Impact of Transparency and Explanations on Trust and Situation Awareness in Human–Robot Teams

Akuadasuo Ezenyilimba, Margaret Wong, Alexander Hehr, Mustafa Demir, Alexandra Wolff, Erin Chiou and Nancy Cooke, Arizona State University

Urban Search and Rescue (USAR) missions continue to benefit from the incorporation of human–robot teams (HRTs). USAR environments can be ambiguous, hazardous, and unstable. The integration of robot teammates into USAR missions has enabled human teammates to access areas of uncertainty, including hazardous locations. For HRTs to be effective, it is pertinent to understand the factors that influence team effectiveness, such as having shared goals, mutual understanding, and efficient communication. The purpose of our research is to determine how to (1) better establish human trust, (2) identify useful levels of robot transparency and robot explanations, (3) ensure situation awareness, and (4) encourage a bipartisan role amongst teammates. By implementing robot transparency and robot explanations, we found that the driving factors for effective HRTs rely on robot explanations that are context-driven and are readily available to the human teammate.

Keywords: urban search and rescue, human robot teams, robot explanations, robot transparency, trust in robot, situation awareness

Introduction

Urban Search and Rescue (USAR) missions are often physically demanding and dangerous for human responders, involving tasks that can risk their overall health and safety. At the same time, the immediate aftermath of a disaster situation is often in dire need of rich and accurate information that can help to quickly improve rescue teams' situation awareness and their ability to carry out their missions effectively. Previous efforts have been made to minimize human exposure to precarious and preventable dangers while increasing their team's capabilities through the use of robotic counterparts, altogether known as human–robot teams (HRTs).

In response to Hurricane Harvey and Irma in 2017, unmanned aerial vehicles were deployed for a search and rescue effort post-disaster (Greenwood et al., 2020). Unmanned aerial vehicles helped to enhance situation awareness (SA) by equipping first responders with relevant information to assess the magnitude of the disaster, with minimal added risk to first responders' health and safety (Greenwood et al., 2020). Additionally, in June of 2021, in response to the Miami Beach condominium collapse, first responders were able to team up with semi-autonomous throwable military robots (Brown, 2021). In short, robot teammates allowed human responders to reduce information uncertainty in compromised buildings (Brown, 2021).

Furthermore, ongoing initiatives within the realm of USAR have demonstrated that semi-autonomous robot teammates can be valuable assets (Hong et al., 2018). For instance, robot counterparts can aid in USAR environments where visibility is diminished by displaying transparency through specialized robot sensors and navigation or mapping functions. Environmental hazards that can result in diminished visibility, such as fires, smoke, or dust, can not only put responders at risk but also the potential victims as well. These situational limitations can severely constrain responders' movements and ability to make timely progress in the rescue effort (Couceiro et al., 2020).

Each of these instances involves some combination of humans and robots working

Address correspondence to Akuadasuo Ezenyilimba, Human Systems Engineering, Arizona State University, 150 Santa Catalina, Mesa, AZ 85212, USA.
Email: aezenyi1@asu.edu

Journal of Cognitive Engineering and Decision Making
Vol. 0, No. 0, ■■ ■, pp. 1-19
DOI:10.1177/15553434221136358

interdependently toward a common goal. We refer to this entity as an HRT. However, even with the introduction of robot technology in USAR environments, many USAR missions rely on robots that have limited communication abilities or are only capable of displaying simple elements of robot status to the human operator. Although robot status information is pertinent, this type of communication does not require the HRT to engage beyond the surface level, relegating the human operator to more of a monitor or a supervisor rather than a teammate, with team interactions being limited to inquiring about or working within the constraints of their robot teammate.

The limited or non-existent communication capabilities of USAR robots can therefore stagnate the effectiveness of HRTs, resulting in robots being viewed as tools in a dependent relationship rather than as a teammate in an interdependent relationship. This objectification, while important under some decision contexts such as appropriate calibration of trust in a tool, could also impede the ability of human counterparts to establish a basis for trust and interactive team cognition with increasingly capable robots. Trust and interactive team cognition are essential components of effective teams (Cooke et al., 2013; Schaefer, 2016).

When it comes to integrating robots into HRTs, there is a justifiable concern of using the teaming lens rather than the tool lens due to the inevitable nature of robot automation failures (Honig & Oron-Gilad, 2021). Unexpected robot failures—even when those failures could be attributed to environmental factors rather than technological factors—typically result in diminished human trust and can result in the robot teammates being viewed as unpredictable or unreliable. Robot failures of this nature shed light on the need for solutions to address the shortcomings of semi-autonomous robot teammates to better realize their potential as useful teammates.

With this in mind, the purpose of this study was to explore how robot explanation and transparency affect human trust and situation awareness within HRTs and to quantify the best modes of HRT communication within a simulated USAR environment. We also address to what extent team member roles promote SA.

Robot Transparency

SA can be categorized into three ordinal levels: perception of elements within a given environment, comprehension of the present situation, and projection of the situation's future status (Endsley, 1995). These three levels contain essential information relevant to the state of a dynamic environment. To support trust, the Chen et al. (2014) SA-based agent transparency (SAT) model expresses that an agent needs to convey transparent interaction through (1) information that captures an agent's purpose, process, and performance; (2) support of the human's understanding of the agent's reasoning process and constraints; and (3) information that helps with projecting the agent's expected future state (Chen et al., 2014; Lee & See, 2004). The SA-based agent transparency model also supports SA for the HRT.

Related studies have shown that high levels of transparency support SA (Selkowitz et al., 2016), assist in the proper calibration of trust in a robot (Mercado et al., 2016; Selkowitz et al., 2016), and lead to increased task performance of an HRT. Inadequate transparency can lead to mistrust or inappropriate levels of trust in the automation when humans and agents interact (Lewis et al., 2018). If trust is low, the human teammate will underuse the assistance of their robot teammate. For example, miscalibrated trust that results in the improper use of automation can lead to failures of omission or commission (Chen et al., 2014).

On the other hand, placing too much trust in a robot teammate can result in over-reliance on robots for tasks outside the robot's design parameters, which can negatively impact trust (Yagoda & Gillan, 2012). Proper use of transparency may help counteract miscalibrated trust within an HRT and subsequently improve HRT performance (Selkowitz et al., 2016). Therefore, an interface that provides a teammate with information pertaining to the three SAT levels is expected to improve situation awareness, trust calibration, and the overall performance of the team.

Although high levels of transparency would seemingly allow information to flow without unnecessary delay in team communication, research has shown that team interaction (i.e., communication and coordination) mediates the appropriate transfer of information that supports effective team performance (Cooke et al., 2013). It is one thing to be aware of elements within one's environment, but without knowing how to incorporate information within the environment, it can be hard to decipher how to effectively use said information. This is especially so in USAR, in which the team is often juggling multiple channels of information. This is where robot transparency and robot explanations become essential. However, clarity is needed on how to properly balance the levels of transparency and explanation that impact the human teammate's trust in their robot teammate.

Robot Explanations

Explanations are explicit communication that provide a reason behind a decision or action occurring relative to a counterpart's understanding (Miller, 2019). Explanations may be used to provide further context or justification for a task or plan deviation. Robot-driven, context-based explanations within HRTs have been shown to improve trust (Wang et al., 2018).

Explanations are particularly useful in dynamic, interactive tasks to correct misalignment of expectations and to support team cognition (Cooke et al., 2013; Miller, 2019). Without explanations, the robot will have limited ability to serve in a responsive role; thus, serving more like a tool rather than a teammate (Chiou & Lee, 2021). Although the use of explanations can be used to improve transparency in an HRT, explanations and transparency are separate concepts. In decision support systems, transparency is more often conveyed through real-time information-exchange modalities (e.g., live status indicators) while explanations are considered retrospective, contrastive, and responsive reasoning frequently conveyed through verbal or text-based communication (Endsley, 1995; Miller, 2019).

However, the perennial challenge in designing HRTs is to support the transfer of the right information at the right time and to avoid designs that contribute to information overload. Whereas transparency displays are readily available, they primarily present information that has been predetermined as important. But, in dynamic task environments, explanations may even be perceived to be more useful than immediately transparent information (Bartlett & Cooke, 2015). If robot transparency is operationalized as a *continuous state of information about the robot teammate, including environmental factors impacting the robot's ability to function*, and robot explanation is operationalized as *explicit* ad hoc *communication that conveys contrastive information with respect to a human teammate's understanding,* then the explicit communication that draws attention to critical information may be more productive than that of more implicit or passive communication such as status indicators (Shah & Breazeal, 2010).

Our previous HRT research focused on different robot explanation strategies and found that moderate levels of proactive information updates, including explanations, were positively related to team effectiveness in a simulated USAR environment (Chiou et al., 2021). Insights from this previous study suggest that when relevant information is conveyed proactively, human teammates are better able to focus on the task at hand, improving individual and overall team performance (Chiou et al., 2021). Thus, the current study focuses on investigating the line between too much communication and robot explanations and reduced communication and robot explanations in a similar task environment. The current study also attempts to determine the appropriate balance of information conveyed to a human operator through explanation and transparency.

Trust in the Robot

In considering the development of an effective HRT, the implementation of robot transparency and robot explanation is important for trust. Trust can be defined as a willingness to be vulnerable based on the expectations of the behavior of another (Rousseau et al., 1998). Increasingly autonomous robots may need humans to place higher trust in them, due to the

inability of humans to fully monitor all aspects of the robot in action. Because trust often requires a sense of mutual understanding, this suggests that team communication that contributes to this sense of mutual understanding, including transparency and explanation, is critical for both trust as well as teaming (Edmonds et al., 2019).

One concern with respect to increasing trust with increasingly autonomous robots is that human operators may be more keen to attribute blame to their robot teammate when unexpected errors within the team occur (Lyons, 2013). This mindset can prevail even when transparency is used to convey the inner workings of the robot teammate. To overcome the conflation between robot and human teammates' responsibility within an HRT, it is important to explicitly make human operators aware of their responsibility within the team dynamics (Lyons, 2013).

As of this writing, robots in USAR are still predominantly used or viewed as assistive tools, even when traversing environments deemed too dangerous for human responders (Greenwood et al., 2020). This is because the promised or envisioned autonomous robot capabilities are not yet deployable in real-world USAR settings. Yet, even with robots originally being viewed as tools, human teammates often attribute responsibility to the robot, especially when responsibility is an explicit capability of the robot (Yagoda & Gillan, 2012). Human perception of robot teammates relies heavily on their associated mental models of how the robot communicates (Phillips et al., 2011). Meaning if the human teammate views the robot as being capable of communicating ongoing errors, there may also be a false preconception that the robot teammate is capable of identifying how to address said errors, even when the robot's capabilities have been explicitly outlined beforehand (Phillips et. al, 2011).

With this in mind, establishing trust while maintaining a sense of responsibility may come to depend on the means for maintaining SA between the human and robot teammates. Even with unexpected events, robots need to be recognized as a valuable asset, while avoiding contributing to a human's absolution of responsibility during a task (Lyons, 2013). By providing the human teammate with appropriate levels of SA, their role can be explicitly outlined to further instill what they are responsible for within the HRT, in order to find success as a team. Therefore, to support SA, trust, and perceived responsibility, we posit that the robot must be capable of conveying its status clearly and succinctly (e.g., transparently).

But the question remains: How to implement robot transparency and robot explanations to effectively support human–robot trust and HRT effectiveness? To answer this question, we implemented varying levels of robot transparency and robot explanation as four conditions. The intent of the study was to understand the role and best modes of (1) *robot explanation* and (2) *robot transparency* in support of trust, situation awareness, and perceived workload to determine the appropriate levels of team communication. Assuming a teaming configuration with a semi-autonomous robot, we also wanted to explore (3) the role of perceived workload on *trust in the robot* teammate. We hypothesize that SA will be positively impacted by information availability for both transparency and explanation. Secondly, we hypothesize that trust will increase with more frequent robot explanations. Thirdly, we hypothesize that trust will be negatively affected as the human teammates' workload increases.

SIMULATED Minecraft Urban Search and Rescue Study

Virtual Task Environment

To explore the effects of different modes of robot transparency and robot explanations in an HRT, our study team designed and deployed a virtual and remotely operated testbed for a USAR mission context. The testbed interface used prerecorded Minecraft mission videos (as seen in Figure 1) that were shown to the human operator, leveraging the Wizard-of-Oz (WoZ) method (Bradley et al., 2009). The testbed interface included an embedded map and status indicators that varied based on condition. These Minecraft videos were then paused by an experimenter throughout the mission, according to a mission script that was not known to participants, conveying "live" stoppages involving the robot teammate.

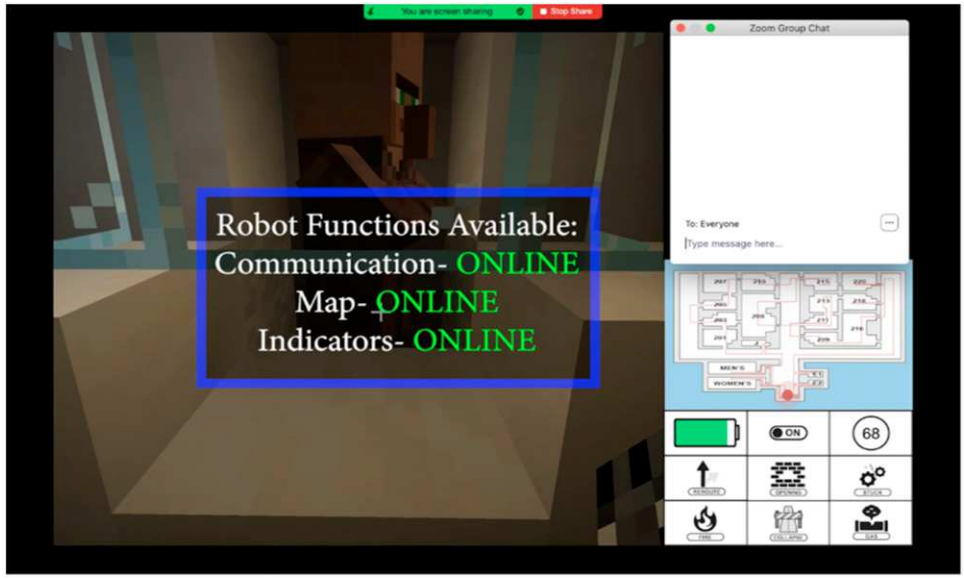


Figure 1. Participant view: Minecraft simulation via Zoom (Wong et al., 2021).

An administrative application (Figure 2) was custom-built for this study and used to stop and start the previously recorded Minecraft video, play sounds associated with implemented interactive tasks (robot disruptions), and to display a live timer countdown and the participant's mission points. Each mission had a fixed timer displayed to participants to count down the allotted mission time. In the midst of COVID-19, this entirely virtual testbed of digital tools enabled the experimenter team to smoothly run the study in a virtual setting with consistency and without risking the health of participants or experimenters.

Inclusion of Situation Awareness Within Testbed

To answer our research questions regarding HRT communication, it was important to incorporate all three levels of SA into the experimental environment because our assessment of team effectiveness (i.e., team performance) was dependent on our measures of SA (i.e., made equivalent). Therefore, as part of our study design, we developed a complex task environment with dynamic scenarios and tools that would involve the three levels of SA. We describe below how these levels were operationalized within the simulated HRT task environment.

Level one SA (perception) focuses on how elements within an environment are discerned. To actualize this first level, disaster victims were presented to participants in the task environment in three colors: red, blue, and yellow. The variation in victim color was to assess if participants were capable of recognizing the correct information within their environment (note that color-blind participants were excluded from the study). Also, the inclusion of environmental changes (fires, collapses, and openings) enabled the need for participants to perceive the attributes associated with elements in the simulated environment (Endsley, 1995). Identifying environmental changes was pertinent to the success of creating a useful map for rescue after the search was complete.

Level two SA (comprehension) was implemented through the incorporation of a map (dynamic and static; dependent on condition; see Figure 2) and the interactive tasks (described in more detail in Methods). The combination of these two components within the missions required participants to understand the associated significance with ongoing events and objects within the intended environment (Endsley, 1995).

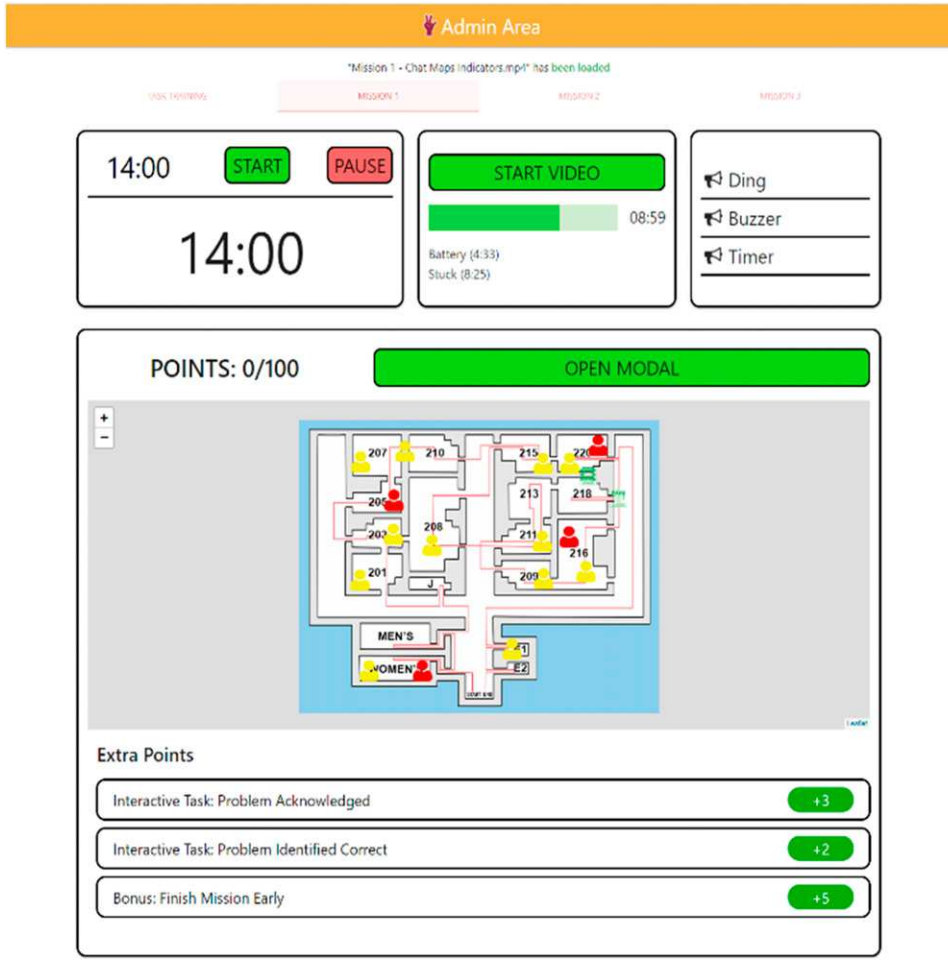


Figure 2. An administrative application, moderator control panel (Wong et al., 2021).

The inclusion of the map gave participants knowledge related to the robot's location and intended route of travel by the robot. The interactive task required participants to use their knowledge from ongoing events and not only just notice but also understand that something about their robot teammate required their attention.

Lastly, level three SA (projection) is associated with an adequate forecast of future actions based on elements within your environment (Endsley, 1995). By tasking participants with finding solutions for the interactive tasks, we were able to gauge participants' ability to use varying levels of robot transparency (displayed through status indicators) and robot explanation in conjunction with elements within their environment (e.g., fire blocking the route) to then address the ongoing issue by identifying a solution. Through the interactive tasks, we wanted participants to use the various modes of transparency and robot explanation capabilities to identify the appropriate next steps in the event of robot malfunction or unexpected event.

A baseline condition was also implemented to represent an HRT task environment that supported level one SA only. The robot was instructed to only provide location-based information to the participant (see Table 1, Experimental Procedure for more details). In contrast, the transparency and explanation conditions were designed to support levels two and three of the Chen SAT model. Within the

Table 1. Experimental Conditions: Levels of Explanation and Transparency.

Condition	Level of available explanation	Level of available transparency
Full	Full explanation	Dynamic map and status indicators
Transparency-only	Limited to location-only questions	Dynamic map and status indicators
Explanation-only	Full explanation	Static map only
Baseline	Limited to location-only questions	Static map only

transparency-only condition, battery status, a dynamic map with preplanned labeled routes, and the reroute icon indicators provided reasoning for certain actions and future status of the robot. Within the explanation-only condition, the robot was allowed to answer questions pertaining to the mission and robot status.

Methods

The following sections will discuss the participant sample, experimental procedure, study design, and various components of the mission task that study participants faced across two mission sessions. During the missions, participants were expected to work as a team with a virtual robot in a partially collapsed building. Participants were responsible for working with their robot teammate to locate and identify potential victims, identify and locate hazardous environmental changes, and to assist the robot during specific interactive tasks which were designed into the simulated environment (see Interactive Tasks and Mission Success section).

Participants and Power Analysis

An a priori power analysis was conducted using G*Power3 (Faul et al., 2007) to test the difference between the means of 4-group by 2-mission using an *F*-test, a medium effect size ($\eta_p^2 = .06$), and an alpha of .05. According to the results, a total sample of 48 participants with four equal-sized groups of $n = 12$ is required to achieve a power of .80. We also considered that there may be issues during data collection, and we increased the sample size.

Participants signed up to participate in this study via Signup Genius, and they were recruited through Arizona State University Slack channels, which include faculty, staff, students, and other community members with an @asu.edu email address. The study took approximately 1-hour and 15-min, and participants were randomly assigned to one of four conditions (full, explanation-only, transparency-only, and baseline—described in more detail in the next section) following informed consent.

Experimental Procedure

To complete the USAR mission in the virtual Minecraft environment, participants were informed that they were a part of an HRT search team that was responsible for generating a useful map for the rescue team to be able to save victims trapped inside. The goal of the mission and various task roles involved were communicated to participants during a study onboarding process and reinforced during a one-mission training walk through (see Table 2 for more information). The Minecraft missions and walk-through tutorial were facilitated by Zoom and a three-person team of researchers executing the "live" events. Next, participants completed two actual missions with the robot teammate in Minecraft with the same research team, before completing questionnaires capturing trust, demographic information, and the NASA-Task Load Index (TLX; Hart & Staveland, 1988).

Study Design

In this study, we focused on testing four different conditions, which we describe as full, explanation-only, transparency-only, and baseline. In essence, robot explanation and robot transparency were implemented at varying levels (see Table 1).

Based on the condition, participants were responsible for using the available robot transparency and robot explanation to assist their robot teammate (see Table 2).

To support sustained task engagement and the collection of high-quality interaction and

Table 2. Experimental Sequence of Events and Components.

Mission	Elements	Interactive tasks (reasons)
Task training (5 min 30 s)	Victims: 7; environmental changes: 2; duration: 5 min 30 s	1 robot stoppage (stuck)
Mission 1 (14 min)	Victims: 16; environmental changes: 2; duration: 14 min	2 robot stoppage (stuck)
Mission 2 (21 min)	Victims: 24; environmental changes: 8; duration: 21 min	2 robot stoppage (overheat, stuck, gas leak, unexplainable robot malfunction, battery, and robot vision impairment)

communication data, participants completed two missions with varying levels of workload (i.e., number of victims, environmental changes, and interactive tasks the human operator had to face). The first mission had a lower workload (16 victims to identify, 2 environmental changes, and 2 interactive tasks), and the second mission had a higher workload (24 victims to identify, 8 environmental changes, and 6 interactive tasks). Increasing workload between missions allowed us to study the effect of increased workload on trust within the HRT.

Team Roles

Mission success was designed to require active participation from both the robot and human teammates. The **role of the robot** was to physically search the building room by room, and the **role of the human** was to watch the "live" feed from their robot's camera. To complete the mission, participants needed to report different types and locations of victims and environmental changes. Two team aids were involved to facilitate the mission; an **Incident Commander** was responsible for recording identified victims and environmental changes communicated by participants via Zoom text chat and providing recommendations in the event of an unexpected stoppage. If necessary, a **Mechanic** was responsible for repairing the robot teammate based on the Incident Commander's recommendation in the event of an unexpected stoppage, only after participants identified the correct reason as part of the interactive task. The robot and Incident Commander were played by experimenters who followed predetermined scripts to communicate with participants via Zoom's text chat feature, based on participant responses. The Mechanic interactions were preprogrammed into the video recording.

Interactive Tasks and Mission Success

Interactive tasks (Table 3) were incorporated to encourage participants to be more active in accomplishing the team mission and responsible for reporting the robot's performance status, thereby encouraging a sense of team interdependence. By having each team member rely on one another for the retrieval of relevant information that contributes to the overarching search and rescue goal furthered the need for effective team communication. During the interactive tasks, participants needed to gather knowledge from their perceived environment, use it to comprehend the situation, and project the future state of the robot to assist in the team's progress throughout the missions.

Mission success relied on how efficiently HRTs progressed through the interactive tasks. At certain points throughout the missions, robot stoppages would occur and require direct assistance from the human for the robot to move forward in its task. Robot stoppages were caused by pausing the prerecorded video representing the robot's camera view as it traversed the partially collapsed building, so that it would appear as if the robot could not move. Once participants noticed their robot teammate had stopped, they would need to use the available levels of robot transparency (check the status indicators) and robot explanation capabilities (address the robot via text chat) to identify the issue to the Incident Commander.

Table 3. Interactive Tasks and Requirements.

Interactive task	What happens	Status indicator operation	Correct participant response	Incident commander response
Stuck	Robot becomes stuck in doorway or opening	Stuck status indicator turns yellow while robot is stuck	Incident commander or IC robot is stuck	The robot should attempt to free itself by clearing a path and wiggling out of the opening
Low battery	Robot begins moving slow/ choppy due to diminished battery level	Battery changes from green to red once it reaches critical condition	Incident commander or IC robot battery is low	I would recommend the robot to reroute and find an outlet. It looks like there is an outlet at the top of the right hallway
Overheating	Robot encounters fire attempt to avoid and then stops moving	Internal temperature increases from 68 to 100	Incident commander or IC the robot is overheating	I would recommend the robot to reroute to the mechanic before severely damaging itself
Gas leak	Limited to location-only questions	Gas leak status indicator turns yellow while gas is detected	Incident commander or IC the robot has detected a gas leak	I would recommend the robot to reroute and search the rooms with gas first in case there are still any victims alive. Victims in those rooms are more likely to be in a critical condition
Robot malfunctions	Camera begins to spin around and reaches a stop facing the ceiling	All indicators light up yellow	Incident commander or IC the robot is malfunctioning	I would recommend the robot reroute to the mechanic to fix any malfunctions
Loss camera signal	Screen becomes static	No status indicator changes	Incident commander or IC camera signal is lost	The mechanic can fix the loss of camera signal remotely within 15 s. In the meantime, I would recommend the robot reroute and go around

Participants were informed during the onboarding process that if the robot happened to incur an unexpected event, they would be unable to continue the mission until they correctly identified the issue and received a recommendation from the Incident Commander. However, in the event that participants were unable to identify the correct issue, the Incident Commander would still provide a recommendation after a one-minute time limit had passed, effectively releasing them from the task. Although this one-minute time limit could potentially result in a performance floor effect in terms of time to complete the mission, it was more relevant for our research objective to eventually allow all participants to get through as much of the mission as possible, even if they happened to be really poor at troubleshooting the issue with

the robot. (The goals of this study are less focused on traditional production measures of performance like time to complete the mission and more on team effectiveness measures like situation awareness, trust, communication, and workload). It was assumed that the countdown timer fixed in the upper left-hand corner of the participant's screen created some time pressure during these interactive tasks and simulated a sense of urgency to correctly identify the issue before the one-minute mark.

To further facilitate sustained attention and active engagement during the interactive task (i.e., to prevent participants from simply waiting for the one-minute mark in the event they happened to encounter the Incident Commander releasing them from this task early on in a mission), participants could hear up to three different sounds as an interactive task occurred: (1) a ticking timer acknowledging an interactive task is taking place, (2) a positive ding indicating the participant has correctly identified the ongoing issue, or (3) a negative buzzer indicating the participant has incorrectly identified the ongoing issue. All participants were able to enter as many attempts as they desired (i.e., they had unlimited guesses) within the allotted one-minute time frame.

Explanations Through Participant-Generated Questions

Communication. During the study onboarding process, participants were informed of types of communication and possible questions they could use when communicating with their robot teammate. Participants were also informed that certain functions may or may not be available and that it "would be up to them to figure out the best way to relay information." Depending on the participant's condition, communication availability would vary as follows, see Table 4.

Communication Implementation. The purpose of the explanation-only and full conditions is to see how participants communicate with their robot teammate given increasing levels of complete robot responsiveness capabilities. During the study onboarding process, participants were shown examples of mission-related questions they could ask their robot teammate. Upon completing the onboarding process participants were guided through a training mission, where the experimenters walked participants through the types of tasks that would be completed as part of the mission, including practicing communication with the robot and the other team aids.

Robot explanations were primarily linked to the designed interactive tasks and were initiated by the human teammate's questions. This provided the requisite communication data needed to see where or what questions were typically asked when there were robot deviations to planned or expected behaviors. Participants were not restricted in terms of how they should ask the question as long as the question was relevant to the mission and the randomly assigned condition and ended in a question mark. If those criteria were met, then the robot would respond depending on the condition.

In the event participants asked questions that were not relevant to the mission, in conditions with high explanation (full and explanation-only) the robot teammates would inform their human teammate they could not answer the question asked (i.e., *"I do not have the capabilities to answer that"*). In low-explanation conditions (baseline and transparency-only), participants would not receive a response from their robot teammate. These instructions, framed as HRT communication requirements due to the robot's limitations with respect to natural language processing, served a dual purpose for analyzing the resulting communication data (text chat transcripts) in an unambiguous way.

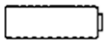
Transparency Indicators

Within the low transparency conditions (explanation-only and baseline), the level of robot transparency was limited to a static map that displayed the robot teammates' preplanned route. Participants in the high-transparency conditions (full and transparency-only) had access to a map that displayed a preplanned route with a continuous live update of the robot's location and status indicators. The status indicators were used to convey higher transparency information that was not afforded in the

Table 4. Types of Explanation Questions.

Condition	Level of robot responsivity	Example question
High explanation: Full and explanation-only	Address all questions pertaining to the mission that end in a question mark	1. Why are you stopped? 2. Are you stuck? 3. Is that a fire? 4. Can you reroute?
Low explanation: Transparency-only and baseline	Address only questions that refer to location information and end in a question mark	1. Where are you? 2. What room are we in? 3. What is our location?

Table 5. Transparency Status Indicators With Explanation of Needed Action (Wong et al., 2021).

Indicator icon		Description	Action
Off	On		
		Robot's internal temperature	Team need to attempt to find a new path to avoid heat damager; if the robot is damaged, the team will reroute to the mechanic
		Robot's operational status	Additional information about robot status; no additional action
		Robot's battery life	When the battery is low, the team will reroute to a charging station
		Any environmental collapses are impeding the robots route	Team will reroute to get to rooms blocked by collapse to incident commander
		The presence of a fire in the environment	Team will reroute around the fire to continue with the mission; teammate reports fire to incident commander
		Robot is unable to move due to environmental conditions	Participants will identify the problem with the robot teammate. The robot will then wiggle to release itself
		Robot is deviating from the preplanned mission route	Additional information about robot status; no additional action
		Robot has encountered an opening in the building environment	Team will enter opening to ensure room is clear: Teammate reports opening to incident commander
		Presence of gas leak in the environment	Team will find a new route to search rooms with gas leak safely and identify most critical victims first

baseline or explanation-only condition, see Table 5. Status indicators, operationalized as dynamic icons (see Table 3), were specifically designed and selected for the robot to communicate task-relevant information to participants in an efficient manner; during the onboarding process participants were made aware of each status indicator's meaning.

Measures

Three dependent measures were used in this study: SA, trust, and workload. Trust and workload questionnaires were given to participants after Mission 1 and Mission 2. Two different trust instruments were used to measure trust in human–robot interaction (HRI) and trust in teams (described in more detail later).

Situation Awareness

A point system (see Table 6) was created and implemented to assess participants' overall situation awareness within the simulated USAR environment. All conditions had three main components that comprised our situation awareness measure in this study: victims, environmental changes (fire, collapses, and openings), and interactive tasks; bonus points were awarded for completing a mission early.

Identification of victims and environmental changes contributed to the assessed situation awareness in terms of perception, whereas acknowledgment of interactive tasks, listed in Table 3, contributed to the assessed comprehension-related situation awareness. The point system was influenced by two factors: level of SA and weighted value for reproducing a useful map for the rescue team based on our pilot studies. Participants were made aware of the point system during the onboarding and training mission prior to the start of Mission 1.

In developing our point system, identification of victim type (i.e., color) was considered the lowest level of difficulty among the tasks. One point was equally awarded and taken away when victims were correctly and incorrectly identified. Points awarded to identifying environmental changes varied based on correct and incorrect identification. The number of awarded and deducted points was based on the difficulty and the importance of proper identification of the type of environmental change to the overall mission (see Table 6).

The interactive tasks were considered the highest due to the level of task difficulty and importance to the team's success. There were two essential parts to completing these interactive tasks: acknowledgment and proper identification of the robot stoppage. Acknowledgment was a necessary component, but not the entirety of the mission task. In the event that participants acknowledged but did not correctly identify the interactive task, this could negatively affect the HRT, possibly resulting in running out of time and not being able to finish the search.

SA scores were calculated for both Mission 1 (52 possible points) and Mission 2 (118 possible points) separately. The sum of all participants' SA scores was then calculated and divided by the total possible points and then converted to a decimal form.

Table 6. Situation Awareness Scoring System.

Description	Correct	Incorrect
Victim	2+	2-
Environmental change	5+	2-
Interactive task acknowledged	3+	3-
Interactive task identification	2+	2-

Trust in the Robot Teammate

To measure trust in the robot teammate, Schaefer's (2016) Trust Perception Scale-HRI was included as part of the study questionnaire. We used an adapted version of the recommended 14 item Subject Matter Experts (SMEs) subscale. Of the 14 items we used 12 of the questions, omitting two that were not relevant to our study:

1. What percent of the time did the robot teammate function successfully?
2. What percent of the time did the robot teammate act consistently?
3. What percent of the time was the robot teammate reliable?
4. What percent of the time was the robot teammate predictable?
5. What percent of the time was the robot teammate dependable?
6. What percent of the time did the robot teammate follow directions?
7. What percent of the time did the robot teammate meet the needs of the mission?
8. What percent of the time did the robot teammate perform exactly as instructed?
9. What percent of the time did the robot teammate make errors?
10. What percent of the time did the robot teammate provide appropriate information?
11. What percent of the time was the robot teammate unresponsive?
12. What percent of the time did the robot teammate malfunction?

We also changed the tense of the questions from future to past tense and changed "robot" to

"robot teammate." We used a sliding bar response scale that ranged from 0% to 100%, rather than having a set selection of potential responses. This permitted participants to have finer-grained control over their response to the question items.

Participants' trust in robot scores was calculated for both missions by finding the sum of positive language questions asked on the questionnaire (e.g., "*What percent of the time did the robot teammate act consistently?"*) and then subtracting that total from the sum of negative language-based questions (e.g., "*What percent of the time did the robot teammate make errors?"*).

Trust in a Team

To measure trust in a team, we used six items from Costa and Anderson's (2009) 21-item trust questionnaire to assess trust within teammates regardless of the USAR task context.

1. In this team, we can rely on each other.
2. In this team, we have complete confidence in each other's ability to perform tasks.
3. In this team, we do not hesitate to help each other when in need.
4. In this team, we work in a climate of cooperation (i.e., environment in which you work together).
5. In this team, some of us hold back relevant information.
6. In this team, we minimize what we tell about ourselves.

For the purposes of this study, we used questions from three of the four original categories: propensity to trust, perceived trustworthiness, and cooperative behaviors. The six questions chosen to represent those three categories were considered the most applicable to both a human and an automated teammate and avoided redundancy with the questions from the trust in the robot teammate questionnaire.

Participant's trust in a team was calculated for both missions by finding the sum of positively valanced question items (e.g., "*In this team, we can rely on each other"*) and then subtracting the total from the sum of the negatively valanced items (e.g., "*In this team, some of us hold back relevant information"*).

Workload

NASA-TLX (Hart & Staveland, 1988) assessed participants' perception of workload during both missions. Workload scores were calculated by totaling the sum of the questionnaire responses for both missions.

Data Analysis and Results

The first split-plot analysis of variance (ANOVA) addresses how *situation awareness* differed across conditions and missions. Although there were significant condition, $F\,(3, 57) = 5.75$, $p = .002$, and mission main effects, $F\,(1, 57) = 12.04$, $p = .001$, condition by mission interaction effect was not significant, $F\,(3, 57) = .49$, $p = .693$. According to the significant condition main effect, teams in the baseline condition had significantly lower SA scores than the other conditions (transparency-only: $p = .027$, explanation-only: $p = .001$, and full: $p = .001$), whereas the other conditions did not differ on the SA score, $p > .050$, see Figure 3(a). According to the significant mission main effect, SA scores significantly decreased from Mission 1 to Mission 2, $p = .001$, across all conditions, likely due to the high workload in Mission 2.

The second split-plot ANOVA addresses how *trust in the robot* differed across conditions and missions. Although there was a significant mission main effect, $F\,(1, 57) = 312$, $p < .0001$, and the condition main effect, $F\,(3, 57) = 2.05$, $p = .117$, condition by mission interaction effect was not statistically significant, $F\,(3, 57) = 1.01$, $p = .394$. According to the significant mission main effect, the operator's trust in the robot significantly decreased from Mission 1 to Mission 2, $p < .0001$, likely due to the increase in interactive tasks with the robot in Mission 2 (Figure 4).

The next split-plot ANOVA addresses how *team trust* differed across conditions and missions. Although there was a significant condition, $F\,(3, 57) = 3.37$, $p = .025$, and mission main effects, $F\,(1, 57) = 30.4$, $p < .0001$, the condition by mission interaction effect was not significant,

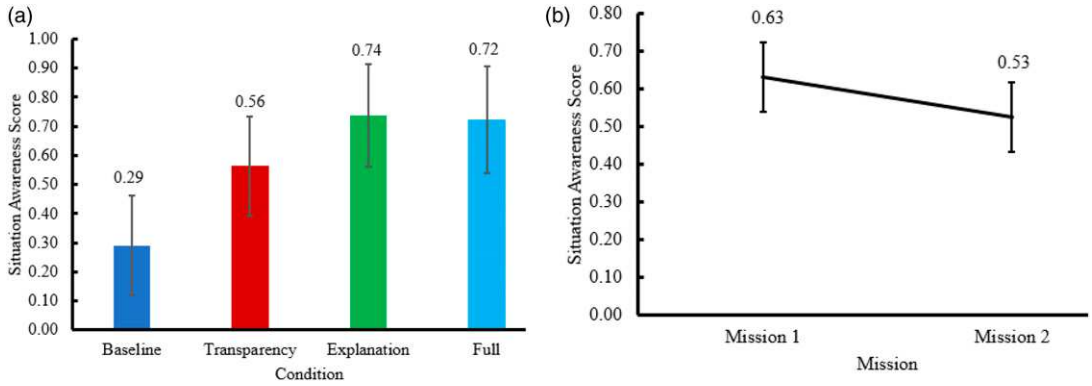


Figure 3. Mean SA across (a) conditions and (b) missions (error bars are 95% confidence intervals).

F (3, 57) = .35, p = .788. According to the significant condition main effect, teams in the baseline condition had significantly lower team trust than other conditions (transparency-only: p = .038, explanation-only: p = .011, and full: p = .007), whereas the other conditions did not differ on the team trust score, p > .050, see Figure 5(a). According to the significant mission main effect, team trust in the robot significantly decreased from Mission 1 to Mission 2, p < .0001, likely due to the robot's abnormal behaviors in Mission 2, see Figure 5.

The final split-plot ANOVA addresses how the operator's *perceived workload* differed across conditions and missions. Although there were significant condition, F (3, 57) = 3.54, p = .020, and mission main effects, F (1, 57) = 21.4, p < .0001, the condition by mission interaction effect was not significant, F (3, 57) = 1.88, p = .143. According to the significant condition main effect, teams in baseline condition had significantly higher perceived workload than transparency-only, p = .034, and full conditions, p = .003, but comparable amounts in the explanation-only condition, p = .149. All other conditions were not different in terms of perceived workload, p > .050, see Figure 6(a). According to the significant mission main effect, participant perceived workload significantly increased from Mission 1 to Mission 2, p < .0001, possibly due to the robot's abnormal behaviors during interactive tasks and environmental changes in Mission 2, see Figure 6.

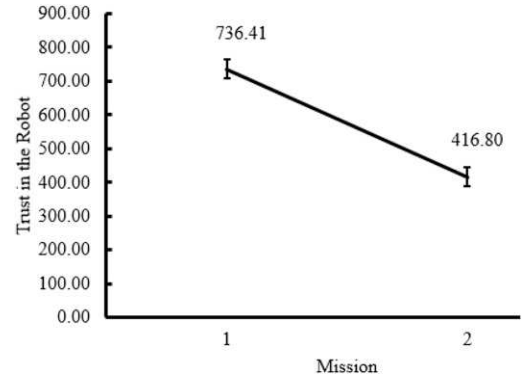


Figure 4. Mean trust in the robot across missions (error bars are 95% confidence intervals).

Discussion

Implications for Robot Transparency and Explanations

This study sought to understand which modes of robot explanation and transparency best support trust, situation awareness, and perceived workload and also explored how participants' perceived workload impacted trust in the robot teammate. Results from this study support hypothesis one; SA was positively impacted by the availability of information in both transparency and explanation conditions. The results also supported hypothesis three, that increasing workload can negatively affect trust. Although hypothesis two, that trust will increase with more frequent robot explanations, was not directly supported, there was some support showing that

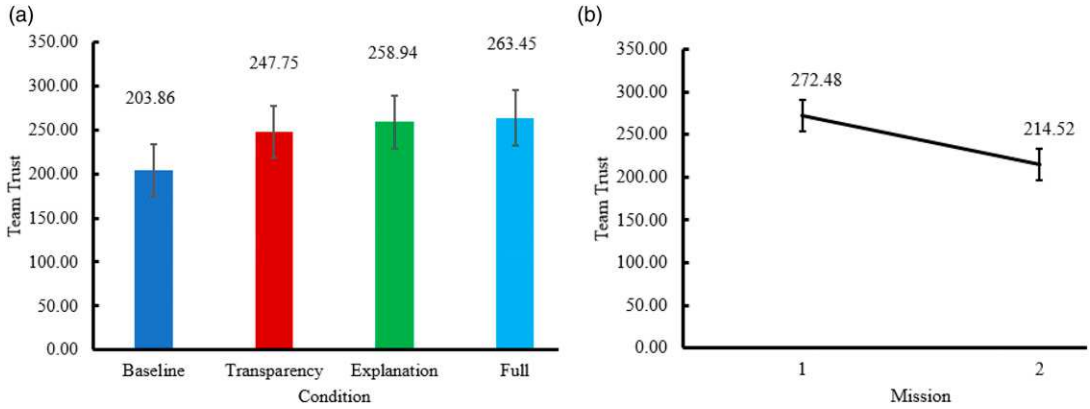


Figure 5. Mean team trust across (a) conditions and (b) missions (error bars are 95% confidence intervals).

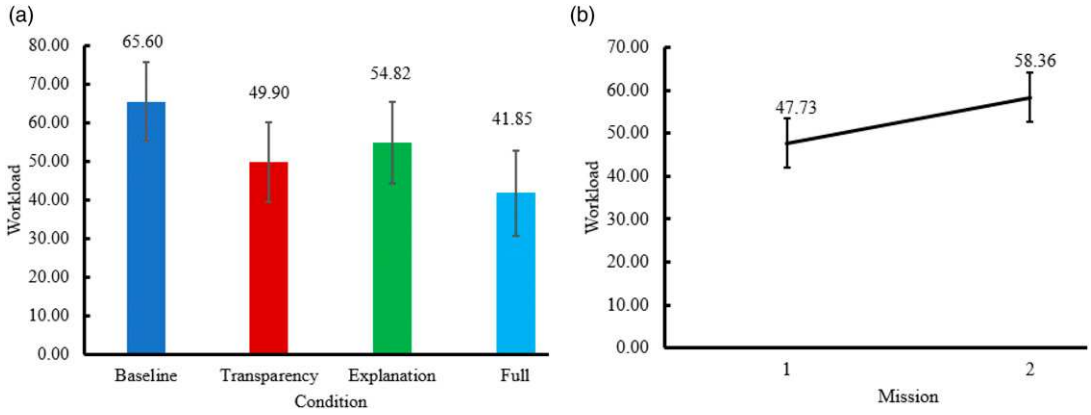


Figure 6. Mean workload across (a) conditions and (b) missions (error bars are 95% confidence intervals).

information, including both transparency and explanations, improved trust in the robot.

Previous research would suggest that the most successful conditions would be those containing robot transparency (Lakhmani et al., 2016). However, in the transparency-only condition, we noticed HRTs did not have as high of an SA score when in comparison to HRTs in condition with more detailed explanations. Due to the way transparency was operationalized across conditions, and the use of robot explanations that provided reasoning for unplanned robot behaviors, human teammates were able to have increased levels of trust, higher levels of SA, and more successful missions. By using a point system to measure SA within the USAR environment, we were able to evaluate team performance as individual components needed to complete the overall mission goal.

Because our point system for measuring SA was developed using face validity for addressing our research questions, further research should be conducted to validate it prior to broader adoption. That being said, our SA measure enabled us to assess whether the elements of transparency (map and status indicators) were effective. At the same time, our results indicate that conditions reliant on this transparency information alone, operationalized as status indicators, did not result in HRTs with the highest SA. The full condition (full transparency and explanation) showed that robot transparency through status indicators and a continuously updated map had little to no impact when compared

to the explanation-only condition (full explanation and limited transparency). Although this was due to how we designed the study and the subsequent point system, it essentially demonstrates that context-based robot explanations in team task environments characterized by uncertainty may be essential to improving SA. Robot explanations provided human counterparts with the necessary information for quicker decision-making in the interactive tasks, enabling human teammates to be more active members within the HRT and better able to support team effectiveness.

Implications for Trust of the Robot

The purpose of the interactive tasks was not only to simulate the type of dynamic and uncertain environment that characterizes USAR but also to encourage a more active role for study participants as well as interactivity within the HRT. In previous pilot studies, we found that participants tended to take on more passive roles in this task environment when left to their own devices, initiating fewer communication threads than expected and that would be required within a more dynamic team task (Chiou et al., 2021). Through the interactive tasks, we were able to see how these unexpected events—and the robots' ability to handle these events with respect to their human teammate—affected trust within an HRT and the overall effect on human trust in their robot teammate.

The increase in interactive tasks from Mission 1 (two interactive tasks) to Mission 2 (six interactive tasks) led to an increase in perceived workload for human teammates, as well as a decrease in team trust with the robot. It is possible that the combination of increased perceived workload as well as the increase in number of interactive tasks was attributed to the robot performing poorly at its task rather than due to factors in the environment or other causes. Future studies would benefit from explicitly measuring attribution of blame and trust as it relates to HRT in dynamic and uncertain task environments (Hsiung & Chiou, 2019).

Higher levels of team trust in the explanation and full conditions had a notable tradeoff in terms of increased perceived workload. Perhaps having to read and understand the additional text of the robot explanations contributes to the overall perceived workload, and it is possible that with additional communication comes additional cognitive overhead. At the same time, the baseline condition had the highest workload despite enjoying the least amount of communication—a consequence of having to figure out most of the task elements on your own without the team benefits that come with increased communication. Yet, trust in the robot teammate was only significantly lower in the baseline condition relative to the other conditions. The baseline condition was also a more extreme condition in which participants had access to limited robot information; these limitations left abnormal robot behavior unexplained, reinforcing lower trust in the HRT as they progressed throughout the missions and the interactive tasks.

Implications for Workload

In USAR missions, workload is another critical factor that can hinder appropriate human trust in HRTs (Khasawneh et al., 2019). USAR is a notoriously high workload environment characterized by human operator fatigue due to the urgent and critical nature of the task environment. This subsequently creates the potential for fatigued operators to distrust and disuse technology that cannot immediately demonstrate reliability (Khasawneh et al., 2019). We were able to see some of these same trends play out within the human–robot teams in our study, in that our conditions with lower workload generally also had higher trust. When the robot encountered a stoppage, it consequently had a negative impact on the HRT's SA score because it cut into the time that the teams had to complete their SA-related mission tasks. Thus, with the interactive tasks contributing to a higher workload for the HRT, compounded with being a potential barrier to obtaining a higher SA score, we were able to see that when technology requires additional work to maintain, monitor, or operate, it is often associated with distrust or mistrust. As a result, the increased workload related to addressing complexities in the environment may not only hinder trust development in teams but also an HRT's overall team performance (Khasawneh et al., 2019). This presents a challenge for designing and

implementing future robots for complex and uncertain task environments, especially without the ability to provide responsive explanations.

Limitations and Future Research

Within our remote testbed, we used prerecorded missions. Prerecorded videos were a tool that allowed us to simulate dynamic events while keeping control over what information was displayed to each participant. However, prerecorded videos also limited our study in terms of being able to observe a greater variety of naturalistic interactions between teammates. Prerecorded videos limited the involvement of the participant with the robot's task to relatively short bursts within the designed interactive tasks, each lasting 1 min long or less. Future studies could extend our work by investigating human–robot interactions with real-time teaming.

Conclusion

Human trust within HRTs benefits from readily available, context-driven robot explanations or transparency information (Wang et al., 2018). Although our study demonstrated workload to be higher in conditions with full explanation, HRTs attained higher levels of SA and were more successful at completing the task at hand than HRTs in the control condition.

As expected in the baseline condition with limited robot transparency and explanation capabilities, SA scores were the lowest. Therefore, HRTs with robot teammates that are capable of providing explanations that are responsive to their human counterparts may not only improve human trust but also improve SA. Furthermore, from our study, we were able to see that robot explanations also enable teammates to take on more active roles that allow them to assist one another, serving as a fundamental building block for productive HRTs.

Acknowledgement

We acknowledge the guidance of Dr. Mica Endsley and the assistance of Nathan Kelly and Andrew Gin in developing the testbed and Shawaiz Bhatti, Felix Raimondo, and Chris DiMino who assisted with data collection.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This research was partially supported by a grant from the Air Force Office of Scientific Research [FA9550-18-1-0067].

ORCID iDs

Akuadasuo Ezenyilimba https://orcid.org/0000-0001-7341-9847

Mustafa Demir https://orcid.org/0000-0002-5667-3701

References

Bartlett, C. E., & Cooke, N. J. (2015). Human-robot teaming in urban search and rescue. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *59*(1), 250–254. https://doi.org/10.1177/1541931215591051

Bradley, J. C., Mival, O., & Benyon, D. (2009). *Wizard of oz experiments for companions*. https://doi.org/10.14236/ewic/hci2009

Breazeal, C., Gray, J., Hoffman, G., & Berlin, M. (2004). Social Robots: Beyond tools to partners. In: *13th IEEE international workshop on robot and human interactive communication*. RO-MANIEEE Catalog No.04TH8759). https://doi.org/10.1109/roman.2004.1374820

Brown, D. (2021). *June 30). Throwable military robots sent to assist with Florida condo collapse*. https://www.washingtonpost.com/technology/2021/06/30/throwable-robot-florida-condo-collapse/

Bryson, J. J., Diamantis, M. E., & Grant, T. D. (2017). Of, for, and by the people: The legal lacuna of synthetic persons. *Artificial Intelligence and Law*, *25*(3), 273–291. https://doi.org/10.gdwpkq

Chen, J. Y. C., & Barnes, M. J. (2014). Human–agent teaming for multirobot control: A review of human factors issues. *IEEE Transactions on Human-Machine Systems*, *44*(1), 13–29. https://doi.org/10.1109/thms.2013.2293535

Chen, J. Y. C., Lakhmani, S. G., Stowers, K., Selkowitz, A. R., Wright, J. L., & Barnes, M. (2018). Situation Awareness-based agent transparency and human-autonomy teaming effectiveness. *Theoretical Issues in Ergonomics Science*, *19*(3), 259–282. https://doi.org/10.1080/1463922x.2017.1315750

Chen, J. Y. C., Procci, K., Boyce, M., Wright, J., Garcia, A., & Barnes, M. (2014). Situation awareness based agent transparency. In: *Report No. ARL-TR-6905, aberdeen proving*. U.S. Army Research Laboratory.

Chiou, E., Demir, M., Buchanan, V., Corral, C., Endsley, M. R., Lematta, G., Cooke, N. J., McNeese, N. J., Couceiro, M. S., Portugal, D., Rocha, R. P., & Araújo, A. (2021). Explanation based communication strategies in a simulated human-robot search and rescue team taskInternational Journal of Social RoboticsFostering human–robot cooperative architectures for search and rescue missions in urban fires. *Simulation*, *97*(3), 177–194. https://doi.org/10.1177/0037549720964548

Chiou, E. K., & Lee, J. D. (2021). *Trusting automation: Designing for responsivity and resilience. Human Factors*. https://doi.org/10.gjvcr2

Cooke, N. J., Gorman, J. C., Myers, C. W., & Duran, J. L. (2013). Interactive team cognition. *Cognitive Science*, *37*(2), 255–285. https://doi.org/10.gf69qb

Edmonds, M., Gao, F., Liu, H., Xie, X., Qi, S., Rothrock, B., Zhu, Y., Wu, Y., Lu, H., & Zhu, S. (2019). A tale of two explanations: Enhancing human trust by explaining robot behavior. *Science Robotics*, *4*(37). https://doi.org/10.1126/scirobotics.aay4663

Endsley, M. R. (1995). Toward a theory of situation awareness in dynamic systems. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *37*(1), 32–64. https://doi.org/10.1518/001872095779049543

Faul, F., Erdfelder, E., Lang, A., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175–191. https://doi.org/10.3758/bf03193146

Greenwood, F., Nelson, E. L., & Greenough, P. G. (2020). Flying into the hurricane: A case study of UAV use in damage assessment during the 2017 hurricanes in Texas and Florida. *Plos One*, *15*(2). https://doi.org/10.1371/journal.pone.0227808

Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (task Load Index): Results of empirical and theoretical research. *Advances in Psychology*, 139–183. https://doi.org/10.1016/s0166-4115(08)62386-9

Hong, A., Igharoro, O., Liu, Y., Niroui, F., Nejat, G., & Benhabib, B. (2018). Investigating human-robot teams for learning-based semi-autonomous control in urban search and rescue environments. *Journal of Intelligent & Robotic Systems*, *94*(3–4), 669–686.

Honig, S., & Oron-Gilad, T. (2021). Expect the unexpected: Leveraging the human-robot ecosystem to handle unexpected robot failures. *Frontiers in Robotics and AI*, *8*. https://doi.org/10.3389/frobt.2021.656385

Hsiung, C.-P., & Chiou, E. (2019). Attribution biases and trust development in physical human-machine coordination: Blaming yourself, your partner or an unexpected event. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *63*(1), 211–211. https://doi.org/10.ghzbnz

Khasawneh, A., Rogers, H., Bertrand, J., Madathil, K. C., & Gramopadhye, A. (2019). Human adaptation to latency in teleoperated multi-robot human-agent search and Rescue Teams. *Automation in Construction*, *99*, 265–277. https://doi.org/10.1016/j.autcon.2018.12.012

Lakhmani, S., Abich, J., Barber, D., & Chen, J. (2016). A proposed approach for determining the influence of multimodal robot-of-human transparency information on human-agent teams. *Lecture Notes in Computer Science Foundations of Augmented Cognition: Neuroergonomics and Operational Neuroscience*, 296–307. https://doi.org/10.1007/978-3-319-39952-2_29

Lee, J. D., & See, K. A. (2004). Trust in automation: Designing for appropriate reliance. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *46*(1), 50–80. https://doi.org/10.1518/hfes.46.1.50_30392

Lewis, M., Sycara, K., & Walker, P. (2018). The role of trust in human-robot interaction. *Foundations of Trusted Autonomy*, *117*, 135–159. https://doi.org/10.1007/978-3-319-64816-3_8

Lyons, J. B. (2013). *Being transparent about transparency: A model for human-robot interaction*. AAAI spring symposium series.

Mercado, J. E., Rupp, M. A., Chen, J. Y. C., Barnes, M. J., Barber, D., & Procci, K. (2016). Intelligent agent transparency in human–agent teaming for multi-UxV management. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *58*(3), 401–415. https://doi.org/10.1177/0018720815621206

Miller, T. (2019). Explanation in artificial intelligence: Insights from the social sciences. *Artificial Intelligence*, *267*, 1–38. https://doi.org/10.1016/j.artint.2018.07.007

Norman, D. A. (1994). How might people interact with agents. *Communications of the ACM*, *37*(7), 68–71. https://doi.org/10.1145/176789.176796

Phillips, E., Ososky, S., Grove, J., & Jentsch, F. (2011). From tools to teammates: Toward the development of appropriate mental models for intelligent robots. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *55*(1), 1491–1495. https://doi.org/10.1177/1071181311551310

Rousseau, D. M., Sitkin, S. B., Burt, R. S., & Camerer, C. (1998). Not so different after all: A cross-discipline view of trust. *Academy of Management Review*, *23*, 393–404.

Schaefer, K. E. (2016). Measuring trust in human robot interactions: Development of the "trust perception scale-HRI". In: *Robust intelligence and trust in autonomous systems* (pp. 191–218). https://doi.org/10.1007/978-1-4899-7668-0_10

Selkowitz, A. R., Larios, C. A., Lakhmani, S. G., & Chen, J. Y. (2016). Displaying information to support transparency for autonomous platforms. *Advances in Intelligent Systems and Computing*, 161–173. https://doi.org/10.1007/978-3-319-41959-6_14

Shah, J., & Breazeal, C. (2010). An empirical analysis of team coordination behaviors and action planning with application to human–robot teaming. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *52*(2), 234–245. https://doi.org/10.1177/0018720809350882

Wang, N., Pynadath, D. V., Rovira, E., Barnes, M. J., & Hill, S. G. (2018). Is it my looks? Or something I said? The impact of explanations, embodiment, and expectations on trust and performance in human-robot teams. *Persuasive Technology Lecture Notes in Computer Science*, 56–69. https://doi.org/10.1007/978-3-319-78978-1_5

Wong, M., Ezenyilimba, A., Wolff, A., Anderson, T., Chiou, E., Demir, M., & Cooke, N. J. (2021). *A remote synthetic testbed development for human-robot teaming: An iterative design process during the COVID-19 pandemic*. 65th Annual Meeting of the Human Factors and Ergonomics Society.

Yagoda, R. E., & Gillan, D. J. (2012). You want me to trust a ROBOT? The development of a human–robot interaction trust scale. *International Journal of Social Robotics*, *4*(3), 235–248. https://doi.org/10.1007/s12369-012-0144-0

Akuadasuo Ezenyilimba is a Human Systems Engineering PhD student at Arizona State University. She holds an M.S. in Human Systems Engineering from Arizona State University and an M.S. in Applied Psychology from Sacred Heart University. She is currently a National Science Foundation Research Trainee and assists in the Center for Human AI Robot Teaming (CHART) lab, led by Dr. Nancy Cooke. Her research interests are human computer interaction, traumatic brain injury rehabilitation, and smart cities.

Margaret Wong is a master's student at Arizona State University studying Human Systems Engineering. She currently works as a graduate research assistant in ASU's Polytechnic School at the Center for Human, AI, & Robot Teaming (CHART) lab, which is led by Dr. Nancy Cooke. Her research interests include human–robot interaction and teaming, healthcare usability, education research, and user experience research.

Mr. Alexander Hehr is a master's student in the Human Systems Engineering department at Arizona State University. His research interests are focused on human–robot teaming, and he hopes to eventually find a place in the aerospace industry.

Mustafa Demir is currently an assistant research scientist working at Ira. A. Fulton Schools of

Engineering at Arizona State University. Dr. Demir received his Ph.D. in Simulation, Modelling, and Applied Cognitive Science, focusing on team coordination dynamics in Human-Autonomy Teaming from Arizona State University in Spring 2017. His current research interests are human-autonomy teaming, team cognition, dynamical systems theory, and advanced statistical modeling.

Alexandra T. Wolff is a Human Systems Engineering master's student and graduate research assistant in the Center for Human AI Robot Teaming (CHART) lab at Ira A. Fulton Schools of Engineering at Arizona State University. She holds a Bachelor of Science in Human Systems Engineering from Arizona State University. Her current research interests include human–autonomy teaming, team cognition, and unmanned vehicles.

Erin K. Chiou is an assistant professor of Human Systems Engineering at Arizona State University where she directs the Automation Design Advancing People and Technology (ADAPT) laboratory. Erin earned a B.S. in psychology and philosophy from the University of Illinois at Urbana-Champaign and an M.S. and Ph.D. in industrial and systems engineering from the University of Wisconsin-Madison. Her current research interests center around social factors in human-automation performance, including trust, accountability, and interaction structures.

Nancy Cooke is a professor of Human Systems Engineering and directs the Center for Human, Artificial Intelligence, and Robot Teaming at Arizona State University. She received her Ph.D. in Cognitive Psychology from New Mexico State University.