed participants employing techniques to take advantage of the peripheral view to orient themselves and improve turning efficiency. These results demonstrate that even in low-bandwidth situations, the additional cues offered by a peripheral camera can improve operator performance and efficiency in mobile collaboration. In addition to improving task performance, our results showed improved situation awareness for participants using a low-quality peripheral view of either width. However, we found no differences in participants' feelings of presence between conditions. Prior literature in presence describes how increased focus in the local environment can decrease feelings of presence in the remote environment [8]. Participants in our study relied heavily on the diagrams of the task goal, which may have reduced feelings of presence across all conditions.

Although we found no differences in task completion time, collisions, or feelings of presence between the wide-angle and panoramic users, we found three main distinctions between the wide-angle and panoramic view conditions.

First, participants in the panoramic view condition made significantly fewer errors when recalling the layout of the room than those in the wide-angle condition. This difference provides support for the hypothesis that wider fields-of-view improve situation awareness via the formation of mental maps. Previous work has shown that impoverished representations from video-feeds can leave out essential cues for building teleop-

Prior result	Comparison	Explanation
Keyhole effect	*Supported*	Increased collisions, slower completion times in narrow view
Cognitive tunneling	*Supported*	Errors in distance/depth judgments increased collisions in narrow view
Wide views increasing cognitive workload	*Supported*	Perceived interface difficulty increased in panoramic condition
Wide views distort velocity perception, reducing driving speed	*Unsupported, Contrasting*	Wide-angle and panoramic views support faster task completion than narrow views
Wider views associated with motion sickness	*Unsupported*	No participants commented on feeling motion sickness
Impoverished video inhibits mental map formation	*Unsupported, Contrasting*	Low-quality periphery improved mental map formation over wide-angle and narrow views

Table 1. A comparison of our results to prior findings.

erators' mental models of the environment [2, 4, 40], but our findings suggest that the addition of low-quality peripheral views can significantly aid in mental map formation.

Second, participants found the panoramic interface to be significantly more difficult to use than the wide-angle interface. We believe that this result supports the theory that wider fields-of-view require greater amounts of cognitive processing to synthesize the larger quantity of visual information provided, leading to perceptions of increased interface difficulty.

Third, we found that participants in the wide-angle condition perceived their collaboration as significantly less successful than in the narrow and panoramic conditions; we suggest two possible explanations for this based on research in organizational behavior that suggests that being vulnerable in a collaboration can increase trust between team members [25].

Dependence. In the narrow-view condition, participants may have had to rely heavily on the confederate to get around. In the panoramic view, participants had significantly more difficulty using the interface than in the wide view, possibly requiring participants to rely more heavily upon the confederate than in the wide-view. Wide-angle view participants may have therefore felt less dependent upon (and less collaborative with) the confederate than in the other two conditions. Future work could explore this mediation effect by measuring interdependent behaviors, vulnerability, and trust between remote and local users.

Collaborative Strategies. Participants in the narrow condition may have employed a collaborative strategy where they explicitly requested assistance from the confederate during the task, asking that he move things into their field of view or aid in image identification, thus building collaborative trust. In the panoramic condition, participants may have engaged in a more implicit collaboration strategy, relying on the confederate to carry out more complex instructions because of their ability to track the state of the room. The participant's ability to verify that instructions were correctly followed may have also aided in building collaborative trust. However, in the wide-angle view, participants may not have engaged fully in either strategy—neither explicitly requesting assistance due to their enhanced view nor trusting in the ability to of the confederate to carry out complex instructions because of their inability to monitor the results. As a consequence, participant perceptions in the success of the collaboration may have been lower than in either the narrow or panoramic views. Future work could explore these potential differences in collaborative strategies by coding recordings for directing speech acts or task behaviors.

Design and Research Implications

Our results offer several implications for research and for the design of camera and display systems for future telepresence robots. This study is a first step toward better understanding the effects of field-of-view on collaborative outcomes for robotic telepresence users in mobile contexts. We found that both augmented conditions experienced collaborative performance gains over the narrow condition by reducing both the number of collisions and the time to complete the task. These results suggest that designers of bandwidth-limited telepresence systems can still significantly improve the performance of remote users by offering low-quality peripheral vision.

While the augmented views improved performance over the narrow view, we detected no significant differences between the augmented views across measures of task completion time, collisions, or feelings of presence. We believe that this lack of differences has two key implications.

First, it opens the door for future work to study the effects that varying angles-of-view might have at a finer granularity. The width of the augmented fields-of-view in our study (180° and 360°) are vastly different. Are there no further collaborative benefits for telepresence users beyond 180° angles-of-view? If not, what is the minimum field-of-view width required to benefit users? Otherwise, is there some performance maximum located within the degree space between 180° and 360°?

Second, it demonstrates how changes that do not affect performance measures can still affect other elements of the collaboration. Our results show that interfaces with wider fields-of-view improve users' ability to form a mental map of the remote environment, but they perceive such interfaces as more difficult to use. In situations where having a correct mental map of the environment is critical—such as doctors using a telepresence robot to visit multiple urgent-care patients—our results suggest that systems should provide wider fields-of-view.

Our observation from the eye-tracking data of distinct user "classes" demonstrates that different types of users have different strategies for maintaining their situation awareness and consequently may benefit from different video configurations. For example, the active users we observed did not appear to suffer from increased cognitive load caused by too much perceptual information in the 180° or 360° views. These users may see even greater performance gains from using a peripheral view with higher-quality video, if bandwidth allows. Users who have already traversed their remote environment and formed accurate mental maps (akin to our "indirect" users), such as regular remote workers, may not require a periphery at all. For these users, it may be a more beneficial use of bandwidth to maximize the resolution of the forward facing camera. While our results demonstrate that the augmented peripheral views of either width supported better performance over a narrow view, these observations suggest that there may not be a "one (view) size fits all" solution for telepresence users in mobile collaborative situations, and that factors such as user familiarity with the remote environment play an important role in selecting the optimal view configuration.

Hollan et al. [14], in their seminal paper, encouraged researchers to not limit communication research by focusing on face-to-face interactions as the gold standard. Our results support this idea, showing that collaborating using a 360° panoramic field-of-view, while unnatural compared to a face-to-face interaction, offers strong advantages for telepresence users operating in new environments, allowing them to form more accurate mental maps. Additionally, this better-than-human view may not fall prey to lower perceptions of collaborative success seen in our wide-angle condition.

Our work opens the door for exploring other facets of telepresence using our methodology by examining the effects of a single interface element. Our results also highlight that designers of these systems may not need to limit themselves into a forced similarity with face-to-face communications.

LIMITATIONS AND FUTURE WORK

There are four key limitations in our study that reveal potential avenues for future research. First, future studies must explore how particular features of the robotic platform affect the view requirements. For example, the telepresence robot used in our study has a zero-degree turn radius, which makes it easy to turn to view objects without traveling laterally. We may find that maintaining awareness of objects behind the robot is of increased importance on a platform which requires more maneuvering to rotate. Second, our study explored the width of the operator's field-of-view for novice users in an unfamiliar location. While this exploration still has value, as it represents situations like someone visiting a museum on a telepresence robot, it may not generalize to all situations. Future studies can explore how view requirements change in known and unknown environments and across varying levels of operator experience. Third, this study used a particular environment and task, but it is possible that richer environments (e.g., with more visual landmarks) or a more engaging task could produce different results, particularly in terms of self-reported situation awareness and presence. Finally, having recruited a small sample from a university campus, there are limitations to the generalizability of our study. Future replication work could be performed using larger samples to study the effects of view width on more diverse populations, such as with remote workers in an actual organization.

CONCLUSION

Our work examined the effects of different widths of operator field-of-view—*narrow* (45°), *wide-angle* (180°), and *panoramic* (360°)—on collaborative outcomes for users of telepresence robots in a mobile context. Our results showed that participants using the wide-angle and panoramic views completed the task more quickly and made fewer collisions compared to the narrow view users. Panoramic view use improved participants' ability to form accurate mental maps of the remote environment over the other views, but participants perceived the interface to be more difficult to use. Participants also perceived their collaboration as significantly less successful in the wide-angle condition compared to the other conditions. Our findings have implications for the designers and researchers of robotic telepresence systems, highlighting that even low-quality augmentations to the operator's field-of-view may aid in supporting collaborative outcomes and user interactions. Telepresence systems are still relatively new and their adoption in a variety of settings is growing, spurred by advances in technology and a desire for the ability to work with others in remote locations. While the first steps toward developing telepresence systems have been taken, we hope that our results aid in informing the designs of a future in which telepresence systems are ubiquitous, and where visiting a museum across the world is as easy as just logging in.

ACKNOWLEDGMENTS

This research was supported by National Science Foundation award 1117652. We would like to thank Brandi Hefty, Jieni Peng, and Catherine Steffel for their help in our work.

REFERENCES

1. Baudisch, P., Good, N., and Stewart, P. Focus plus context screens. In *Proc. UIST'01* (2001), 31–40.
2. Chen, J. Y., Haas, E. C., and Barnes, M. J. Human performance issues and user interface design for teleoperated robots. *IEEE Trans Syst, Man, Cybern C, Appl Rev 37*, 6 (2007), 1231–1245.
3. Darken, R. P., Kempster, M. K., and Peterson, B. Effects of streaming video quality of service on spatial comprehension in a reconnaissance task. In *Proc. I/ITSEC'01* (2001).
4. Darken, R. P., and Peterson, B. Spatial orientation, wayfinding, and representation. In *Handbook of Virtual Environment Technology*, K. Stanney, Ed. 2002, 493–518.
5. DeJong, B. P., Colgate, J. E., and Peshkin, M. A. Improving teleoperation: reducing mental rotations and translations. In *Proc. ICRA'04* (2004), 3708–3714.
6. Desai, M., Tsui, K. M., Yanco, H. A., and Uhlik, C. Essential features of telepresence robots. In *Proc. TePRA'11* (2011), 15–20.
7. Double Robotics. Double, 2014 (accessed Sept. 20, 2014). http://www.doublerobotics.com/.
8. Draper, J. V., Kaber, D. B., and Usher, J. M. Telepresence. *Hum Factors 40*, 3 (1998), 354–375.
9. Drury, J. L., Scholtz, J., and Yanco, H. A. Awareness in human-robot interactions. In *Proc. SMC'03* (2003), 912–918.
10. Endsley, M. R. Measurement of situation awareness in dynamic systems. *Hum Factors 37*, 1 (1995), 65–84.
11. Fels, D. I., Waalen, J. K., Zhai, S., and Weiss, P. Telepresence under exceptional circumstances. In *Proc. Interact'01* (2001), 617–624.
12. Gorawara-Bhat, R., Cook, M. A., and Sachs, G. A. Nonverbal communication in doctor–elderly patient trans. *Patient Educ Couns 66*, 2 (2007), 223–234.
13. Gutwin, C., and Greenberg, S. A descriptive framework of workspace awareness for real-time groupware. *CSCW 11*, 3-4 (2002), 411–446.
14. Hollan, J., and Stornetta, S. Beyond being there. In *Proc. CHI'92* (1992), 119–125.
15. Howard, I. P. *Binocular vision and stereopsis*. Oxford University Press, 1995.
16. Ishiguro, H., and Trivedi, M. Integrating a perceptual information infrastructure with robotic avatars. In *Proc. IROS'99* (1999), 1032–1038.

17. Johnson, S., Gibson, M., and Mutlu, B. Handheld or handsfree? Remote collaboration via lightweight head-mounted displays and handheld devices. In *Proc. CSCW'15* (2015).

18. Jouppi, N. P. First steps towards mutually-immersive mobile telepresence. In *Proc. CSCW'02* (2002), 354–363.

19. Keyes, B., Casey, R., Yanco, H. A., Maxwell, B. A., and Georgiev, Y. Camera placement and multi-camera fusion for remote robot operation. In *Proc. of IEEE Workshop on Safety, Security and Rescue Robotics* (2006), 22–24.

20. Kraut, R. E., Miller, M. D., and Siegel, J. Collaboration in performance of physical tasks: Effects on outcomes and communication. In *Proc. CSCW'96* (1996), 57–66.

21. Kress, G., and Almaula, H. Sensorimotor requirements for teleoperation. *FMC Corporation, San Diego, CA, Report R-6279* (1988).

22. Kristoffersson, A., Eklundh, K. S., and Loutfi, A. Measuring the quality of interaction in mobile robotic telepresence: a pilot's perspective. *Int J Soc Robot 5*, 1 (2013), 89–101.

23. Lee, M. K., and Takayama, L. Now, I have a body: Uses and social norms for mobile remote presence in the workplace. In *Proc. CHI'11* (2011), 33–42.

24. McGovern, D. E. Experiences in teleoperation of land vehicles. *Spatial Displays and Spatial Instruments* (1989).

25. Meyerson, D., Weick, K. E., and Kramer, R. M. Swift trust and temporary groups. In *Trust in organizations: Frontiers of theory and research*, vol. 166. 1996, 195.

26. Michaud, F., Boissy, P., Labonte, D., Corriveau, H., Grant, A., Lauria, M., Cloutier, R., Roux, M.-A., Iannuzzi, D., and Royer, M.-P. Telepresence robot for home care assistance. In *Multidisciplinary Collaboration for Socially Assistive Robotics* (2007), 50–55.

27. Osaka, N. Speed estimation through restricted visual field during driving in day and night: naso-temporal hemifield differences. In *Proc. VIV II* (1988).

28. Paulos, E., and Canny, J. Delivering real reality to the world wide web via telerobotics. In *Proc. ICRA'96*, vol. 2 (1996), 1694–1699.

29. Rae, I., Mutlu, B., and Takayama, L. Bodies in motion: Mobility, presence, and task awareness in telepresence. In *Proc. CHI'13* (2013), 1921–1930.

30. Rae, I., Venolia, G., Tang, J., and Molnar, D. A framework for understanding and designing telepresence. In *Proc. CSCW'15* (2015).

31. Revolve Robotics. Kubi, 2014 (accessed Sept. 20, 2014). http://www.kubi.me/.

32. Scholtz, J. C. Human-robot interactions: Creating synergistic cyber forces. *Multi-Robot Systems: From Swarms to Intelligent Automata* (2002), 177–184.

33. Schubert, T., Friedmann, F., and Regenbrecht, H. The experience of presence: Factor analytic insights. *Presence 10*, 3 (2001), 266–281.

34. Scribner, D. R., and Gombash, J. W. The effect of stereoscopic and wide field of view conditions on teleoperator performance. *DTIC Document* (1998).

35. Shiroma, N., Sato, N., Chiu, Y.-h., and Matsuno, F. Study on effective camera images for mobile robot teleoperation. In *Proc. RO-MAN'04* (2004), 107–112.

36. Smyth, C. C. Indirect vision driving with fixed flat panel displays for near unity, wide, and extended fields of camera view. In *Proc. HFES'00* (2000), 541–544.

37. Suitable Technologies. Beampro, 2014 (accessed Sept. 20, 2014). https://www.suitabletech.com/beam/.

38. Thomas, G., Robinson, W. D., and Dow, S. Improving the visual experience for mobile robotics. In *Proc. Seventh Annual Iowa Space Grant* (1997), 10–20.

39. Thomas, L. C., and Wickens, C. D. Effects of display frames of reference on spatial judgments and change detection. *DTIC Document* (2000).

40. Tittle, J. S., Roesler, A., and Woods, D. D. The remote perception problem. In *Proc. HFES'02*, vol. 46 (2002), 260–264.

41. Tsui, K. M., Desai, M., Yanco, H. A., and Uhlik, C. Exploring use cases for telepresence robots. In *Proc. HRI'11* (2011), 11–18.

42. Tsui, K. M., and Yanco, H. A. Assistive, rehabilitation, and surgical robots from the perspective of medical and healthcare professionals. In *Proc. AAAI Workshop on Human Implications of Human-Robot Interaction* (2007).

43. Van Erp, J. B., and Padmos, P. Image parameters for driving with indirect viewing systems. *Ergonomics 46*, 15 (2003), 1471–1499.

44. Venolia, G., Tang, J., Cervantes, R., Bly, S., Robertson, G., Lee, B., and Inkpen, K. Embodied social proxy. In *Proc. CHI'10* (2010), 1049–1058.

45. Voshell, M., Woods, D. D., and Phillips, F. Overcoming the keyhole in human-robot coordination. In *Proc. HFES'05* (2005), 442–446.

46. Wickens, T. D., and Keppel, G. Design and analysis: a researchers handbook, 2004.

47. Witmer, B. G., and Sadowski, W. J. Nonvisually guided locomotion to a previously viewed target in real and virtual environments. *Hum Factors 40*, 3 (1998), 478–488.

48. Woods, D. D., Tittle, J., Feil, M., and Roesler, A. Envisioning human-robot coordination in future operations. *IEEE Trans Syst, Man, Cybern C, Appl Rev 34*, 2 (2004), 210–218.

49. Yanco, H. A., and Drury, J. "Where am I?" acquiring situation awareness using a remote robot platform. In *Proc. SMC'04* (2004), 2835–2840.