Comparison of Display Modality and Human-in-the-Loop Presence for On-Orbit Inspection of Spacecraft

Hannah Weiss, University of Michigan, Ann Arbor, USA, Andrew Liu, Massachusetts Institute of Technology, Cambridge, USA, Amos Byron, Jonathan Blossom, California Institute of Technology, Pasadena, USA, and Leia Stirling, University of Michigan, Ann Arbor, USA

Objective: To investigate the impact of interface display modalities and human-in-the-loop presence on the awareness, workload, performance, and user strategies of humans interacting with teleoperated robotic systems while conducting inspection tasks onboard spacecraft.

Background: Due to recent advancements in robotic technology, free-flying teleoperated robot inspectors are a viable alternative to extravehicular activity inspection operations. Teleoperation depends on the user’s situation awareness; consequently, a key to successful operations is practical bi-directional communication between human and robot agents.

Method: Participants (n = 19) performed telerobotic inspection of a virtual spacecraft during two degrees of temporal communication, a Synchronous Inspection task and an Asynchronous Inspection task. Participants executed the two tasks while using three distinct visual displays (2D, 3D, AR) and accompanying control systems.

Results: Anomaly detection performance was better during Synchronous Inspection than the Asynchronous Inspection of previously captured imagery. Users’ detection accuracy reduced when given interactive exocentric 3D viewpoints to accompany the egocentric robot view. The results provide evidence that 3D projections, either demonstrated on a 2D interface or augmented reality hologram, do not affect the mean clearance violation time (local guidance performance), even though the subjects perceived a benefit.

Conclusion: In the current implementation, the addition of augmented reality to a classical egocentric robot view for exterior inspection of spacecraft is unnecessary, as its margin of performance enhancement is limited in comparison.

Application: Results are presented to inform future human–robot interfaces to support crew autonomy for deep space missions.

Keywords: teleoperation, interface evaluation, assistive technologies, situation awareness, augmented reality

Address correspondence to Hannah Weiss, Industrial and Operations Engineering Center for Ergonomics, University of Michigan, Ann Arbor, MI 48109-1382, USA; e-mail: hanweiss@umich.edu

Author(s) Note: The author(s) of this article are U.S. government employees and created the article within the scope of their employment. As a work of the U.S. federal government, the content of the article is in the public domain.

HUMAN FACTORS
Vol. 00, No. 0, Month XXXX, pp. 1-15
DOI: 10.1177/00187208211042782
Article reuse guidelines: sagepub.com/journals-permissions

BACKGROUND
Teleoperation, the remote control of robots through a visual display and control interface, remains the dominant control paradigm for human interaction with robotic systems in complex and hazardous environments. Various applications include aerial and aquatic unmanned vehicles for urban search and rescue (Jones et al., 2011) and underwater exploration (Xu et al., 2019), surgical robots (Okamura, 2004), nuclear plants (Fong & Thorpe, 2001), and planetary surface exploration (Biesiadecki et al., 2007). The current study focuses on the teleoperation of a free-flying robot inspector for on-orbit space operations and considers how interface design and human–robot synchronicity influence anomaly detection and navigation. Human–robot interaction is a vital component of current and future space exploration. Long-duration deep-space exploration missions (LDEMs) will require astronaut crews to operate more autonomously from ground support and rely on complex robotic systems to perform critical tasks such as spacecraft inspection. Current low-earth orbit missions and extended exploration to the lunar surface necessitate the restriction of extravehicular activity for astronaut safety. Understanding human–robot interaction requirements in the context of spacecraft inspection is beneficial for all mission types regardless of duration or proximity to earth.

Synchronous teleoperated control tasks where the human and robot are spatially separated without direct visual contact typically require continuous monitoring of multiple displays, time-sharing between tasks, and may include decision making under time pressure and uncertainty (Almeida et al., 2020; Chen...
Due to the operator’s remotely located position relative to the robot, natural human perceptual processing abilities are decoupled from the physical environment posing considerable challenges to the operator’s situation awareness (SA), that is, their understanding of information, events, and actions that impact mission objectives in dynamic environments (Endsley, 2000; Wickens, 2008). Situation awareness is defined within three hierarchical phases: perception of elements in the environment, comprehension of the current situation for rapid decision making, and projection of future status for short-term planning (Endsley, 1995). Prior work has identified common factors related to interface design that affect remote perception and manipulation, depending upon these SA levels. The limited field of view (FOV) of assistive cameras has been identified as the primary factor affecting the judgment of vehicle speed, collision avoidance, robot orientation, and future vehicle states (Chen et al., 2007). Suggested solutions include alterations to interface design through multiple camera viewpoints, adjustable field of view, and restrictions to minimum image frame rate and time delays (Thomas & Wickens, 2001). However, it is recognized that specific display formats are better suited for specific tasks (Fong & Thorpe, 2001). For example, egocentric information (from the robot’s perspective) best supports local navigational guidance (maintenance of desired navigational path), whereas exocentric information (observing from an external perspective) promotes global awareness, the knowledge of objects in space concerning one’s position and orientation (Barfield et al., 1995; Ferland et al., 2008; Wickens & Prevett, 1995). While much of the research regarding operational frame of reference has been conducted in the aviation domain, recent work on drones has demonstrated similar results with improved obstacle avoidance using an exocentric perspective compared to the drone’s egocentric perspective (Cho et al., 2016). The maintenance of the operator’s SA through appropriate interface design is imperative to effective teleoperation of robotic systems, particularly in space operations where poor navigation can damage mission-critical hardware.

Traditional interfaces rely on computer monitors, which require the operator to synthesize 2D projections of 3D environments, often challenging the user’s ability to build accurate mental models, a prerequisite for achieving SA (Sarter & Woods, 1991). Recent advancements in consumer-grade extended reality (XR) technologies, such as augmented reality (AR) and virtual reality (VR), enable novel forms of intuitive 3D visual communication between humans and machines. Researchers have demonstrated VR as a viable alternative to traditional monitor interfaces with advantages to task completion times, lower cognitive workload, and higher usability scores in dexterous robot manipulation tasks (Hetrick et al., 2020; Whitney et al., 2020). AR implementation in collocated robot teleoperation and robot path planning via waypoint control have exhibited advantages over traditional methods with objective and subjective improvements in navigation and task completion, especially for novice users (Fang et al., 2014; Walker et al., 2019). XR provides opportunities to support mental model formation through augmented temporal and spatial awareness. However, a deeper understanding of how XR impacts robotic teleoperation across different applications and tasks, such as space exploration specific operations, is required to develop guidelines that inform the engineering and design processes of complex robotic systems, which are dependent on effective human–machine interaction.

The National Aeronautics and Space Administration (NASA) Human Integration Design Handbook provides interface and interaction guidelines but currently lacks guidance on XR design criteria for display devices. During increasingly earth-independent missions, evidence is needed to evaluate and set standards for human–automation–robotic teams. While visual perception has been studied for computer-based systems (e.g., icon density, color contrast, and luminance; Shen et al., 2015), it is not clear how perception generalizes to new modalities and tasks, such as AR in the context of an inspection task. It is widely acknowledged that visual attention patterns...
and fixations are task-dependent (Rothkopf et al., 2007; Tatler et al., 2010), highlighting the importance of both the scene and task goals as factors for design decisions to support perception performance. To inform best practices in function allocation and communication modalities, it is essential to characterize human performance involving mixed-agent teams and automation in exploration mission simulations.

The present study evaluates display and human-in-the-loop presence on an operator’s performance and strategy during the telerobotic inspection of spacecraft to address the aforementioned gaps, specifically the appropriate allocation of functions across human–robot teams and the effect of display modality on spacecraft inspection operations to ultimately inform design criteria. Synchronous Inspection requires real-time human input, whereas Asynchronous Inspection requires the operator to examine previously captured images without the need for direct satellite control. The degree of human-in-the-loop presence is essential for mission operations planning. As robotic technology advances, there is a potential for fully autonomous free-flying inspectors with operators reviewing captured imagery. In an urban search and rescue application, synchronous telerobotic control yielded better victim search performance than asynchronous control (Velagapudi et al., 2008), suggesting the importance of operator attention during the data-collection phase.

The simulated telerobotic inspection considered in this study requires the operator to detect anomalies on the exterior of spacecraft when the inspector information is presented as (1) 2D camera images, (2) a 3D reconstruction on a 2D projection, and (3) a 3D reconstruction in an AR environment. The anomaly detection task is considered during a Synchronous Inspection task and an Asynchronous Inspection of previously collected imaging. The research considered two main questions: (1) how the operator utilized each interface, and (2) how effective this utilization was in completing task goals. It was hypothesized that the AR and 3D interfaces would (1) lead to better clearance (at least 2 m distance) between the inspector and the station, (2) enable a higher percentage of the station to be inspected, and (3) result in greater anomaly detection accuracy. For Synchronous Inspection, whereby operators could select automatic or manual command modes, it was expected that (4) participants would employ both modes to optimize performance. We hypothesized that the (5) 3D and AR interfaces would result in longer session lengths within task and across task, specifically for the Synchronous Inspection where teleoperation would require more cognitive demand. Lastly, (6) between tasks, we hypothesized that anomaly detection performance would be superior during Synchronous Inspection. This study will inform the engineering and design process of these complex systems from the interface and mission planning perspectives to optimize telerobotic inspection operations.

**METHODS**

**Participants**

Nineteen University of Michigan students participated in this experiment (14 males, females; mean age 24.26 years, SD = 3.07). This research complied with the American Psychological Association Code of Ethics and was approved by the University of Michigan Institutional Review Board for Health Sciences and Behavioral Sciences (UM IRB-HSBS). All subjects reported normal to corrected-to-normal vision and reported no musculoskeletal, auditory, or vestibular disorders. The subjects were informed of the purpose and procedures before the training session and signed the informed consent approved by the UM IRB-HSBS. Twelve subjects reported previous limited experience with XR (i.e., VR or AR) and eight with robotic platforms (i.e., drones or flight simulators); the remaining subjects were considered novice users.

Each participant performed a 2-day protocol, in which the first day consisted of training and familiarization with the software, and the second day comprised evaluative testing. Upon admittance, each participant was randomly assigned a contextually relevant mission role, either commander, flight engineer, or mission specialist 1 or 2. The designation allowed for the systematic order-based randomization of
independent variables between participants. Two subjects’ data for the dependent variables (session length and anomalies correctly detected) were omitted from the analysis due to software malfunctions.

Software and Hardware Configuration

The simulated environment incorporated an inspector with a single front-facing camera (60º FOV) and a space station with distinct module geometries and various attachments (e.g., docked spacecraft, engine, and communication antenna; Figure 1). The design mimics characteristics of the International Space Station (ISS) and presents various degrees of collision risk (Figure 1). Collisions could occur; however, proximity warnings were displayed when the satellite was maneuvered within 2 meters of the station’s exterior.

A series of 7–10 simulated anomalies, varying in size (3), edge characteristics (3), and location, were mapped to the station’s exterior. Twelve distinct anomaly sets of similar difficulty were generated and randomly assigned between the trials. The anomalies remained hidden until uncovered by the inspector’s spotlight, which corresponded with the inspector camera’s FOV. The anomalies remained visible once uncovered. The operators were instructed to tag all anomalies detected using a bounding box with the interface provided (Figure 1).

During Synchronous Inspection, the inspector could be flown in automatic mode along a predetermined path or manual mode to move off the nominal path. Inspector motion was limited to a 2D plane through the vehicle’s center. Unlimited toggling between control modes were permitted. The inspector was controlled with a 3-DOF joystick (Logitech Extreme 3D Pro), including 12 programmable buttons, an eight-way hat switch, and a trigger. Four programmable buttons on the joystick were used to enter picture mode, capture pictures, resize the bounding box, and change command mode. The hat switch allowed fixed manipulation of the inspector’s camera view. The Synchronous Inspection controls included direct control of the inspector’s position around the station’s circumference, pitching the inspector’s camera up/down to inspect distal features, and capturing images of anomalies. During Asynchronous Inspection, playback commands included starting/stopping video feedback, manipulating the viewpoint, and capturing images.

A combination of joystick and keyboard controls were utilized for teleoperation and anomaly tagging; however, additional input controls were introduced in 3D and AR interfaces (Figure 2). The 3D projection was manipulated by a mouse, while AR gestures controlled the holograms. In AR, the standard HoloLens gestures were used to select, resize, relocate, and rotate the holograms.

The interfaces were equipped with the inspector’s egocentric view on the left monitor and the respective display modality on the right monitor or AR headset. The egocentric view was either synchronous with operator commands

Figure 1. Left: Virtual space station and inspector. Right: Egocentric inspector robot view with the picture mode activated for anomaly detection via a red bounding box.
Comparison of Display and Human Presence

or asynchronous interactive video playback or reconstructed model of previously collected inspector data. The right auxiliary monitor featured inspection metrics (proximity of inspector to the station, number of pictures taken, control mode, total task time, and time elapsed) and the additional viewpoints to augment the egocentric view.

The 2D and 3D interfaces were displayed on two 24” computer monitors (Figures 2 and 3). In 2D configuration, four fixed camera views (60° FOV) mounted on the space station’s exterior provided exocentric but not comprehensive views of the space station. The 3D configuration featured an interactive 3D model of the station reconstructed on the 2D monitor.

The AR configuration holographically displayed a custom-built simulation designed in Unity on a Microsoft HoloLens 2, with eye-tracking and gesture recognition to allow the user to interact with the augmented environment. During AR tasks, the space station was anchored in the viewer’s physical space, allowing participants to move around the hologram and reposition, scale, and rotate the hologram relative to their physical environment (Figure 4).

Figure 2. Experimental hardware and desktop configuration with two monitors in the 3D configuration, AR headset (left), keyboard, mouse, and joystick (right).

Experimental Design and Protocol

The independent variables in this study were interface (2D, 3D, AR) and task type (Synchronous Inspection, Asynchronous Inspection). In all conditions and tasks, participants were instructed to inspect the entire exterior of the spacecraft and identify and mark all of the anomalies within 15 min. Detected anomalies were documented by taking a picture of the anomaly within an appropriately resized bounding box where the anomaly size was roughly equal to one quadrant of the bounding box. It was discovered that the user captured images were cropped from the monitor’s original aspect ratio, and subsequently 37 out of 307 total pictures could not be formally assessed. Our participants were instructed to center the anomaly for image capture; however, several failed to do so leading to anomalies located in the cropped regions of the images. The images in which the bounding box and anomaly were not visible were excluded from the anomaly detection results. In Synchronous Inspection, participants were required to balance aspects of inspection as well as navigation by flying the inspector robot around the spacecraft’s circumference with dynamic control capabilities. They were instructed
to avoid collisions by maintaining a 2-meter distance between the station and satellite. Navigation strategies needed to support competing goals of maintaining a safe distance due to concern of damage as well as being close enough to recognize anomalies. In Asynchronous Inspection,
participants analyzed imagery captured during a previous inspection flight, so navigation constraints such as the 2-meter rule, were not relevant. In all training and testing sessions, participants completed the Synchronous Inspection first followed by Asynchronous Inspection, where the interface order was balanced across subjects. Each participant performed the anomaly detection task six times, once per interface and once per task. After each task, participants answered subjective surveys of system usability, perceived workload, and situation awareness.

**Performance Measures**

We considered standard measures for human–robot interaction performance to evaluate our hypotheses, including task completion time, percentage of station inspected, distance to station, and detection accuracy (Steinfeld et al., 2006). To address our first hypothesis regarding navigational clearance performance, we selected distance metrics describing the portion of the session length the inspector was within 2 meters of the station and the average, minimum, and maximum distances. The percentage of the station inspected addressed our second hypothesis with respect to the effects of interface type during synchronous inspection. Our third and sixth hypotheses relate to the primary performance metric of anomaly detection accuracy and were addressed through the percent-correct score of anomalies identified and documented by the user. For Synchronous Inspection, participant strategy was assessed by the inspector path length, time within control modes, and picture mode, which served to evaluate our fourth hypothesis. Inspector idle time, defined as periods when the inspector was stationary and not in picture mode, was logged along with total session length to evaluate our fifth hypothesis.

The NASA Task Load Index (TLX) survey captured perceived workload scores during each task without differentiation of interface. The TLX scores aimed to address differences in perceived workload with respect to human-in-the-loop presence rather than address differences between interface. A custom Qualtrics survey administered immediately following all three interfaces for a given task considered subjective comparisons of the interfaces. Participants indicated which station regions were challenging to inspect across tasks and interfaces and concluded the questionnaires by ranking their interface preferences within task type and a brief explanation of their top choice.

**Statistical Analysis**

Session length and the number of anomalies detected were assessed using an n-way Analysis of Variance (ANOVA) with factors Subject (random), Task (fixed), and Interface (fixed). Dependent variables specific to Synchronous Inspection (i.e., average and peak inspector speed, time stopped, path length, coverage map percentage, control mode, and distance metrics) were assessed using a two-way ANOVA with factors Subject (random) and Interface (fixed). If a significant effect was found in the ANOVA, post-hoc paired t-tests (with Bonferroni correction) were performed to assess the effect. Eleven post-hoc evaluations were considered: the three comparisons within task, three across task with the same interface, and all five combinations with inspection 2D (the baseline). Only the raw TLX scores from each of the subscales were obtained between tasks then assessed using dependent t-tests. K-means clustering was applied to the Synchronous Inspection control modes to determine whether groups of subjects adopted common control strategies. The optimal value of K clusters was determined using the Calinski-Harabasz clustering criterion with a maximum of six clusters.

**RESULTS**

**Teleoperation Strategies**

During Synchronous Inspection, the 2D interface yielded faster average inspector speeds and less inspector idle time (Figure 5). There was a significant main effect of the interface type on the average speed of the inspector during Synchronous Inspection ($F(18,36) = 13.37, p < .0005$). Post-hoc paired t-tests showed that the average inspector speed for the 2D interface was significantly faster than the 3D interface ($t(18) = 4.52, p < .0005$) and AR interface ($t(18) = 4.21, p < .005$). There was a significant main effect of the interface type on the inspector idle time ($F(18,36) = 20.76, p < .0005$). Post-hoc paired t-tests showed that the inspector’s idle time was significantly less for
the 2D interface as compared with the 3D interface ($t(18) = -5.54, p < .0005$) and AR interface ($t(18) = -5.88, p < .0005$). There was no significant effect of interface on the peak speed, inspector path length, control mode utilization (manual mode, auto mode, picture mode), or the station’s total percentage uncovered by the inspector’s spotlight (Figure 5).

The participants successfully maintained 2 meters from the space station across interfaces with less than 3.54% (pooled across displays) of the session time flying the inspector closer than 2 meters. There was a significant main effect of the interface type on the average minimum distance of the inspector during Synchronous Inspection ($F(18,36) = 3.51, p < .05$). Post-hoc paired $t$-tests support that the average minimum distance for the 3D interface was significantly longer than the 2D interface ($t(18) = 2.44, p < .05$) and AR interface ($t(18) = 2.19, p < .05$). There was no significant interface effect on the average distance, the maximum distance, and the total time within 2 meters of the station.

The 2D interface yielded faster session lengths across task types and superior anomaly detection accuracy in Synchronous Inspection than all other combinations (Figure 6). There was a significant interaction effect between task and interface on the session length ($F(18,89) = 6.64, p < .005$). Within Asynchronous Inspection, the session lengths for all three interfaces were significantly different (2D vs. 3D $t(18) = -3.95, p < .005$, 2D vs. AR $t(17) = -6.20, p < .0005$, and 3D vs. AR $t(17) = -5.73, p < .0005$). Within Synchronous Inspection, the session lengths for the 2D and AR interfaces were significantly different ($t(18) = -3.31, p < .005$), as well as the 3D and AR interfaces ($t(18) = -3.33, p < .005$). There was no significant difference in the session length for the 2D versus 3D interface within Synchronous Inspection. A significant interaction of task and interface on the percent correct for anomaly detection was also seen ($F(18,90) = 7.51, p < .005$). The post-hoc paired $t$-tests within and across tasks revealed that the mean response for Synchronous Inspection 2D was significantly different from the mean response for the other group combinations. There were no other significant differences within or across task.

Figure 5. Synchronous Inspection task metrics. (a) Average inspector speed. (b) Portion of session length when the inspector was idle. (c) Minimum distance between inspector and station. (d) Coverage map percentage where 71% was the maximum coverage achievable due to the motion restriction within a 2D plane. Statistically significant contrasts notated with an asterisk (*).
The k-means clustering yielded four comparable strategies for automation usage with the 3D and AR interfaces (Figure 7). The 2D interface resulted in five clusters, one additional strategy characterized by little automatic mode utilization and less manual mode utilization due to more time spent in picture mode. The four strategies range from no automatic mode utilization to minimal manual mode utilization while navigating around the spacecraft. The middle two clusters are characterized by varied emphasis between automatic and manual mode utilization. The remaining portion of control time accounts for the inspector’s time in picture mode. The number of subjects who flew the inspector primarily in automatic mode (greater
than 50% of session length) was the largest for the 3D interface (subjects), followed by the 2D interface (subjects), and lastly, the AR interface (subjects). Across the interfaces, participants maintained a similar control mode utilization strategy as their behavior remained within the same or the next nearest cluster.

**NASA TLX**

There was a significant decrease in the mean perceived physical demand while performing Synchronous Inspection ($M = 54.21$, $SD = 24.05$) compared with Asynchronous Inspection ($M = 46.84$, $SD = 19.52$), $t(18) = -2.37, p < .05$. There was no significant difference between tasks for the perceived mental demand, temporal demand, performance, effort, or frustration (Figure 8). Interestingly, the raw TLX scores yielded low levels of perceived performance (rated out of 100%) in both Synchronous ($M = 27.63$, $SD = 19.53$) and Asynchronous ($M = 30.26$, $SD = 26.43$) Inspection.

**Posttrial Questionnaire**

The space station features that protrude in the zenith and nadir directions were reported as the most challenging regions to inspect across task types, especially with the 2D interface. For both the Synchronous and Asynchronous Inspection tasks, 79% of the subjects chose the AR interface as their preferred interface to complete task goals. The 3D interface followed as the second most preferred for both tasks (21%). No participants listed the 2D interface as their preferred interface. The participants indicated four reasons supporting their preference for the AR and 3D interfaces: better spatial awareness, ease of identifying satellite location, collision avoidance, and determining how the satellite would respond to control input. Criticism of the AR interfaces included difficulty manipulating the hologram and lens artifacts affecting the user’s ability to see detail. For the 2D interface, subjects reported actively ignoring the 2D interface’s fixed camera views and focusing their attention on the egocentric view. One subject reported difficulty manipulating the 3D interface.

**DISCUSSION**

The research considered two main questions: how the operator utilized each interface and how effective this utilization was in completing task goals. We hypothesized that the AR and 3D
Comparison of Display and Human Presence

interfaces would result in (1) better clearance, (2) enable a higher percentage of the station being inspected, and (3) better anomaly detection accuracy. We hypothesized that within Synchronous Inspection, (4) the participants would employ both control modes to optimize performance. Due to the interactive quality of the 3D and AR views, we expected (5) longer session lengths for these interfaces. Finally, we expected (6) improvements in anomaly detection performance during teleoperation with synchronous human input.

Our results provide mixed support for our first hypothesis, which looked at local situation awareness through the metric of better clearance, which was defined as the maintenance of 2 meters from the space station and the inspector for collision avoidance. In the post-trial surveys, most participants indicated that the 3D models from the 3D and AR interfaces supported more straightforward navigation with the ability to determine satellite location relative to the spacecraft. The 3D interface yielded extra caution when locally navigating with a significantly higher average minimum distance. However, there was no evidence of a difference in average distance to the station, and the total time within 2 meters of the station suggests that the 3D and AR interfaces did not considerably improve user’s local navigation. In the context of this paper and the referenced literature (Barfield et al., 1995; Ferland et al., 2008; Wickens & Prevett, 1995), local navigation refers to the maintenance of the desired navigational path. Conversely, global navigation refers to the knowledge of objects in space concerning one’s position and orientation. While the local navigation SA performance was not improved, global SA performance was not evaluated in this study. Participants implemented different teleoperation strategies resulting in large variations in path length, where nine subjects consistently inspected the station in a single rotation and the remaining ten subjects with two rotations. Similarly, participants used manual mode to support additional station coverage beyond the automatic capabilities; however, the median coverage did not vary across interface, and thus hypothesis 2 was not supported.

In practice, participants utilized the additional viewpoints afforded by the 3D and AR interfaces to identify anomalies before flying the inspector to specific locations for further inspection, as evident by verbal comments and the increased inspector idle time and session lengths. However, even though the anomalies were visible on both the egocentric robot view and the 3D views, anomaly detection was not better using the 3D and AR interfaces. Literature supporting hypothesis 3 suggests improved search performance and object identification with multiple camera viewpoints and separate cameras controlled independently from the robot’s orientation (Chadwick, 2005; Hughes & Lewis, 2005). Unexpectedly, Synchronous Inspection task performance using the 2D interface yielded significantly higher anomaly detection accuracy within and across tasks, and thus hypothesis 3 was not supported in the present study. The results suggest that the participant’s detection abilities were superior when their full attention was devoted to the egocentric view. It is essential to reiterate that numerous subjects reported deliberately ignoring the supplemental screen on the 2D interface to focus their sole attention on the egocentric view, resulting in participants flying closer to the station on average. This response could be an example of cognitive tunneling where a single view captures the operator’s attention, and the other views are actively ignored to reduce workload (Thomas & Wickens, 2001). However, the information provided on the supplemental monitor (inspector speed, distance from station, control mode, and time elapsed) was consistent across interfaces, and only the distance from station information was needed for manual control navigation; therefore, cognitive overload is not likely. The egocentric robot camera view may simply provide a closer proximity view to detect anomalies with lower motor and attentional demands than the 3D and AR views. As a result, users may not have exploited sizing and scaling capabilities in the exocentric views to the degree of the egocentric view. In our study, collision avoidance and anomaly detection were prioritized, and this may have influenced the participant to rely more heavily on the 3D and AR views for navigational cues rather than anomaly.
detection. These results fit the trade-off pattern observed by Wickens and Prevett (1995) in which higher global SA performance is coupled with exocentric displays, and higher local navigation performance is associated with egocentric displays. Furthermore, our study shows that exocentric views in conjunction with traditional egocentric views did not improve local guidance and anomaly detection, at least for the dynamic environment considered here.

For Synchronous Inspection, in which operators could select command modes, it was expected that participants would employ both modes to optimize performance. The results yielded user strategies ranging from zero utilization of automation to a large dependence on automation. The clustering results provided mixed support for the fourth hypothesis, indicating that most of the subjects employed both automation and manual mode to optimize performance. Our results contrast previous studies that found user’s generally preferred exclusive command modes when given flexible command capabilities (Todd et al., 2020; Wang & Lewis, 2007). Several subjects in our study did not use automation even though the system was highly trustworthy. The majority of these subjects noted manual command as more enjoyable; the remaining mentioned the automation moved too fast for inspection. Implementing adjustable automation speeds may shift the command mode strategies by encouraging greater automation use and also improving proximity maintenance as a secondary response. The utilization of manual control increases collision risk; however, automation has other emergent considerations due to the dynamics of trust, overreliance, and unexpected failures (Lee & See, 2004). The varying command strategies identified suggest that future human-in-the-loop teleoperation interfaces necessitate flexibility in automation command capabilities and considerations for speed and control engagement to achieve task goals.

In Asynchronous Inspection, the 3D and AR interfaces yielded longer session lengths due to increased interaction with the 3D models, supporting our fifth hypothesis. The session lengths were shorter for Asynchronous Inspection as a whole; however, based on the anomaly detection results, the Asynchronous Inspection also yielded worse detection performance supporting our sixth hypothesis. These results suggest that despite inspection time being longer, human presence through synchronous anomaly detection is recommended for operations. These findings are consistent with a similar study considering synchronous and asynchronous video in multi-robot search and rescue exhibited better victim search performance with synchronous systems (Velagapudi et al., 2008). For mission operation designs where human-in-the-loop presence is not feasible, methods to support anomaly detection performance will need to be evaluated.

Limitations and Future Work

This study did not evaluate the movement of the inspector in all 6 degrees of freedom but rather constrained inspection to a 2D plane. For movement in 3D over the entire surface of the space station, the 3D and AR implementations may prove to be more helpful in inspection and navigation. Each subject was allotted 2 hr for training on the system 1 day before evaluative testing. While participants were able to complete the task and acquire practical knowledge of the system in this timeframe, we recognize real-life analogous systems such as the ISS Remote Manipulator System (Currie & Peacock, 2002) require extensive training. Future considerations for increased training protocols and learning effects over extended utilization of the system should be considered. The AR system we used features a wider field of view than other market devices; however, the field of view was still considerably restricted compared to normal human vision. While participants were comfortable manipulating the system, task timing can be influenced by the interface used. For example, Yau et al. (2008) observed effects of input devices (touchscreen and three trackball designs) and motion types on a cursor positioning task, revealing that touch screen performance was the fastest, but less accurate than trackball inputs across movement types. Thus, there is both a need and opportunity to further characterize AR task performance based on the input mappings, as well as the visual
field available, to improve integration of this modality for human–automation–robotic teams. Despite recruiting from a specific demographic (students with an engineering background), we acknowledge that the participants are not fully representative of the education, capabilities, experience, and other qualities of astronaut populations (for which the system is intended). The simulation software will be further tested in a longitudinal study at NASA’s Human Exploration Research Analog (HERA) facility, serving as a high-fidelity LDEM analog and enabling evaluation of differences in inspection strategies over prolonged use. To address gaps revealed in the present study, future efforts should evaluate global SA measures, 3D inspection without movement constraints, and more complex spacecraft structures.

**CONCLUSION**

Current and future space missions require the identification of information needs, function allocation, and the effect of display modality in human–automation–robotic systems. These design concepts, including novel uses of XR, are not currently included in NASA’s Human Information Design Handbook. Our research characterized teleoperation performance to support defining design guidance for AR in mission display devices. This study examined the use of three distinct display interfaces and considered human-in-the-loop presence for performing the critical task of spacecraft inspection for micrometeoroid or orbital debris impact sites. For the measures we analyzed, the results do not suggest a benefit of AR over the 2D egocentric view for improving local situation awareness. While there may be AR benefits for other global navigation and performance metrics, we did not see a benefit of adding AR for anomaly detection operations in the context of 2D planar motion within 3D operation spaces. In future studies, implementing global SA criteria and evaluation, such as the situation awareness global assessment technique (SAGAT; Endsley, 1988), could further evaluate this balance. Although there was no evidence for an AR benefit in these inspection operations, alternative space-specific applications for AR include intravehicular virtual assistants (Braly et al., 2019; Helin et al., 2018) to replace reliance on mission control, head-up displays for the spacesuit architectures (Mitra, 2018), and augmented training capabilities in ground operations (Nuernberger et al., 2020). From a mission operations perspective, our results suggest that when feasible, synchronous human presence is the preferred method for spacecraft inspection. This study informs the engineering and design process of these complex systems from the interface and mission planning perspectives to support teleoperation inspection operations.

**ACKNOWLEDGMENTS**

This study was supported by the National Aeronautics and Space Administration (NASA) Human Research Program Award 80NSSC20K0409.

**KEY POINTS**

- The advantages and disadvantages of three distinct visual displays and human-in-the-loop presence were evaluated for a simulated exterior inspection of spacecraft in nominal and off-nominal scenarios.
- Detection accuracy was the highest for the 2D display for the Synchronous Inspection task. Additional interactive 3D viewpoints decreased detection accuracy and increased task completion time. Augmented Reality provided no significant improvement to local navigation, i.e., minimum distance to station or portion of time within two meters, suggesting that the technology did not enhance the perception level of situation awareness.
- In the current implementation, the integration of Augmented Reality with a traditional egocentric view is not recommended for interface design in inspection operations.
- Mission planning operations, when applicable, should include synchronous human-in-the-loop presence for telerobotic inspection of spacecraft.


Comparison of Display and Human Presence

Hannah Weiss is a master’s candidate in the Department of Industrial and Operations Engineering Center for Ergonomics at the University of Michigan. She earned her BS in 2018 from the University of Southern California in Human Biology with an emphasis in Applied Physiology.

Andrew Liu is a research scientist in the Human Systems Laboratory within the Department of Aeronautics and Astronautics at the Massachusetts Institute of Technology. He earned his PhD in 1993 from the University of California, Berkeley, in Bioengineering.

Amos Byon is a User Interface Software Engineer at the Jet Propulsion Laboratory. He earned his BS in 2016 from the University of Southern California in Computer Science with an emphasis in Game Design.

Jonathan Blossom is the Group Supervisor of the User Interfaces group and Manager of the Ops Lab at the Jet Propulsion Laboratory Ops Lab. He earned his BA in 1992 from Yale University in Religious Studies.

Leia Stirling is an associate professor in the Department of Industrial and Operations Engineering at the University of Michigan, Core Faculty in the Center for Ergonomics, and Core Faculty in the Robotics Institute. She earned her PhD in 2008 from the Massachusetts Institute of Technology in Aeronautics and Astronautics.

Date received: May 6, 2021
Date accepted: August 6, 2021