International Journal of Human–Computer Interaction

ISSN: 1044-7318 (Print) 1532-7590 (Online) Journal homepage: www.tandfonline.com/journals/hihc20

Should AI Teammates Give All the Answers? Examining the Role of Different AI Information-Sharing Techniques on Team Cognition in Human-AI Teams

Beau G. Schelble, Rohit Mallick, Allyson Hauptman & Nathan McNeese

To cite this article: Beau G. Schelble, Rohit Mallick, Allyson Hauptman & Nathan McNeese (24 Jul 2025): Should AI Teammates Give All the Answers? Examining the Role of Different AI Information-Sharing Techniques on Team Cognition in Human-AI Teams, International Journal of Human–Computer Interaction, DOI: 10.1080/10447318.2025.2528988

To link to this article: https://doi.org/10.1080/10447318.2025.2528988

Published online: 24 Jul 2025.

Submit your article to this journal

Article views: 60

View related articles

View Crossmark data

Full Terms & Conditions of access and use can be found at
https://www.tandfonline.com/action/journalInformation?journalCode=hihc20

INTERNATIONAL JOURNAL OF HUMAN–COMPUTER INTERACTION
https://doi.org/10.1080/10447318.2025.2528988

Taylor & Francis
Taylor & Francis Group

Check for updates

Should AI Teammates Give All the Answers? Examining the Role of Different AI Information-Sharing Techniques on Team Cognition in Human-AI Teams

Beau G. Schelble[a], Rohit Mallick[b], Allyson Hauptman[c] and Nathan McNeese[b]

[a]Industrial and Systems Engineering, The University of Tennessee, Knoxville, Tennessee, USA; [b]Division of Human-Centered Computing, Clemson University School of Computing, Clemson, South Carolina, USA; [c]Electrical Engineering & Computer Science, West Point Military Academy, Army Cyber Institute, West Point, New York, USA

ABSTRACT

This study investigated how the information-sharing traits of an AI teammate, such as augmenting team memory (ATM) and changes in intra- and extra-team information (IET), can enhance situation awareness (SA) and responses to unexpected events in human-AI teams. Thirty-one teams of two participants and one AI flew a simulated UAV task with the AI's information-sharing attribute set to ATM, IET, or control. Results showed that IET teams, whose AI provided direct solutions to task disruptions, were the most likely to overcome them. The ATM teams, whose AI helped teammates find the solution themselves, also outperformed the control group, but also exhibited more action communication, a better perceived shared mental model with the AI, and a more positive perception of SA compared to the IET condition. These results highlight that while providing direct, immediate AI solutions can improve responses to task disruptions, there can be a cost to team development.

KEYWORDS

Human-AI team; situation awareness; team cognition; AI information-sharing; human systems integration

1. Introduction

In recent years, technology has quickly progressed from brittle automation tools to flexible AI systems capable of accurately perceiving and adapting to dynamic situations (Endsley, 2017). This growth in applicability has allowed humans to recognize AI's potential, increasingly trusting them with heightened roles and responsibilities due to their computational advantages. In doing so, the perception of AI as a limited tool has undergone a paradigm shift where it has merit to be considered a collaborative teammate to humans (Lyons et al., 2021). This human-AI team (HAT) organizational structure has demonstrated effectiveness across several domains, including military applications, healthcare, and cybersecurity (Sarker et al., 2023; Schelble et al., 2020). Such fields have defined and expressed a need for AI teammates; however, each is a demanding dynamic environment calling for the best in precision, teamwork efficiency, and performance, as lives, significant financial sums, or sensitive data rely on the team's performance. In these use cases, humans and AI execute decision-making processes by proactively communicating information with each other (O'Neill et al., 2022; Demir et al., 2017), with the best teams anticipating information-sharing needs without even needing to inquire. This anticipatory information-sharing relies on adequate situation awareness (SA) throughout the team, which AI teammates currently struggle to implement effectively (Zhang et al., 2023). Specifically, AI teammates have difficulty anticipating information-sharing needs, what information is necessary, when, and why, especially across novel contextual circumstances (Demir et al., 2018; Endsley, 2023). As such, for HATs to eventually be deployed successfully into novel high-stakes environments, AI teammates' ability to understand their place within team SA and contribute to it through effective information-sharing must improve as HATs continue to be deployed worldwide.

CONTACT Beau G. Schelble bschelbl@utk.edu 515 John D. Tickle Engineering Building, 851 Neyland Drive, Knoxville, TN 37996-2315
© 2025 Taylor & Francis Group, LLC

One of the core components of facilitating SA is the effective sharing of information between teammates, but AI teammates will need to be designed to do so in a unique way. Efficient teams that demonstrate strong SA share benefits in understanding why decisions are made, how they advance towards task completion, and anticipate what actions they are required to execute to support their teammates (Cannon-Bowers et al., 1993). However, challenges arise when team SA is inaccurate and thus produces ineffective team decision-making with maladaptive actions (Endsley, 2023). Since HATs are often deployed in dynamic environments, mistakes can be costly and even endanger human lives. AI teammate information-sharing must be well understood to implement in HATs effectively (Cooke et al., 2024; Jiang et al., 2023), and task-related information-sharing by AI teammates stands as an excellent starting point given AI's existing ability to excel in taskwork (Koch, 2016). Prior studies also indicate that the content of the information shared has been noted to vary in effectiveness (Schelble et al., 2025), meaning that the nature of the information shared can impact a team's collective ability to navigate disruptive situational events (Demir et al., 2018). How an automated or autonomous system should respond to perceived task disruptions has long been a topic of interest in SA research, as overreliance and related challenges, such as the lumberjack effect, can stunt humans' SA and shared mental model (SMM) development, reducing human capabilities, knowledge, and performance (Endsley, 2015; Jones & Endsley, 1996). New AI systems and AI teammates face an even greater challenge as their ability to step in rapidly and address task disruptions is a significant advantage, but overreliance remains a challenge, as detailed in the seven human-computer interaction (HCI) grand challenges (Stephanidis et al., 2025). An example of this overreliance in collaborative AI systems is LLMs as coding assistants, which improve programming speed and accuracy but reduce developer comprehension and creative problem-solving for the actual system built alongside the AI (Liu et al., 2023). As such, research that explicitly investigates how AI teammates may be designed to better contribute to team SA through effective information-sharing is critical to developing better AI teammates, especially when understanding how AI should communicate the proper response to task disruptions.

Furthermore, AI information-sharing needs to go beyond just task completion, as teams bolster SA in between tasks during what is known as a transition phase. Understanding exactly when and how AI can meaningfully contribute to SA within HATs outside of action phases (i.e., during tasks) thus becomes a critical process to support and reduce information inaccuracy. These teaming phases dictate how teams interact, with action phases focusing directly on goal-related tasks, while transition phases see teams plan and review (Marks et al., 2001). These transition phases can be seen as concrete planned events, such as debriefs and after-action reviews, where teams engage in strategy formation, goal setting, and planning (Marks et al., 2001). Interestingly, the effect of transition phases on HATs' effectiveness has yet to be studied in detail alongside action phases, let alone its impact on team SA development in concert with AI information-sharing strategies (O'Neill et al., 2022). While current research has not made the leap to include transition phases with action phases for HATs, the utility of AI information-sharing in action phases is clear. Current AI information-sharing research occurring solely in the action phase centers on topics such as decision-making support, transparency, and explainability (Mercado et al., 2016; Yan & Gurkan, 2023), which show generally positive effects of information-sharing by an AI teammate but primarily focus on enhancing human teammates' shared knowledge of the AI teammate exclusively. Further, explainability and transparency information on the AI teammate can sometimes reduce team performance, especially with more experienced human teammates (Paleja et al., 2021). As such, it is critical to examine AI information-sharing's ability to contribute to HAT effectiveness and shared knowledge throughout *all* phases of teaming, given that the unique dynamics of the transition phase can induce effects on the outcomes produced in the action phase (Tannenbaum & Cerasoli, 2013).

Using these gaps in the literature as motivation, the following research questions focusing on how AI teammates can effectively engage in information-sharing to contribute to shared knowledge are proposed:

RQ1: Do AI teammates designed to engage in different forms of information-sharing affect performance and shared knowledge constructs in HATs?

RQ2: Are shared knowledge constructs affected by AI information-sharing providing direct versus indirect solutions to task disruptions?

RQ3: How does AI information-sharing during transition phases affect shared knowledge constructs and performance in HATs across their life cycle?

To address these gaps, 30 HATs were analyzed after each team completed four missions of an aerial system task (CERTT-RPAS) with three types of AI teammate information-sharing designs and two transition phase participation schedules. These manipulations allowed the study to understand how these implementations of AI information-sharing design can strengthen shared knowledge constructs, such as SMMs and SA within HATs. This work provides three key contributions across several concepts in HCI. First, by investigating various AI information-sharing strategies an AI teammate employs, this study clarifies how AI can better actively contribute to their team's shared knowledge. Second, this work empirically sheds light on the role transition periods play within HATs by detailing the impact and role of AI teammates participating in these critical team information-sharing processes. Finally, this study demonstrates the benefits and detriments of AI teammates responding to task disruptions with direct and indirect solutions on SMM development, coordinated perception and action communication, and even perceived situation awareness. Together, these contributions can strengthen the relationship between AI systems and humans as they shift from pure taskwork expertise and begin to support more and more of the complex teamwork behaviors that teams rely heavily on.

2. Background

Reviewing the relevant HCI research is necessary to outline the motivations behind and the design of this study. The current section will first discuss SA and how it can be classified before reviewing the role and challenges of SA development in HATs. Finally, it will address the research and possibilities for designing AI to support better team SA. Each of these topics motivates the current study to improve the effectiveness and applicability of HATs.

2.1. Situation awareness in human-AI teams

SA is vital to overall task effectiveness across various domains and contexts (Bolstad et al., 2005). SA consists of three distinct levels: 1) Level 1 is the person's perception of their surrounding environment; 2) Level 2 is the person's comprehension of their current situations; and 3) Level 3 is the person's ability to project the situation into the future (Endsley, 1995; Endsley & Jones, 1997). The concept of SA extends from the individual level to the team level as a unique form of shared understanding known as team SA. Team SA builds on and takes into account each team member's awareness of their environment, tasks, time, and projection of future events (Endsley et al., 2003). While team SA stems from each individual's awareness, it represents more than a mere summation of each member's SA. Team SA goes beyond a summation of individual SA by including the team's common perspective and understanding of each member's unique role and responsibilities (Bolstad et al., 2005; Wellens, 1993). These additional elements of team SA depend on team behaviors, such as monitoring and backup behaviors, and the specific type and composition of the team (Salas et al., 1995). As such, a critical difference between individual SA and team SA is the requirement for communication and coordination amongst the team members (McNeese et al., 2021).

Many types of teams differ in one way (e.g., composition, goal, life span); however, one particular kind that is increasingly important is the HAT. These teams are composed of at least one human and one autonomous agent to fulfill unique, interdependent roles in pursuit of a common goal (O'Neill et al., 2022). HATs can leverage the technical strength of AI to fulfill capability gaps in new, data-intensive environments (Nyre-Yu et al., 2019). When humans and AI work together, such teams can increase their collective awareness of their goals and overall productivity (Damacharla et al., 2019; Schaefer et al., 2017). When humans and AI work together, such teams can increase their collective awareness of their goals and overall productivity based on their communication. Previous examples discuss the need to communicate intent to engender trust in human-agent teams via transparency implementations in the interface (Chen et al., 2018; Schaefer et al., 2017) or through linguistic emotional expressions (Mallick et al., 2024). These works utilize AI in contexts where the AI performs passive communication updates when missions are performing as expected, without further examining how an AI teammate may contribute during an unexpected event. As a result, previous explorations demonstrate various SA-enhancing communication strategies (Chen et al., 2018; Chiou et al., 2022; Ezenyilimba et al., 2023); however, further analysis is needed in contexts where AI takes an active role in and outside of catastrophic events that disrupt team SA simultaneously. This study addresses this gap by examining information-sharing, communicative needs, and strategies during unexpected events alongside typical functioning.

2.2. AI support for team situation awareness

While understanding how various types of information-sharing help enhance SA, communication, by and large, is multifaceted, such that the timing and context of when communication is provided factor into how that information is perceived (Demir et al., 2017; Zhang et al., 2023). The Input-Process-Output (IPO) and Input-Mediator-Output-Input (IMOI) models of team effectiveness are used to characterize the episodic nature of teaming, where teams experience two distinct phases in accomplishing their goals (Marks et al., 2001). The first of these phases is the transition phase, where the team evaluates past phases and plans to guide members towards their shared goal (Marks et al., 2001). This phase includes processes such as mission analysis (Fleishman & Zaccaro, 1992; Prince & Salas, 1993), goal specification (Dickinson & McIntyre, 1997; Prussia & Kinicki, 1996), and strategy formulation (Gladstein, 1984; Prince & Salas, 1993). Next, the action phase is when teams are actively engaged in acts that directly contribute to accomplishing the team's overall goal (Marks et al., 2001); such acts include monitoring progress towards goals (Jentsch et al., 1999), team monitoring, back-up behaviors (Dickinson & McIntyre, 1997; Kude et al., 2019), and coordination behaviors (Demir et al., 2015, 2019). It is important to note the cyclical nature of these phases, where the results of the action phase feed the discussions and focus of the subsequent transition phase. This transition phase cycle is integral to allowing teams to analyze and improve prior performance in the following action phase as it formally engages in strategy formation, planning, and goal setting, such as an after-action review or debriefing (Marks et al., 2001; Weger et al., 2022). Effective communication and analysis of information during this phase have been shown to increase team performance in human-only teams (Weger et al., 2022).

In teams relying heavily on technology, additional visual cues and communication can be integral to establishing key team processes (Gergle et al., 2013). The success of these strategies has led to recent research on how to design AI and interfaces for HATs that can better build and support SA for their human teammates (Mercado et al., 2016; Schelble et al., 2022b). Like humans, AI agents maintain a conception of what is happening around them, an incomplete viewpoint contributing to overall team SA (Munir et al., 2022). To contribute to a team's SA, the AI must share relevant information through focused design features, such as transparency and explainability (Endsley, 2023). This sharing not only contributes to team SA but also to team memory, which is an essential component of successful knowledge sharing in teams (Fiore et al., 2006). However, while some research studies have shown that increased explainability and transparency can enhance team SA (National Academies of Sciences, Engineering & Medicine, 2021), too much can also cognitively overload human teammates and harm team performance (Wright et al., 2016). Thus, AI's unique capabilities and information-sharing must be carefully managed to enhance team SA in HATs. One example is an AI agent's ability to quickly and comprehensively process information and distribute it amongst the team (Skaug & Busch, 2022). AI agents can be designed to provide critical, timely feedback to teammates efficiently and accurately (Schulman et al., 2017). Several frameworks exist regarding how AI teammates can support key elements of team cognition, such as team SA, working memory, and sensemaking. This study utilizes these frameworks focused on the autonomous support of team cognition constructs (Cooke et al., 2013; Cuevas et al., 2007; Endsley, 2023). Notably, the work by Cuevas and colleagues in 2007 outlines a framework where autonomous teammates assist in team cognition by directly supporting sensemaking, working memory, team SA, SMMs, monitoring, and other related team processes that are critical for achieving effective team outcomes (Cuevas et al., 2007).

AI teammates can enhance these cognitive team processes by facilitating information sharing that raises awareness of team members' statuses and actions, correcting team errors, tracking progress, and updating team members on significant changes occurring within and outside the team. These strategies align closely with the essential role of effective and timely communication in improving team SA, sensemaking, and overall performance, as emphasized by Cooke and colleagues concerning interactive team cognition and Endsley regarding SA (Cooke et al., 2013; Endsley, 1995; 2023). However, despite such technical advantages, HATs have difficulty performing the traditional push-pull communication strategies integral to effective team situational awareness development and sustainment (Demir et al., 2018). These limitations leave significant gaps in the literature to explore how AI teammates may improve their coordination strategies throughout the entire teaming life cycle.

3. Methods

The current study represents an in-person experiment utilizing a mixed-methods design. It involved teams of three completing four missions together over three hours (Schelble, 2023).

3.1. CERTT UAS-STE task

The experimental task used was the Cognitive Engineering Research on Team Tasks Uncrewed Aerial System-Synthetic Task Environment (CERTT UAS-STE). This task environment is based on the United States Air Force Predator drone or uncrewed aerial system (UAS) ground control system, giving it extreme practical applicability and external validity (Cooke & Shope, 2004). Further still, the CERTT UAS-STE has a long track record in team research (Cooke & Shope, 2004; Demir & Cooke, 2014; McNeese et al., 2018) and team cognition research (Cooke et al., 2007, 2013; Gorman et al., 2020). The system has also been extensively used to analyze HATs (Demir et al., 2018; McNeese et al., 2018).

The CERTT UAS-STE task contains three distinct interdependent roles: 1) the pilot (AVO) operates the UAS heading, airspeed, and altitude according to a flight plan sent by the navigator role (see Figure 1); 2) the navigator (DEMPC) develops and provides the flight plan and any specific flight restrictions such as speed and altitude (see Figure 1); and 3) the photographer (PLO) monitors the sensors and is in charge of taking photographs of the actual targets following the current altitude, airspeed, and target specific camera requirements (e.g., zoom level and camera type) (see Figure 1). The goal of each mission was to identify targets within restricted zones (ROZ), fly the UAS to those ROZs through the entry waypoint, follow the airspeed and altitude restrictions for each target, photograph each target using the required settings, and fly the UAS out of the ROZ through the exit waypoint as an interdependent team. Each mission was developed with three ROZs that contained either six or seven targets (two to three targets per ROZ) and mirrored one another in difficulty. The equipment used for the CERTT UAS-STE included five computers, with each participant utilizing three monitors to display the two role-specific displays and a chat screen. Two participants recruited for the study took the AVO and PLO roles, while the AI teammate took the DEMPC role. The AI teammate was placed into the DEMPC role given AI systems' success in navigation and route planning in solo and collaborative environments (Baaj & Mahmassani, 1991; Hu et al., 2020; Loske & Klumpp, 2021). The AI teammate was also designed to provide information with perfect accuracy and high speed for their base-level DEMPC role and additional SA-related information-sharing. These tasks typically require a human to engage in more prolonged training and unnecessarily expose the team to further errors. Using the CERTT UAS-STE, the current experiment developed four 20-minute missions, and one training mission 15 min in length.

Because the CERTT task consisted of a series of missions spread over three hours, teams could engage in formal transition phase discussions between missions. These discussions focused on improving team strategy, goal alignment, and engaging in performance review (Marks et al., 2001). Implementing this transition phase in the CERTT task meant that participants received a description of the purpose of the transition phase from the researcher beforehand (e.g., to analyze and discuss past

Figure 1. View of the pilot/AVO's console (Orange-Top left), view of the photographer/PLO's console (Green-Top right), and view of the navigator/DEMPC's console (Blue-Bottom).

performance and plan for future missions). Then, the transition phase started once the researcher navigated the participants' tablets to the Slack application, where a text-based conversation occurred in a locked session-specific channel. These transition phases lasted for six minutes, at which point the researcher directed the participants to conclude their discussion and move on to the next task in the experiment.

3.2. Experimental design

Two between-subjects manipulations were implemented to address the RQs: 1) order of AI transition phase participation and 2) AI information-sharing attribute. Specifically, the AI transition phase participation order was crossed with AI information-sharing attribute, consisting of two levels: 1) first transition participation and 2) second transition participation. The AI information-sharing attribute contained three between-subjects levels: 1) augmenting team memory, 2) intra/extra team information changes, and 3) control. The ATM and IET actions were developed by identifying realistic ways an AI teammate could support tasks such as sensemaking, SA, working memory, attention, SMMs, and monitoring within the CERTT environment. Collaborating with CERTT developers and utilizing existing task analyses, the actions focused on enhancing sensemaking and working memory related to ROZs to augment team memory. Additionally, the study implemented features to notify the team of dusk and aid in correcting errors, addressing intra- and extra-team information changes. These manipulations are described in more detail in Section 3.3. Given the many levels of the manipulations and the length of their names, they will all be referred to using acronyms moving forward (see Table 1).

3.3. AI teammate design and implementation

The AI teammate in all conditions was represented utilizing the Wizard of Oz methodology by having a confederate researcher portray the AI teammate's actions and communication to participants without their knowledge (Maulsby et al., 1993). This communication was done using text-based chat on one of the three computer monitors, the standard communication method for all team members. WoZ was also utilized for the AI teammate as it was highly effective in previous CERTT studies involving HATs and SA (Demir et al., 2019, 2020; McNeese et al., 2018). Specifically, the AI teammate's activities within the simulation and its chat communications were scripted, and these scripts were developed from several pilot sessions. The scripts were developed to align as closely to the standard performance of modern large language models (e.g., ChatGPT, LLaMa, and Gemini) as possible. This meant ensuring high speed and accuracy in the communication experience to ensure participants were convinced they were working with an AI teammate, which could only be achieved using a well-developed script. A manipulation check was conducted at the end of the experiment to ensure this manipulation was convincing to participants. Those participants who did not believe they were working with an AI teammate were not included in the subsequent analysis (one team).

3.3.1. AI teammate information-sharing

The three AI information-sharing attribute levels were pilot-tested by fine-tuning their specific manipulations to adequately accommodate the task and their particular attribute (see Table 2). First, the control AI teammate provided the team with the information required to ensure the duties of the DEMPC were carried out in full and that the team had all the information they needed to accomplish their shared goal effectively and nothing more. The two non-control AIs gave the same required DEMPC-specific information as the control AI, but then provided additional information relative to their specific attribute to enhance team SA. Second, the ATM AI teammate offered additional information to the team members to augment their memory of the task and how the team should coordinate to accomplish the task. This information included the number of targets in the upcoming ROZ, an airspeed range that accommodated the most targets, whether that speed needed to change for a specific target, and when that change would be necessary. Examples of this information-sharing from the ATM AI were as follows: *"Planning ahead for ROZ 2, it contains three targets," "A speed of 200 satisfies the first two targets but then it must be raised to 300-400 after SEN-1," "I will remind you of this change after SEN-1," "Remember, after completing this target, we must increase our speed to at least 300."* During an SA roadblock, the ATM AI would augment the team's memory by assisting with the recognition of information needs and information availability.

Third, the IET AI aided the team through information-sharing by supporting their awareness and ability to quickly respond to changes occurring within and outside the team. The IET AI accomplished this by alerting the team when dusk had fallen to help avoid bad photos, alerting the team whenever the UAS was not meeting the airspeed or altitude restrictions of a target waypoint, and suggesting a fix for any airspeed or altitude violations. Examples of this information-sharing attribute were as follows: *"Our current airspeed is too fast AVO, the airspeed restriction is XX," "Please decrease airspeed to be within XX," "As a reminder, this UAV should not exceed 380 kts, and cannot exceed 400 kts," "Dusk has now fallen PLO. Your aperture settings may need to be corrected to account for this change to take good photos of targets."* During an SA roadblock, the IET AI would support the team's awareness of the change within the team by detailing the obstacle and the specific fix needed. As such, the two non-control AI information-sharing attributes each focus on individual aspects of teaming. Specifically, the ATM AI provides information supporting team memory relating to upcoming task-related events and general task and team knowledge for shared mental models (i.e., who needs what information to accomplish their taskwork). Alternatively, the IET AI directly supports team awareness of environmental and internal changes affecting the team and provides direct solutions that effectively manage those changes.

Table 1. Acronym key for manipulation levels.

Condition	Acronym
AI Participates in the First Transition Phase	**PN**
AI Participates in the Second Transition Phase	**NP**
Augmenting Team Memory	**ATM**
Intra/Extra Team Information Changes	**IET**
Control	N/A

Table 2. The actions taken by the AI teammate for their information-sharing attributes.

AI teammate condition	General SA information-sharing	SA roadblock information-sharing	Transition phase information-sharing
ATM	1. DEMPC Responsibilities 2. Number of Targets in Upcoming ROZ 3. Airspeed Ranges Accommodating Upcoming Targets 4. Reminders for When Airspeed Ranges Changed within ROZs	1. Notify Team of Failure 2. Help Team Understand, Find, and Communicate the Data Needed to Mitigate Roadblock	1. Can Answer Basic Questions About CERTT Rules and Functions 2. Can Answer Questions Directly Related to the ATM Information Shared in Missions
IET	1. DEMPC Responsibilities 2. Notification if Airspeed/Altitude Was Causing a Warning 3. Solution to Airspeed/Altitude Warning 4. Notification if Dusk Had Fallen Changing Photo Settings	1. Notify Team of Failure 2. Provide the Exact Data Team Needs to Overcome Roadblock	1. Can Answer Basic Questions About CERTT Rules and Functions 2. Can Answer Questions Directly Related to the IET Information Shared in Missions
Control	1. DEMPC Responsibilities	1. N/A	1. Can Answer Basic Questions About CERTT Rules and Functions

3.3.2. AI teammate transition phase information-sharing

The AI teammates were not allowed to engage in any communication that did not directly relate to their information-sharing feature and only responded in the affirmative or negative in the case of requests of the AI teammate for future missions. This limitation was implemented given the breadth of the tasks AI teammates would have needed to engage in to support all possibilities of the transition phase discussion, and to simplify the development of the script. For example, the IET AI was described as being capable of responding to questions regarding alarms and system failures. The ATM AI could respond to and provide information on matters such as restrictions for ROZs and the information accessible and needed by each team member. The control AI was described as capable of answering questions related to its team role and the CERTT UAS-STE task in general. The AI teammate also responded in the affirmative when asked to adhere to an effective tactic or strategy, but responded in the negative when asked to adhere to a detrimental tactic or strategy (to avoid any potential effects of confusion or neglect by participants). If participants asked a question outside the purview of the AI teammates' defined abilities, they would respond that they could not answer because it was outside their capabilities. The participants were made aware of these boundaries and capabilities for their AI teammates before and during the hands-on training session.

3.3.3. AI teammate situation awareness roadblock information-sharing

One feature of the CERTT system is the ability to implement task disruptions, or SA roadblocks, to measure teams' SA through their ability to respond to the roadblock. The current study utilized temporary failures of the data readouts for AVO and PLO. Specifically, information related to navigation for AVO and distance from waypoints for PLO. These failures occurred during the second ROZ of each mission, excluding the training mission and occurring roughly halfway through the mission. The roadblock affected the PLO's role in the first two missions and the AVO's role in the last two. During a failure, one of the human teammates had their data readouts fail, and the other did not. To overcome the roadblock, the team needed to communicate what information was occluded and what data they needed to continue working effectively, and then communicate that data to one another. The team always had all the information needed to overcome the failure, even in the control AI teammate condition. To assess the effectiveness of AI information-sharing abilities in supporting SA, the three AI teammates provided distinct information in response to SA roadblocks. The ATM AI would notify the team of the system failure, what information was occluded from which teammate, and what information that teammate needed to accomplish their job effectively. An example of the ATM AI's response to an SA roadblock was as follows: *"I have detected a system failure in the UAV. The AVO screen containing the UAVs airspeed, altitude, and positional data has stopped reporting. I will remind the team what information the AVO needs to do their job given the constraints of this failure," "The AVO must know the correct bearing they should take towards the target if they are deviating," "During system failure, AVO will need to know the correct bearing that can be found on PLOs equipment."* The IET AI notified all team members of the system failure and what information was occluded from what teammate, and then explained to the team that they would provide the necessary information to the teammate to still accomplish their job effectively. An example of the IET AI's response was as follows: *"I have detected a system failure in the UAV. The AVO screen containing the UAVs airspeed, altitude, and positional data has stopped reporting. I will help AVO by providing the correct bearing needed for this target," "The correct bearing to take is XX," "The distance to the target is approximately."* In these examples for the AVO failure, the AVO needed the correct bearing given to them by the PLO to adjust the heading of the UAV properly. The examples also show how the IET AI teammate told the team a failure had occurred, that information had been occluded from AVO, and then gave AVO the exact bearing they needed directly. Alternatively, the ATM AI teammate told the team that a failure had occurred, that the pilot had bearing information occluded, and that the photographer needed to communicate that bearing information to overcome the failure (which they always had on their screen).

Essentially, the IET AI gave teams the answer needed, while the ATM AI helped them find the answer independently. Lastly, the control AI teammate did not automatically provide any information regarding the SA roadblock, which meant teams had to use their unsupplemented SA to overcome it. The process of overcoming the SA roadblock for teams with the control AI would involve the AVO or PLO communicating what information was occluded from them, identifying what information they needed to be supplemented by a teammate, and then communicating that information. For a failure affecting the AVO, they would need the PLO to communicate the current and correct bearing, while a failure for the PLO would require the AVO to provide the current objective and distance to the objective. Lastly, the information needed to overcome the roadblock was available to all teammates, ensuring the AI teammate's actions remained in line with the actual capabilities of the DEMPC role itself. The DEMPC, AVO, and PLO roles had access to the same readouts for Current Waypoint, Queued Waypoint, Time to Waypoint, Distance to Waypoint, Bearing Required for the Current Waypoint, and Current Bearing Course Deviation. As such, the AI teammate in the DEMPC role was purely manipulated in how it shared this information with the rest of the team in response to a roadblock.

3.4. Participants

Seventy-four participants were recruited and divided into 37 teams; however, six teams were dropped entirely, with five teams being dropped due to technical difficulties and the other due to failing the manipulation check. This left 62 participants' data to be analyzed as a total of 31 teams, with at least ten teams per between-subjects condition. The control condition consisted of 11 teams, as one team's survey data was corrupted, but their CERTT data was recorded and retained, while another team was recruited. These 62 participants had an average age of 20.4 ($SD = 3.13$). Forty-five participants identified as women, and the rest identified as men. This research complied with the American Psychological Association Code of Ethics and was approved by the Institutional Review Board at Clemson University. Each participant provided informed consent before participating in the study.

All participants were recruited using flyers and a participant pool located at a large southeastern university in the United States. Those recruited using flyers posted around the university campus and social media posts to the university's related online communities on platforms like Reddit and Discord were compensated with a $30 Amazon gift card for their time. The participants recruited using the participant pool were compensated with course credit for their time.

3.5. Procedure

The study took three hours to complete (see Figure 2), and participants were randomly assigned to a condition and role on the team. They were then given an overview of the study and its purpose before being directed to review the informed consent document, answer demographic questions, and complete an interactive PowerPoint training session. At this point, the researcher opened the CERTT UAS-STE on each of the participants' computers and began the hands-on training mission with the participants, which resembled a real mission but was shorter in length (15 min). The researcher then utilized a training script with the participants to direct them on completing their tasks effectively at the individual and team levels. The hands-on training also allowed the participants to familiarize themselves with their AI teammate and its information-sharing attributes before starting their missions. Participants were also told that the AI teammate's communication ability was in line with current large language models but that it had been developed only to provide answers related to the CERTT system itself and its additional information-sharing attribute, meaning it could not provide information or conversation on unrelated topics. Once the first mission was completed, the participants completed one of the six post-task surveys on separate tablets with integrated keyboards before proceeding to the transition phase discussion using the Slack messaging application on those same iPads. These surveys were conducted on separate tablets with integrated keyboards and trackpads. The participants completed the first six-minute transition phase and were returned to their tablets for the post-transition phase survey. The participants then went on to complete the second mission and their third survey after that mission, before being given a short five-minute break due to the length of the experimental session. After the break, the participants went through the same process to complete missions three and four, with a transition phase between the two missions and a survey following each task. Once the final mission and post-task survey were completed, the researcher engaged in a five to ten-minute focus group interview with the two participants before they were debriefed, compensated, and finished with the study.

3.6. Measures

3.6.1. Team situation awareness (CAST)

Team SA was measured using task disruptions or "roadblocks," implemented in the current study using pre-programmed system failures built into the CERTT UAS-STE (Cooke et al., 2016). These roadblocks were utilized with the Coordinated Awareness of Situation by Teams (CAST) metric developed by Gorman and colleagues for the CERTT UAS-STE environment (Gorman et al., 2005). This method of assessing team SA was created through previous CERTT UAS-STE research and has frequently been utilized in HAT research on SA (Cooke et al., 2016; McNeese et al., 2018). The roadblocks were characterized by a temporary loss of information in one of the team's displays. Accordingly, the CAST metric considers a team's ability to perceive the roadblock, coordinate perception of the roadblock throughout the team, and coordinate action to mitigate the roadblock effectively (Gorman et al., 2010). CAST does this by having an experimenter monitor chat communication within the team to note the presence and direction of communications related to coordinated perception (CP) and coordinated action (CA) of the roadblock. Both CP and CA have their own set of communication boxes (shown in Figure 3), and each score is calculated by dividing the number of marked boxes by the total number of boxes available. However, because the current experiment manipulated the communication from the DEMPC teammate, the CP and CA metrics did not utilize the two boxes indicating DEMPC to PLO and

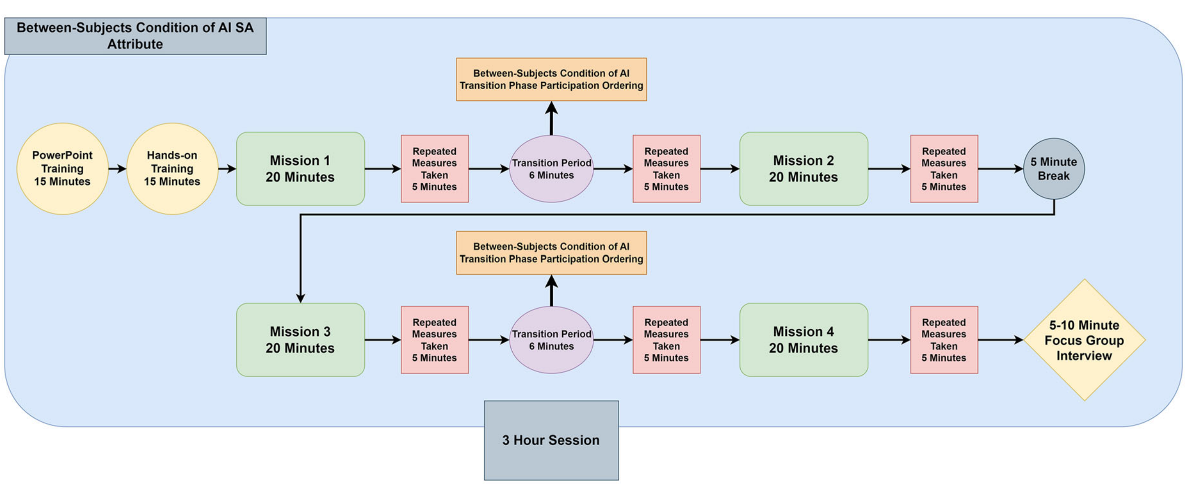

Figure 2. Study session timeline, indicating the frequency of repeated measures and the placement of transition periods along with the time associated for each.

DEMPC to AVO communication, meaning scores were calculated by four total boxes instead of six. As such, all communication recorded for CA and CP measures was from a participant, indicating the total amount of human CA and CP utterances within a team.

Using CP as an example, if Figure 3 showed the top left box for PLO and the left box for AVO as being marked, it would convey that the PLO teammate had received a communication informing them of the roadblock from the AVO teammate and then responded to AVO to confirm. In this example of CP scoring, the team would score 0.5 because two of the four possible human-origin (non-DEMPC) communication boxes were marked. CP is distinguished from CA as it is information related to the nature of the failure, such as what information is occluded, when it was occluded, or asking teammates to clarify the nature of their predicament. CA is then distinguished from CP when the utterances shift to taking action, such as asking another teammate to share specific information, and or sharing that information.

3.6.1.1. Overcoming situation awareness roadblock.

Teams were rated on their ability to successfully overcome the task disruption by whether they could perceive the failure, identify the necessary information, and then successfully communicate and act upon that information. As there was a failure affecting PLO in Rounds 1-2 and AVO in Rounds 3-4, this meant that in Rounds 1-2 the PLO or AVO needed to recognize that they needed to communicate and utilize the distance from the current target information to allow PLO to know when they were in range for a photo. In Rounds 3-4, the team needed to recognize that they needed to share and leverage the bearing information to ensure AVO could keep the UAV on course to the next target. The PLO and AVO roles had this information available throughout the experiment, allowing all teams to overcome these SA roadblocks, provided they had effective SA and a shared understanding of the task. Overcoming the SA roadblock resulted in a binary outcome of yes or no. This measure also represents whether or not a team has developed effective team SA, as demonstrated in past HAT research (McNeese et al., 2018).

3.6.2. Perceived situation awareness

The participants' perceived SA was measured using the Situation Awareness Rating Technique (SART) (Taylor, 1990). SART consists of nine items, each measured on a seven-point Likert scale ranging from "Very Low" to "Very High," and asks participants to rate those questions while retrospectively considering their experience in the previous action phase. Responses to these items were averaged for each participant, with higher values in the SART measure indicating lower levels of team SA.

3.6.3. Team performance

Team performance was measured for each team by mission, allowing for an evaluation of team effectiveness outcomes throughout the study. Specifically, team performance was measured using a series of metrics, which include the amount of time a team spent with alarms and warnings active, the amount of memory used in taking photos, the accuracy of the route sequence, targets visited, and the number of good, unique photographs of targets. Each team starts a mission with a score of 1000, and points are subtracted for each detrimental action or time spent in error. The final metric strongly indicates overall general team performance across the three individual roles, as seen in past HAT studies (Cooke et al., 2016; McNeese et al., 2018).

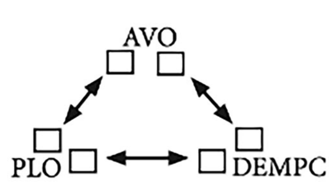

Figure 3. Situation Awareness logger.

3.6.4. Perceived shared mental model with AI

Participants' perceived SMM with the AI teammate was taken using a scale developed by van Rensburg and colleagues (Van Rensburg et al., 2021); however, this adaptation of the survey consisted of 10 items, each rated on a seven-point Likert scale ranging from "Strongly Disagree" to "Strongly Agree." The 10 items came from the execution, interaction, and temporal sub-scales within the five-factor mental model scale. These factors were chosen over equipment and composition as they were not as relevant to the current task, and team and task mental models are primarily the focus of team cognition research (Mathieu et al., 2000). The participants rated their perceived shared knowledge with the AI teammate across 10 items at each repeated measures point depicted in Figure 2. Responses for each set of ten items were averaged for each participant, and higher values indicated a greater perception of a SMM with the AI teammate.

4. Results

The data provided by the study was analyzed and reported as follows in the current section, using multilevel modeling to account for the structured nature of the data. The following analyses apply to all three RQs, each examining the effect of the AI teammate's information-sharing attribute, transition phase participation ordering, and time (as mission or transition phase number). Further, the results pertaining directly to teams' responses to the SA roadblock were organized together, given their unique tie to the information provided by the AI during those events. Conducting these analyses is imperative to understanding how AI teammates can be designed to contribute to team processes in shared knowledge development, such as SA, and whether AI teammates can encourage better SA-enhancing behaviors from human teammates themselves (Cuevas et al., 2007; Schelble et al., 2022a).

4.1. Analysis

The current data analysis utilized participant and team as clustering variables for multilevel random effects, which was supported by an average intraclass correlation coefficient of .534, computed across all of the dependent variables (Fox, 2015; McGraw & Wong, 1996). Only participants' perceived level of SA (SART) and SMM with the AI were clustered by participant and team, while the rest of the variables were measured at the team level and used only team as the clustering variable in the model. The AI SMM and SART variables displayed a left skew. However, no transformations were made given the robustness of multi-level models to such violations of normality (Schielzeth et al., 2020) and the limitations a gamma distribution model brings. A binomial model was used to evaluate teams' ability to overcome the SA roadblock, given that the data for this variable were either successful or unsuccessful. Given the issues in calculating effect sizes for linear mixed-effects models caused by being negative or a failure to increase monotonically as predictors are added, the marginal pseudo-R^2 suggested by Nakagawa and Schielzeth in 2013 was reported for each model (Nakagawa & Schielzeth, 2013). A residual analysis was conducted for all models to ensure assumptions were met (Fox, 2015; Rosopa et al., 2013), and the models were tested for significance using Type III Wald χ^2 ANOVA tests. All post hoc analyses were conducted using estimated marginal means, Tukey HSD corrections, and Kenward-Roger degrees of freedom. However, log odds ratios with turkey corrections were used to analyze teams' likelihood of overcoming the SA roadblock, with means and standard errors being back-transformed from the logit scale for readability.

4.2. CAST metrics in response to the situation awareness roadblock

The following metrics, CP, CA, and the likelihood of overcoming the SA roadblock, represent measures directly tied to teams' response to the SA roadblock occurring in each mission.

4.2.1. Coordinated perception (CAST)

A 3 (AI Information-Sharing Attribute: ATM, IET, Control) x 2 (AI Transition Phase Ordering: PN, NP) x 4 (Mission: M1, M2, M3, M4) linear mixed effects model was conducted to assess the effect of

AI information-sharing attribute (between-subjects), AI transition phase ordering (between-subjects), and mission (within-subjects) on the level of CP messages sent by human teammates. There was a statistically significant main effect of AI information-sharing attribute on CP messages sent by human teammates ($\chi^2(2) = 15.85$, $p < .001$; see Figure 4). The average level of CP was significantly greater with the control condition ($M = 0.37$, $SE = 0.05$) than the IET condition ($M = 0.13$, $SE = 0.05$). There was also a statistically significant main effect of mission on CP messages sent by human teammates ($\chi^2(3) = 25.84$, $p < .001$; see Figure 4). Specifically, Mission 1 ($M = 0.14$, $SE = 0.04$) had significantly fewer CP utterances than Mission 3 (\textit{M} = 0.36, $SE = 0.04$) and Mission 4 ($M = 0.30$, $SE = 0.04$). Lastly, Mission 2 ($M = 0.21$, $SE = 0.04$) had significantly fewer CP utterances than Mission 3 ($M = 0.37$, $SE = 0.04$). The marginal model $R^2 = .31$. There were no significant two-way interaction effects.

4.2.2. Coordinated action (CAST)

A 3 (AI Information-Sharing Attribute: ATM, IET, Control) x 2 (AI Transition Phase Ordering: PN, NP) x 4 (Mission: M1, M2, M3, M4) linear mixed effects model was conducted to assess the effect of AI information-sharing attribute (between-subjects), AI transition phase ordering (between-subjects), and mission (within-subjects) on the level of CA messages sent by human teammates. There was a statistically significant main effect of AI information-sharing attribute on CA messages sent by human teammates ($\chi^2(2) = 8.25$, $p = .016$; see Figure 5). This effect showed that the ATM attribute ($M = 0.40$, $SE = 0.06$) had higher levels of CA utterances than the IET attribute ($M = 0.18$, $SE = 0.06$). The marginal model $R^2 = .12$. There were no significant two-way interaction effects.

4.2.3. Overcoming situation awareness roadblock

A 3 (AI Information-Sharing Attribute: ATM, IET, Control) x 2 (AI Transition Phase Ordering: PN, NP) x 4 (Mission: M1, M2, M3, M4) linear mixed effects model was conducted to assess the effect of AI information-sharing attribute (between-subjects), AI transition phase ordering (between-subjects), and mission (within-subjects) on teams' likelihood of overcoming the system failure. This analysis revealed a significant main effect of AI information-sharing attribute ($\chi^2(2) = 18.54$, $p < .001$; see Figure 6). Estimated marginal means showed that the AI with the IET attribute ($M = 0.97$, $SE = 0.02$) significantly improved teams' likelihood of overcoming the system failure over those working with the ATM attribute ($M = 0.79$, $SE = 0.09$) and control attribute ($M = 0.27$, $SE = 0.09$), with the ATM attribute also giving teams a significant likelihood advantage over the control.

There was also a significant main effect of mission on teams' likelihood of overcoming the SA roadblock ($\chi^2(3) = 11.61$, $p = .009$). Namely, teams were significantly more likely to overcome the SA roadblock in Mission 2 ($M = 0.94$, $SE = 0.04$) than in Mission 3 ($M = 0.41$, $SE = 0.13$). The marginal model $R^2 = .57$. There were no significant two-way interaction effects.

4.3. Perceived situation awareness

The following analyses examine perceived SA across the missions and the two transition phases.

4.3.1. Perceived situation awareness across missions

A 3 (AI Information-Sharing Attribute: ATM, IET, Control) x 2 (AI Transition Phase Ordering: PN, NP) x 4 (Mission: M1, M2, M3, M4) linear mixed effects model was conducted to assess the effect of AI information-sharing attribute (between-subjects), AI transition phase ordering (between-subjects), and mission (within-subjects) on participants' level of perceived SA. A main effect of AI information-sharing attribute was found on participants' perceived SA ($\chi^2(2) = 7.68$, $p = .021$; see Figure 7). Participants' perceived SA across the missions was significantly greater for the ATM condition ($M = 5.41$, $SE = 0.16$) than the IET condition ($M = 4.83$, $SE = 0.16$). There was also a significant main effect of mission on participants' perceived SA across missions ($\chi^2(1) = 99.43$, $p < .001$; see Figure 7). Participants perceived SA was lower in Mission 1 ($M = 4.80$, $SE = 0.10$) than Mission 2 ($M = 5.16$, $SE = 0.10$), Mission 3 ($M = 5.26$, $SE = 0.10$), and Mission 4 ($M = 5.39$, $SE = 0.10$). Further, the perceived SA for Mission 2 was significantly lower {than that} of Mission 4. The marginal model $R^2 = .16$. There were no significant two-way interaction effects.

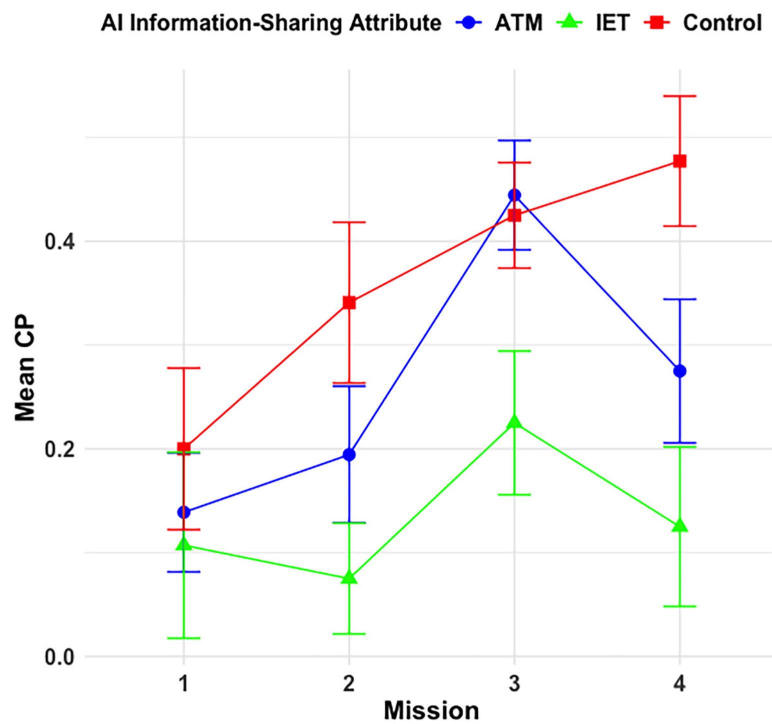

Figure 4. Main effects of mission and AI information-sharing attribute on human teammates' CP utterances in response to a system failure acting as an SA roadblock. Error bars indicate standard error.

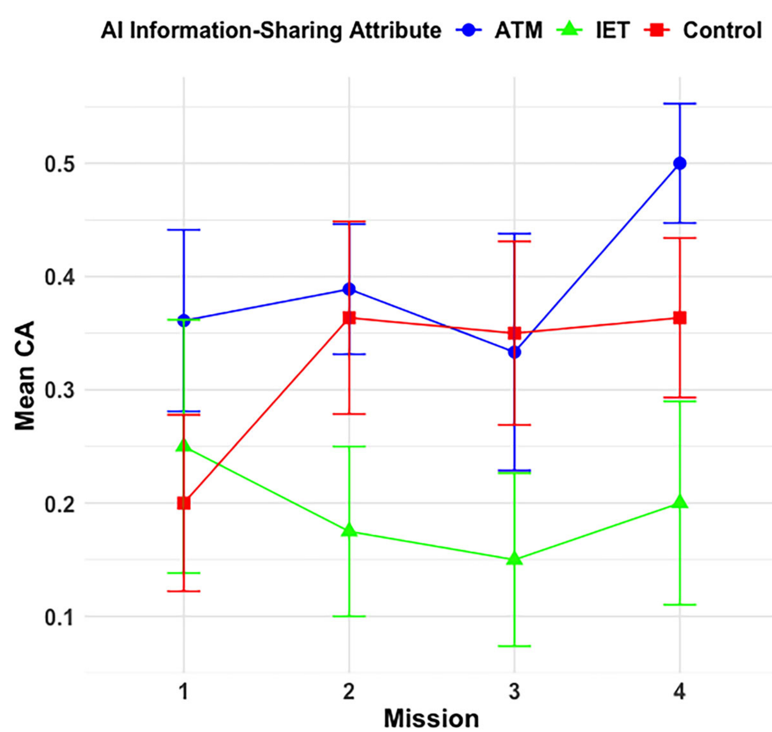

Figure 5. Main effects of mission and AI information-sharing attribute on human teammates' CA utterances in response to a system failure acting as an SA roadblock. Error bars indicate standard error.

4.3.2. Perceived situation awareness across transition phases

A 3 (AI Information-Sharing Attribute: ATM, IET, Control) x 2 (AI Transition Phase Ordering: PN, NP) x 4 (Time: T1, T2) linear mixed effects model was conducted to assess the effect of AI information-sharing attribute (between-subjects), AI transition phase ordering (between-subjects), and

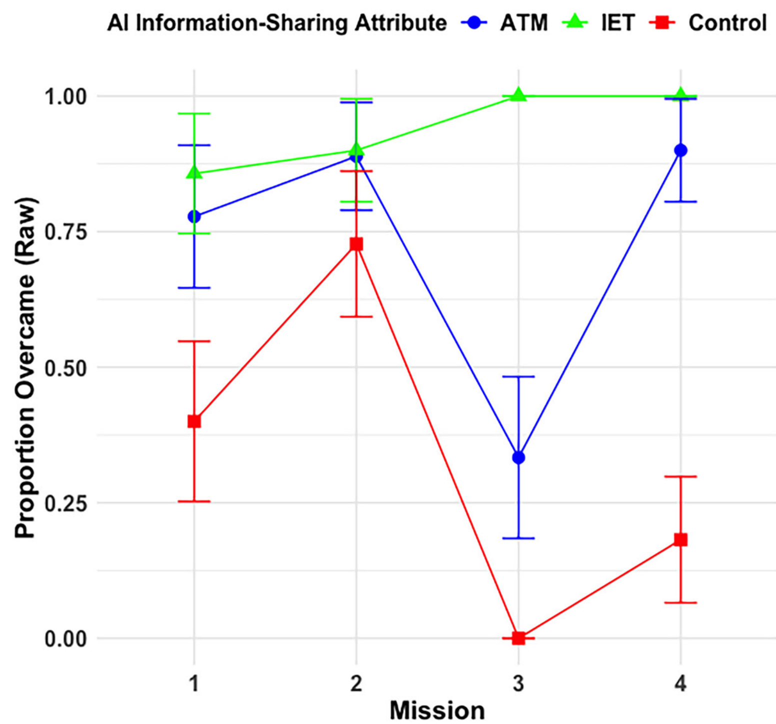

Figure 6. Main effects of mission and AI information-sharing attribute on the proportion of teams' overcoming the SA roadblock. Error bars represent standard error.

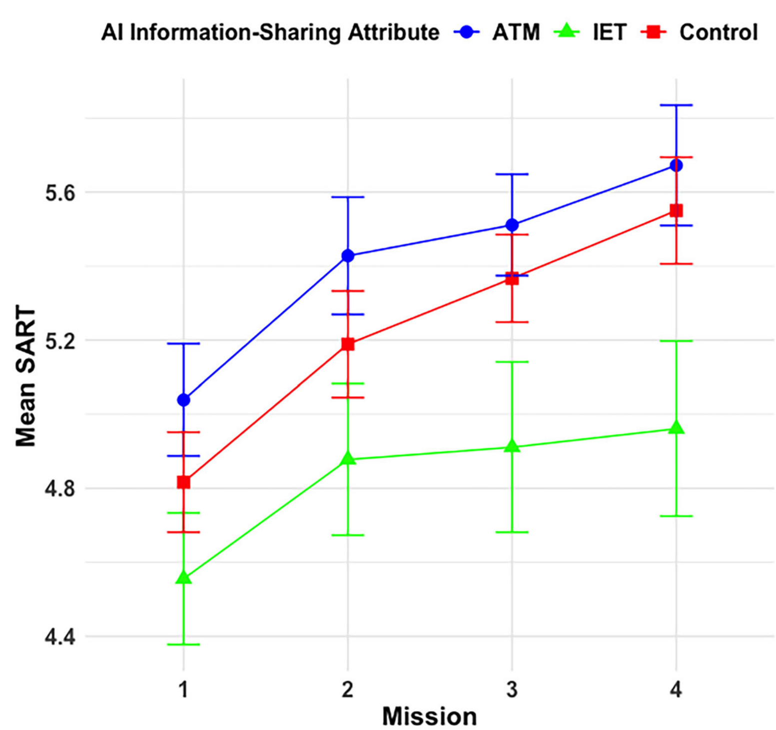

Figure 7. Main effects of AI information-sharing attribute and mission on participants' perceived SA across missions. Error bars represent standard error.

time (within-subjects) on participants' level of perceived SA across the two transition phases. A main effect of time was found on participants' perceived SA across the two transition phases ($\chi^2(1) = 34.89$, $p < .001$; see Figure 8). Participants' perceived SA in the first transition phase ($M = 4.87$, $SE = 0.10$)

was significantly lower compared to the second transition phase ($M = 5.24$, $SE = 0.10$). The marginal model $R^2 = .12$. There were no significant two-way interaction effects.

4.4. Perceived shared mental model with the AI teammate

The following analyses examine participants' perceived SMM with their AI teammate across the missions and the two transition phases.

4.4.1. Perceived shared mental model with the AI teammate across missions

A 3 (AI Information-Sharing Attribute: ATM, IET, Control) x 2 (AI Transition Phase Ordering: PN, NP) x 4 (Mission: M1, M2, M3, M4) linear mixed effects model was conducted to assess the effect of AI information-sharing attribute (between-subjects), AI transition phase ordering (between-subjects), and mission (within-subjects) on participants' perceived SMM with the AI teammate. A significant main effect of mission on participants' perceived SMM with the AI was found ($\chi^2(3) = 8.30$, $p = .04$; see Figure 9). Specifically, participants perceived SMM with the AI was significantly lower in Mission 1 ($M = 5.26$, $SE = .16$) compared to Mission 4 ($M = 5.62$, $SE = .16$). The marginal model $R^2 = .04$. However, this main effect was qualified by a significant ordinal interaction effect between mission and AI information-sharing attribute on participants perceived SMM with the AI $\chi^2(6) = 19.08$, $p = .004$; see Figure 9). Participants in the ATM condition ($M = 6.13$, $SE = .27$) rated their shared understanding with the AI teammate significantly higher than those in the control condition ($M = 5.07$, $SE = .27$) but only for Mission 4. The marginal model $R^2 = .10$.

4.4.2. Perceived shared mental model with the AI teammate across transition phases

A 3 (AI Information-Sharing Attribute: ATM, IET, Control) x 2 (AI Transition Phase Ordering: PN, NP) x 4 (Time: T1, T2) linear mixed effects model was conducted to assess the effect of AI information-sharing attribute (between-subjects), AI transition phase ordering (between-subjects), and time (within-subjects) on participants' perceived SMM with the AI Teammate across the two transition phases. There were no significant main effects, and the marginal model $R^2 = .05$.

However, there was a significant interaction effect between AI information-sharing attribute and time on participants' perceived SMM with the AI teammate ($\chi^2(2) = 6.47$, $p = .04$; see Figure 10).

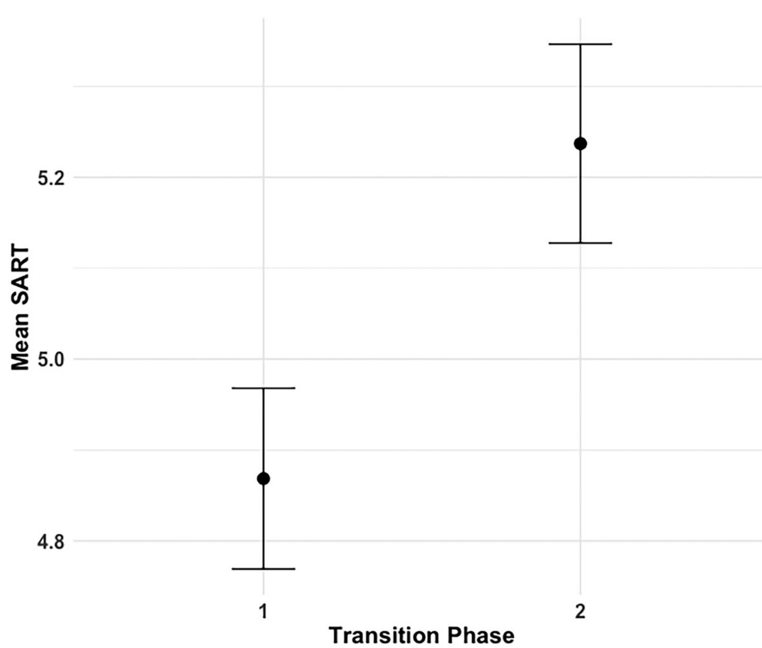

Figure 8. Main effect of time on participants' perceived SA across transition phases. Error bars represent standard error.

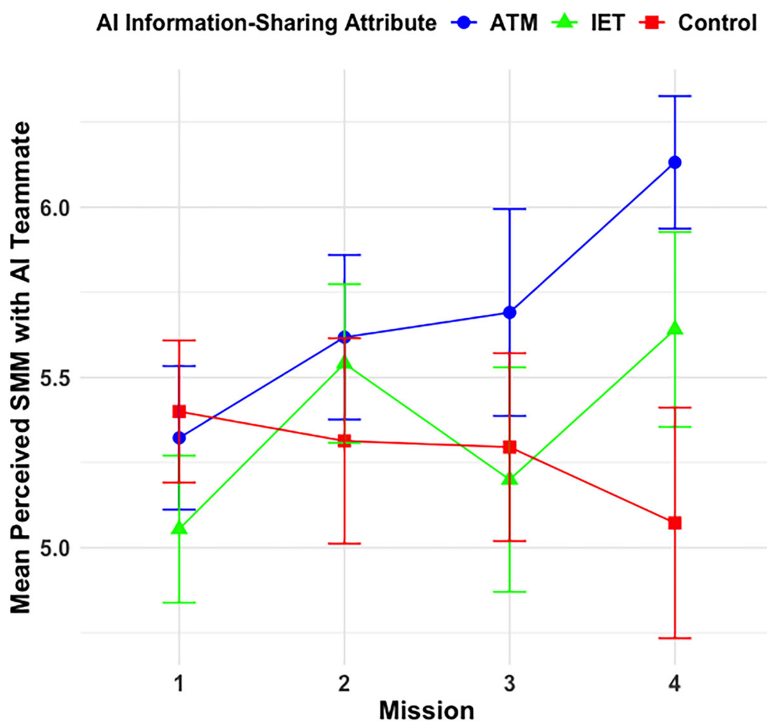

Figure 9. Interaction effect between mission and AI information-sharing attribute on participants' perceived SMM with the AI across missions. Error bars represent standard error.

Specifically, the participants in the ATM condition's perceived SMM with their AI teammate ($M = 6.00$, $SE = .25$) was greater than those in the control condition ($M = 5.11$, $SE = .25$), but only for the second transition phase. The marginal model $R^2 = .10$.

4.5. Team performance

A 3 (AI Information-Sharing Attribute: ATM, IET, Control) x 2 (AI Transition Phase Ordering: PN, NP) x 4 (Mission: M1, M2, M3, M4) linear mixed effects model was conducted to assess the effect of AI information-sharing attribute (between-subjects), AI transition phase ordering (between-subjects), and time point (within-subjects) on team performance. A main effect of mission was found on team performance ($\chi^2(3) = 9.99$, $p = .019$; see Figure 11). Specifically, team performance was higher in Mission 2 ($M = 329$, $SE = 13.6$) than in Mission 1 ($M = 292$, $SE = 13.6$). There was also a main effect of AI transition phase ordering on team performance ($\chi^2(1) = 4.39$, $p = .036$; see 11). Teams that saw the AI participate in the second transition phase ($M = 332$, $SE = 14.5$) had higher scores than those with the AI participating in the first transition phase (M = 291, SE = 15.00). The marginal model $R^2 =$ .13. There were no significant two-way interaction effects.

5. Discussion

The current study combines several AI design features to contribute to SA. These AI information-sharing attributes were tested using an experimental platform designed for complex team interactions over an extended period through several missions with real-world challenges, such as system failures. For RQ1, the results revealed that the ATM AI teammate encouraged significantly more perceived SA and SMM with the AI teammate. Specifically, the ATM AI teammate outperformed the control condition in perceived SMM with the AI in the last transition phase and mission, while the ATM condition outperformed the IET condition in perceived SA across missions. However, when examining RQ2 and the SA information-sharing attributes' effect on teams' response to task disruptions, the ATM condition had more CA utterances than the IET condition, and the control condition had more CP utterances

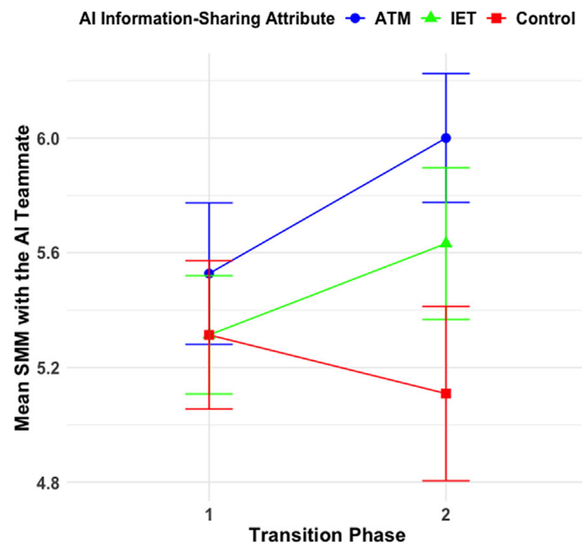

Figure 10. Interaction effect between time and AI information-sharing attribute on participants' perceived SMM with the AI across transition phases. Error bars represent standard error.

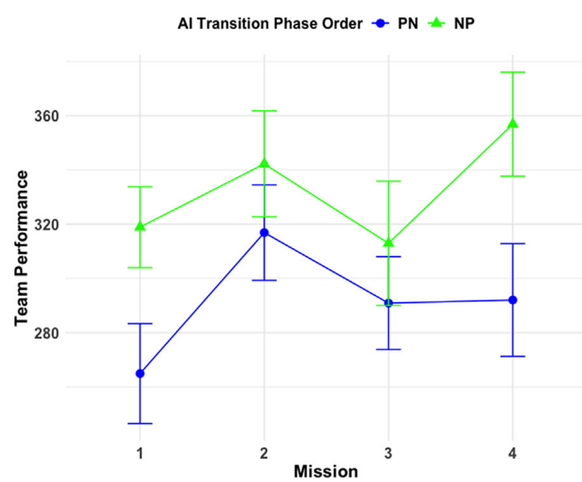

Figure 11. Main effects of mission and AI transition phase ordering on team performance. Error bars represent standard error.

than the IET condition. Still, despite their lower CA and CP utterances, the IET condition was most likely to overcome the task disruption, with the ATM condition also outperforming the control condition. Lastly, only a single result addressed RQ3, which found higher team performance for those teams working with the AI teammate that participated in the second transition phase rather than the first. The following discussion critically examines these findings in light of existing research.

5.1. Designing AI to contribute to shared knowledge is a balancing act between shared knowledge development and resiliency

The impact of the AI information-sharing attributes demonstrated clear improvements in overcoming the roadblock, but affected SA-related communication and perceived measures of shared knowledge differently. Specifically, the IET teams were significantly more likely to overcome the task disruption while having fewer CA and CP utterances than the ATM and control conditions, respectively. These results showcase HATs overcoming a significant task disruption with minimal communication to coordinate perception and action in response to the roadblock. This result for the IET condition mimics the behavior of high-performance teams that have strong, accurate SMMs and team SA, which can overcome obstacles with minimal and efficient communication (Endsley, 1995, 2023; Stout et al., 1999). However, it is doubtful that the teams working with the IET AI teammate truly developed the ability to overcome the task disruptions with minimal communication if the IET AI support was absent. It is more likely that the IET AI teammate allowed them to mimic the expert level of shared knowledge. While SMM development wasn't directly analyzed, the study did measure perceived SMM similarity with the AI teammate, which showed that the IET condition did not outperform the control condition. Alternatively, the ATM AI teams *did* have significantly higher perceived SMM similarity with the AI teammate compared to the control in the final mission and transition phase, suggesting the ATM AI contributed more to shared knowledge in this regard. These results on perceived SMM with the AI and SA are also distinct from the task disruption data (CAST).

Specifically, the ATM AI's non-roadblock information encouraged the participants to use the information to develop their own taskwork and teamwork patterns, thereby improving shared knowledge naturally compared to the direct solutions offered by the IET AI. Based on the current study's results, it is evident that having AI make meaningful contributions to SA involves more than just AI simply giving humans the solution. A tenet of the lumberjack effect is that the higher the level of automation for a task, the worse the failure will be when the system inevitably fails (Onnasch et al., 2014; Parasuraman & Riley, 1997; Sheridan & Parasuraman, 2000). While autonomy failures were not examined by the current study specifically, they did involve the CERTT system's failure, which the AI teammate is attempting to help mitigate. When trying to overcome the SA roadblock and mitigate the system failure, the human teammates benefited from what can be construed as a lower level of autonomy in the form of the ATM condition, which helped the human teammates find the information they needed to overcome the disruption on their own volition. Meanwhile, the IET condition gave teams the answer, improving their ability to overcome the task disruption, but hampered their perception of shared knowledge development within the team as a consequence.

Although not directly studied, the ATM AI may have naturally (could do so without the AI teammate) enhanced teams' resiliency to future related system failures by teaching teammates essential information, when and how to use it, leading to better coordination and decision-making because they made human teammates find and share the correct information themselves instead of directly providing it (Yano et al., 2015). A solid push exists to design AI teammates to take things away from their human teammates, such as workload, risk, and other task-related factors (Shneiderman, 2020). However, these findings may suggest that there is also a place for designing AI capable of improving the way human teammates work with one another and not just the AI itself. The ATM conditions' results relating to task disruptions significantly increased CA utterances supporting team SA in response to those roadblocks. There is a clear trade-off between the IET AI's focus on performance consistency and overcoming SA roadblocks with minimal communication (e.g., direct answers to roadblocks or alerts about altitude and airspeed) compared to the ATM AI's approach emphasizing support for improved shared understanding and planning ahead (e.g., guiding the team to solve roadblocks or providing airspeed ranges for upcoming ROZs). Past research on HATs in the CERTT UAS-STE found that those working with AI teammates struggled to engage in effective pushing verbal behaviors (Demir et al., 2018; McNeese et al., 2018). While the current study is not a one-to-one comparison, the high resiliency of the IET teams, along with the high perceived shared knowledge of the ATM teams, would suggest that this shortcoming may be overcome with more effective AI teammate designs tailored to understand when to push answers or when to push teammates to answer it themselves.

Designing AI teammates to contribute to shared knowledge and overcome unexpected task disruptions should be a priority for the system and team designers, with the highest levels of performance likely coming from AI teammates capable of engaging in both information-sharing techniques adaptively based on the context of the environmental state. Applying these results outside of the CERTT UAS-STE's UAV context, the ability to leverage AI's inherent technical strengths, such as an expert-level understanding of the basic task and information requirements across roles, to quickly bring human teammates up to speed is critical. All team-based activities rely on this shared understanding supported by SMMs and team SA (Cooke et al., 2013; Endsley, 2023; Mathieu et al., 2000), and if AI teammates can utilize their pre-programmed knowledge to help detail the task in action and transition phases, then a host of benefits could be introduced. For example, AI could rapidly bring new team members joining an existing team up to speed, entirely new teams can undergo AI augmented training to reduce training times, or AI can contribute to continuous improvement in teams to improve performance and resiliency over time. With these results, the understanding of AI teammates' information-sharing impact on team development is improved, displaying how AI teammates might be better tailored to the unique needs of the environments in which they find themselves.

5.2. AI participation in transition phases and the role of shared experiences on shared knowledge for human-AI teams

Transition periods are critical to effective teaming in all forms, human-only and human-AI (Kennedy & McComb, 2014; Marks et al., 2001; Tannenbaum & Cerasoli, 2013). Knowing this, the current study examined how the presence of an AI teammate early or late in the team's life cycle affected the team's performance, SA, and SMM with the AI teammate. However, few results spoke to RQ3 and the influence of AI transition phase participation, except for those on team performance. Specifically, the HATs where the AI teammate participated in the second transition phase rather than the first had the highest level of performance, which was the only significant effect on performance in the study. Such results echo the research pointing toward the different types of discussions occurring earlier and later in teams' life cycles having varied effects on teams, potentially extending this finding to HATs for the first time (Gersick, 1988). Specifically, the first transition phase may have been beneficial for essential task learning, which is supported by the fact that the first few missions had the lowest performance and perceived shared knowledge for teams. Initial team discussions for human-only teams have noted this focus on basic task learning (Fransen et al., 2013; Tuckman & Jensen, 1977). However, AI teammate participation in the second transition phase positively affected the teams' performance, which may result from the human teammates needing to develop task and teamwork knowledge independently without the AI teammate in the first transition phase, encouraging stronger self-reliance and confidence. Still, these results are unclear, and additional research is needed to fully understand the role and effect of an AI teammate's participation in transition phase discussions. What is clear is that AI teammates' involvement in these discussions does affect HAT functioning and warrants further study.

HATs and HAT studies benefit from engaging in transition phases, and their effect on shared knowledge development over time is apparent. As teams progress through a task, they develop their skills, shared knowledge, and norms, resulting in markedly better performance (Marks et al., 2001; Salas et al., 1992). A similar effect is seen in the current study for the HATs analyzed, except these HATs were empowered to engage in additional discussion during transition phases compared to previous CERTT and HAT experiments (Hauptman et al., 2023; McNeese et al., 2018). Allowing HATs to engage in these transition phases resulted in significantly improved perceived SA from the first transition phase to the second for all teams and a stronger perceived SMM with the AI teammate for the ATM condition in the second transition phase, which was also present for the mission analysis in the final mission. These results highlight the temporal nature of shared knowledge development and give teams as much time together as possible to develop that shared knowledge, especially as the impact of shared knowledge on team performance increases over time Mohammed et al. (2015). HATs mirror human-only teams' dependence on team cognition to have highly effective transition phases (Bush et al., 2018), and these transition phases are a direct avenue for teams' emergent states to act as inputs on future team effectiveness. Studying HATs as they engage in transition phases also increases the ecological validity of studies, given how common such discussions are in working teams (Tannenbaum & Cerasoli, 2013). The importance of AI participation in transition phases extends beyond the CERTT UAV environment, with team briefings and debriefings a critical part of team effectiveness across industries (Tannenbaum & Cerasoli, 2013). Participation or the lack thereof by AI teammates in team transition phases is very likely to have an effect in other highly interdependent team-based tasks similar to the CERTT UAS-STE in performance-based and affective-based team outcomes. Consequently, as the role of HATs increases, the importance of these transition periods will increase. Thus, the criticality of these shared knowledge constructs of team cognition will be further emphasized for positive HAT outcomes.

5.3. Limitations and future work

The results of this study have significant benefits to the state of HAT, though it has limitations that should be considered when interpreting the results. Specifically, the generalizability of the task disruption, as it is engineered to represent a system failure, cannot be applied to all unexpected events within a task. Though this is a practice utilized in past CERTT UAS-STE experiments (Cooke et al., 2016; Gorman et al., 2005; McNeese et al., 2018), future studies should examine the range of applicability these findings have on unexpected events. Another limitation of the study is the emphasis on the task disruption to assess the types of information-sharing, as only team performance, perceived SA, and the exploratory perceived SMM with the AI teammate were directly affected by the other information-sharing done by the AI teammate. Future studies could examine team responses to the task disruptions as a separate condition to examine both effects. Given that the primary difference between the two information-sharing attributes was the presence or lack of direct support, which found benefits and drawbacks, the question of adaptive autonomy could leverage the best of both. As such, future research should study whether adaptive autonomy could be introduced to intelligently provide direct solutions to teammates depending on the criticality of the failure or problem the team is facing. Such an implementation would allow teams to benefit from improved level 3 SA and shared mental model development, with the additional benefit of quickly addressing critical failures that otherwise risk team performance or safety. Finally, the current study involved college students with varying individual differences. As such, several questions remain on the impact of those differences on human teammates' willingness to accept information from and interact collaboratively with their AI teammate.

6. Conclusion

This study advances our understanding of HATs and how AI teammates influence shared knowledge development, particularly team SA in response to task disruptions, across different AI information-sharing techniques and AI participation in transition phases early versus late in the team's life cycle. The study found that while the IET AI information-sharing attribute, which was designed to give human teammates direct information and solutions on environmental changes, gave HATs the best chance to overcome the task disruption, it also reduced coordinated perception and action communication initiated by human teammates, and, subsequently, did not foster the highest levels of perceived SA and SMM with the AI teammate compared to the ATM and control conditions in some instances. It was the ATM AI information-sharing attribute, which was designed to augment the teams' memory with information they could use to plan and decide their actions, that gave teams the next best chance to overcome the task disruption compared to the control. The ATM AI did this while also fostering higher levels of perceived SMM with the AI and SA compared to the control and IET conditions in some cases. Lastly, the teams that were able to speak with their AI teammate in the second transition phase, towards the end of their life cycle, outperformed those that spoke with it in the first transition phase at the start of their teams' life cycle. These results demonstrate the importance of providing a balance to the information provided by AI teammates as they engage in more shared knowledge supporting teamwork behaviors, such as information-sharing, which contribute directly to team cognition. As such, the current study enhances our understanding of the nuanced development of SA in HATs, making it clear that AI designs are not one-size-fits-all; their effects can be positive or negative depending on the team's contextual challenges at any given moment. Researchers and practitioners can use these insights to design AI teammates that expand the potential for team SA, making HATs more resilient and effective.

Author contributions

CRediT: **Beau G. Schelble**: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing; **Rohit Mallick**: Conceptualization, Formal analysis, Investigation, Methodology, Project administration, Software, Writing – original draft, Writing – review & editing; **Allyson Hauptman**: Conceptualization, Writing – original draft, Writing – review & editing; **Nathan McNeese**: Conceptualization, Methodology, Project administration, Resources, Software, Supervision, Writing – review & editing.

Disclosure statement

No potential conflict of interest was reported by the author(s).

ORCID

Beau G. Schelble http://orcid.org/0000-0003-3704-697X
Rohit Mallick http://orcid.org/0009-0007-8411-2470
Allyson Hauptman http://orcid.org/0000-0001-7785-5996
Nathan McNeese http://orcid.org/0000-0002-9143-2460

References

Baaj, M. H., & Mahmassani, H. S. (1991). An ai-based approach for transit route system planning and design. *Journal of Advanced Transportation*, *25*(2), 187–209. https://doi.org/10.1002/atr.5670250205

Bolstad, C. A., Cuevas, H. M., & Costello, A. M. (2005). *Improving situation awareness through cross-training* [Paper presentation]. Proceedings of the Human Factors and Ergonomics Society Annual Meeting (Vol. 49, pp. 2159–2163). https://doi.org/10.1177/154193120504902509

Bush, J. T., LePine, J. A., & Newton, D. W. (2018). Teams in transition: An integrative review and synthesis of research on team task transitions and propositions for future research. *Human Resource Management Review*, *28*(4), 423–433. https://doi.org/10.1016/j.hrmr.2017.06.005

Cannon-Bowers, J. A., Salas, E., & Converse, S. (1993). Shared mental models in expert team decision making. In *Current issues in individual and group decision making* (pp. 221–247). Lawrence Earlbaum.

Chen, J. Y., 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

Chiou, E. K., Demir, M., Buchanan, V., Corral, C. C., Endsley, M. R., Lematta, G. J., Cooke, N. J., & McNeese, N. J. (2022). Towards human–robot teaming: Tradeoffs of explanation-based communication strategies in a virtual search and rescue task. *International Journal of Social Robotics*, *14*(8), 1933–1933. https://doi.org/10.1007/s12369-022-00886-x

Cooke, N. J., Cohen, M. C., Fazio, W. C., Inderberg, L. H., Johnson, C. J., Lematta, G. J., Peel, M., & Teo, A. (2024). From teams to teamness: Future directions in the science of team cognition. *Human Factors*, *66*(6), 1669–1680. https://doi.org/10.1177/00187208231162449

Cooke, N. J., Demir, M., & McNeese, N. (2016). *Synthetic teammates as team players: Coordination of human and synthetic teammates* (Technical Report). Cognitive Engineering Research Institute Mesa United States. https://apps.dtic.mil/sti/html/tr/AD1017169/

Cooke, N. J., Gorman, J. C., Duran, J. L., & Taylor, A. R. (2007). Team cognition in experienced command-and-control teams. *Journal of Experimental Psychology. Applied*, *13*(3), 146–157. https://doi.org/10.1037/1076-898X.13.3.146

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/10.1111/cogs.12009

Cooke, N. J., & Shope, S. M. (2004). Synthetic task environments for teams: CERTT's UAV-STE. In N. Stanton, A. Hedge, K. Brookhuis, E. Salas, & H. Hendrick (Eds.), *Handbook of human factors and ergonomics methods* (pp. 41–46). CRC Press.

Cuevas, H. M., Fiore, S. M., Caldwell, B. S., & Strater, L. (2007). Augmenting team cognition in human-automation teams performing in complex operational environments. *Aviation, Space, and Environmental Medicine*, *78*(5 Suppl), B63–B70.

Damacharla, P., Dhakal, P., Stumbo, S., Javaid, A. Y., Ganapathy, S., Malek, D. A., Hodge, D. C., & Devabhaktuni, V. (2019). Effects of voice-based synthetic assistant on performance of emergency care provider in training. *International Journal of Artificial Intelligence in Education*, *29*(1), 122–143. https://doi.org/10.1007/s40593-018-0166-3

Demir, M., & Cooke, N. J. (2014). *Human teaming changes driven by expectations of a synthetic teammate* [Paper presentation]. Proceedings of the Human Factors and Ergonomics Society Annual Meeting (Vol. 58, pp. 16–20). SAGE. https://doi.org/10.1177/1541931214581004

Demir, M., Likens, A. D., Cooke, N. J., Amazeen, P. G., & McNeese, N. (2019). Team coordination and effectiveness in human-autonomy teaming. *IEEE Transactions on Human-Machine Systems*, *49*(2), 150–159. https://doi.org/10.1109/THMS.2018.2877482

Demir, M., McNeese, N., & Cooke, N. J. (2017). Team situation awareness within the context of human-autonomy teaming. *Cognitive Systems Research*, *46*(2017), 3–12. https://doi.org/10.1016/j.cogsys.2016.11.003

Demir, M., McNeese, N., & Cooke, N. J. (2018). The impact of perceived autonomous agents on dynamic team behaviors. *IEEE Transactions on Emerging Topics in Computational Intelligence*, *2*(4), 258–267. https://doi.org/10.1109/TETCI.2018.2829985

Demir, M., McNeese, N., & Cooke, N. J. (2020). Understanding human-robot teams in light of all-human teams: Aspects of team interaction and shared cognition. *International Journal of Human-Computer Studies*, *140*(2020), 102436. https://doi.org/10.1016/j.ijhcs.2020.102436

Demir, M., McNeese, N., Cooke, N. J., Ball, J. T., Myers, C., & Frieman, M. (2015). *Synthetic teammate communication and coordination with humans* [Paper presentation]. Proceedings of the Human Factors and Ergonomics Society Annual Meeting (Vol. 59, pp. 951–955). https://doi.org/10.1177/1541931215591275

Dickinson, T. L., & McIntyre, R. M. (1997). A conceptual framework for teamwork measurement. In M. T. Brannick, E. Salas, & C. Prince (Eds.), *Team performance assessment and measurement: Theory, methods, and applications* (pp. 19–43). Lawrence Erlbaum Associates.

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

Endsley, M. (2015). Situation awareness misconceptions and misunderstandings. *Journal of Cognitive Engineering and Decision Making*, *9*(1), 4–32. https://doi.org/10.1177/1555343415572631

Endsley, M. (2017). From here to autonomy: Lessons learned from human–automation research. *Human Factors*, *59*(1), 5–27. https://doi.org/10.1177/0018720816681350

Endsley, M. (2023). Supporting human-ai teams: Transparency, explainability, and situation awareness. *Computers in Human Behavior*, *140*(2023), 107574. https://doi.org/10.1016/j.chb.2022.107574

Endsley, M., Bolte̕, B., & Jones, D. G. (2003). *Designing for situation awareness: An approach to user-centered design* (1st ed.). CRC press. https://doi.org/10.1201/9780203485088

Endsley, M., & Jones, W. M. (1997). *Situation awareness information dominance & information warfare* (Technical Report). LOGICON TECHNICAL SERVICES INC DAYTON OH. https://apps.dtic.mil/sti/pdfs/ADA347166.pdf?trk=public_post_comment-text

Ezenyilimba, A., Wong, M., Hehr, A., Demir, M., Wolff, A., Chiou, E., & Cooke, N. (2023). Impact of transparency and explanations on trust and situation awareness in human–robot teams. *Journal of Cognitive Engineering and Decision Making*, *17*(1), 75–93. https://doi.org/10.1177/15553434221136358

Fiore, S. M., Cuevas, H. M., Schooler, J. W., & Salas, E. (2006). Cognition, teams, and team cognition: memory actions and memory failures in distributed team environments. In C. Bowers, E. Salas, & F. Jentsch (Eds.), *Creating high-tech teams: Practical guidance on work performance and technology* (pp. 71–87). American Psychological Association. https://doi.org/10.1037/11263-004

Fleishman, E. A., & Zaccaro, S. J. (1992). Toward a taxonomy of team performance functions. In R. W. Swezey & E. Salas (Eds.), *Teams: Their training and performance* (pp. 31–56). Ablex Publishing. https://psycnet.apa.org/record/1992-98450-002

Fox, J. (2015). *Applied regression analysis and generalized linear models*. Sage Publications.

Fransen, J., Weinberger, A., & Kirschner, P. A. (2013). Team effectiveness and team development in cscl. *Educational Psychologist*, *48*(1), 9–24. https://doi.org/10.1080/00461520.2012.747947

Gergle, D., Kraut, R. E., & Fussell, S. R. (2013). Using visual information for grounding and awareness in collaborative tasks. *Human–Computer Interaction*, *28*(1), 1–39. https://doi.org/10.1080/07370024.2012.678246

Gersick, C. J. (1988). Time and transition in work teams: Toward a new model of group development. *Academy of Management Journal*, *31*(1), 9–41. https://doi.org/10.2307/256496

Gladstein, D. L. (1984). Groups in context: A model of task group effectiveness. *Administrative Science Quarterly*, *29*(4), 499–517. https://doi.org/10.2307/2392936

Gorman, J. C., Cooke, N. J., & Amazeen, P. G. (2010). Training adaptive teams. *Human Factors*, *52*(2), 295–307. https://doi.org/10.1177/0018720810371689

Gorman, J. C., Cooke, N. J., Pederson, H. K., & DeJoode, J. A. (2005). *Coordinated awareness of situation by teams (cast): Measuring team situation awareness of a communication glitch* [Paper presentation]. Proceedings of the Human Factors and Ergonomics Society Annual Meeting (Vol. 49, pp. 274–277). https://doi.org/10.1177/154193120504900313

Gorman, J. C., Grimm, D. A., Stevens, R. H., Galloway, T., Willemsen-Dunlap, A. M., & Halpin, D. J. (2020). Measuring real-time team cognition during team training. *Human Factors*, *62*(5), 825–860. https://doi.org/10.1177/0018720819852791

Hauptman, A. I., Schelble, B. G., McNeese, N., & Madathil, K. C. (2023). Adapt and overcome: Perceptions of adaptive autonomous agents for human-ai teaming. *Computers in Human Behavior*, *138*(2023), 107451. https://doi.org/10.1016/j.chb.2022.107451

Hu, W.-C., Wu, H.-T., Cho, H.-H., & Tseng, F.-H. (2020). Optimal route planning system for logistics vehicles based on artificial intelligence. *Journal of Internet Technology*, *21*(3), 757–764. https://doi.org/10.3966/160792642020052103013

Jentsch, F., Barnett, J., Bowers, C. A., & Salas, E. (1999). Who is flying this plane anyway? what mishaps tell us about crew member role assignment and air crew situation awareness. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *41*(1), 1–14. https://doi.org/10.1518/001872099779577237

Jiang, J., Karran, A. J., Coursaris, C. K., Léger, P.-M., & Beringer, J. (2023). A situation awareness perspective on human-ai interaction: Tensions and opportunities. *International Journal of Human–Computer Interaction*, *39*(9), 1789–1806. https://doi.org/10.1080/10447318.2022.2093863

Jones, D. G., & Endsley, M. (1996). Sources of situation awareness errors in aviation. *Aviation, Space, and Environmental Medicine*, *67*(6), 507–512.

Kennedy, D. M., & McComb, S. A. (2014). When teams shift among processes: Insights from simulation and optimization. *The Journal of Applied Psychology*, *99*(5), 784–815. https://doi.org/10.1037/a0037339

Koch, C. (2016). How the computer beat the go player. *Scientific American Mind*, *27*(4), 20–23. https://doi.org/10.1038/scientificamericanmind0716-20

Kude, T., Mithas, S., Schmidt, C. T., & Heinzl, A. (2019). How pair programming influences team performance: The role of backup behavior, shared mental models, and task novelty. *Information Systems Research*, *30*(4), 1145–1163. https://doi.org/10.1287/isre.2019.0856

Liu, J., Tang, X., Li, L., Chen, P., & Liu, Y. (2023). Which is a better programming assistant? a comparative study between chatgpt and stack overflow. *arXiv preprint arXiv:2308.13851*. https://doi.org/10.48550/arXiv.2308.13851

Loske, D., & Klumpp, M. (2021). Human-ai collaboration in route planning: An empirical efficiency-based analysis in retail logistics. *International Journal of Production Economics*, *241*(2021), 108236. https://doi.org/10.1016/j.ijpe.2021.108236

Lyons, J. B., Sycara, K., Lewis, M., & Capiola, A. (2021). Human–autonomy teaming: Definitions, debates, and directions. *Frontiers in Psychology*, *12*(2021), 589585. https://doi.org/10.3389/fpsyg.2021.589585

Mallick, R., Flathmann, C., Duan, W., Schelble, B. G., & McNeese, N. (2024). What you say vs what you do: Utilizing positive emotional expressions to relay ai teammate intent within human–ai teams. *International Journal of Human-Computer Studies*, *192*(2024), 103355. https://doi.org/10.1016/j.ijhcs.2024.103355

Marks, M. A., Mathieu, J. E., & Zaccaro, S. J. (2001). A temporally based framework and taxonomy of team processes. *The Academy of Management Review*, *26*(3), 356–376. https://doi.org/10.2307/259182

Mathieu, J. E., Heffner, T. S., Goodwin, G. F., Salas, E., & Cannon-Bowers, J. A. (2000). The influence of shared mental models on team process and performance. *The Journal of Applied Psychology*, *85*(2), 273–283. https://doi.org/10.1037/0021-9010.85.2.273

Maulsby, D., Greenberg, S., & Mander, R. (1993). *Prototyping an intelligent agent through wizard of oz* [Paper presentation]. Proceedings of the interact'93 and chi'93 Conference on Human Factors in Computing Systems (pp. 277–284). https://doi.org/10.1145/169059.169215

McGraw, K. O., & Wong, S. P. (1996). Forming inferences about some intraclass correlation coefficients. *Psychological Methods*, *1*(1), 30–46. https://doi.org/10.1037/1082-989X.1.1.30

McNeese, N., Demir, M., Cooke, N. J., & Myers, C. (2018). Teaming with a synthetic teammate: Insights into human-autonomy teaming. *Human Factors*, *60*(2), 262–273. https://doi.org/10.1177/0018720817743223

McNeese, N., Demir, M., Cooke, N. J., & She, M. (2021). Team situation awareness and conflict: A study of human–machine teaming. *Journal of Cognitive Engineering and Decision Making*, *15*(2-3), 83–96. https://doi.org/10.1177/15553434211017354

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

Mohammed, S., Hamilton, K., Tesler, R., Mancuso, V., & McNeese, M. (2015). Time for temporal team mental models: Expanding beyond "what" and "how" to incorporate "when". *European Journal of Work and Organizational Psychology*, *24*(5), 693–709. https://doi.org/10.1080/1359432X.2015.1024664

Munir, A., Aved, A., & Blasch, E. (2022). Situational awareness: Techniques, challenges, and prospects. *AI*, *3*(1), 55–77. https://doi.org/10.3390/ai3010005

Nakagawa, S., & Schielzeth, H. (2013). A general and simple method for obtaining r2 from generalized linear mixed-effects models. *Methods in Ecology and Evolution*, *4*(2), 133–142. https://doi.org/10.1111/j.2041-210x.2012.00261.x

National Academies of Sciences, Engineering & Medicine (2021). *Human-AI teaming: State of the art and research needs*. The National Academies Press. https://doi.org/10.17226/26355

Nyre-Yu, M., Gutzwiller, R. S., & Caldwell, B. S. (2019). *Observing cyber security incident response: Qualitative themes from field research* [Paper presentation]. Proceedings of the Human Factors and Ergonomics Society Annual Meeting (Vol. 63, pp. 437–441). https://doi.org/10.1177/1071181319631016

O'Neill, T., McNeese, N., Barron, A., & Schelble, B. (2022). Human–autonomy teaming: A review and analysis of the empirical literature. *Human Factors*, *64*(5), 904–938. https://doi.org/10.1177/0018720820960865

Onnasch, L., Wickens, C. D., Li, H., & Manzey, D. (2014). Human performance consequences of stages and levels of automation: An integrated meta-analysis. *Human Factors*, *56*(3), 476–488. https://doi.org/10.1177/0018720813501549

Paleja, R., Ghuy, M., Ranawaka Arachchige, N., Jensen, R., & Gombolay, M. (2021). The utility of explainable ai in ad hoc human-machine teaming. *Advances in Neural Information Processing Systems*, *34*, 610–623. https://proceedings.neurips.cc/paper_files/paper/2021/file/05d74c48b5b30514d8e9bd60320fc8f6-Paper.pdf

Parasuraman, R., & Riley, V. (1997). Humans and automation: Use, misuse, disuse, abuse. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *39*(2), 230–253. https://doi.org/10.1518/001872097778543886

Prince, C., & Salas, E. (1993). Training and research for teamwork in the military aircrew. In E. L. Wiener, B. G. Kanki, & R. L. Helmreich (Eds.), *Cockpit resource management* (pp. 337–366). Academic Press. https://psycnet.apa.org/record/1993-97500-012

Prussia, G. E., & Kinicki, A. J. (1996). A motivational investigation of group effectiveness using social-cognitive theory. *Journal of Applied Psychology*, *81*(2), 187–198. https://doi.org/10.1037/0021-9010.81.2.187

Rosopa, P. J., Schaffer, M. M., & Schroeder, A. N. (2013). Managing heteroscedasticity in general linear models. *Psychological Methods*, *18*(3), 335–351. https://doi.org/10.1037/a0032553

Salas, E., Dickinson, T. L., Converse, S. A., & Tannenbaum, S. I. (1992). Toward an understanding of team performance and training. In R. W. Swezey & E. Salas (Eds.), *Teams: Their training and performance* (pp. 3–29). Ablex Publishing. https://psycnet.apa.org/record/1992-98450-001

Salas, E., Prince, C., Baker, D. P., & Shrestha, L. (1995). Situation awareness in team performance: Implications for measurement and training. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *37*(1), 123–136. https://doi.org/10.1518/001872095779049525

Sarker, I. H., Janicke, H., Mohammad, N., Watters, P., & Nepal, S. (2023). *AI potentiality and awareness: A position paper from the perspective of human-ai teaming in cybersecurity* [Paper presentation]. International Conference on Intelligent Computing & Optimization (pp. 140–149). https://doi.org/10.1007/978-3-031-50887-5_14

Schaefer, K. E., Straub, E. R., Chen, J. Y., Putney, J., & Evans, A. W. III, (2017). Communicating intent to develop shared situation awareness and engender trust in human-agent teams. *Cognitive Systems Research*, *46*, 26–39. https://doi.org/10.1016/j.cogsys.2017.02.002

Schelble, B. G. (2023). *Leveraging artificial intelligence for team cognition in human-ai teams.* (Unpublished doctoral dissertation). Clemson University.

Schelble, B. G., Flathmann, C., Macdonald, J. P., Knijnenburg, B., Brady, C., & McNeese, N. (2025). Modeling perceived information needs in human-ai teams: Improving ai teammate utility and driving team cognition. *Behaviour & Information Technology*, *44*(9), 2069–2092. https://doi.org/10.1080/0144929X.2024.2396476

Schelble, B. G., Flathmann, C., & McNeese, N. (2020). Towards meaningfully integrating human-autonomy teaming in applied settings [Paper presentation]. Proceedings of the 8th International Conference on Human-Agent Interaction (pp. 149–156). https://doi.org/10.1145/3406499.3415077

Schelble, B. G., Flathmann, C., McNeese, N., Freeman, G., & Mallick, R. (2022a). Let's think together! assessing shared mental models, performance, and trust in human-agent teams. *Proceedings of the ACM on Human-Computer Interaction*, *6* (GROUP), 1–29. https://doi.org/10.1145/3492832

Schelble, B. G., Flathmann, C., Musick, G., McNeese, N. J., & Freeman, G. (2022b). I See You: Examining the Role of Spatial Information in Human-Agent Teams. *Proceedings of the ACM on Human-Computer Interaction*, *6*(CSCW2), 1–27. https://dl.acm.org/doi/10.1145/3555099

Schielzeth, H., Dingemanse, N. J., Nakagawa, S., Westneat, D. F., Allegue, H., Teplitsky, C., Réale, D., Dochtermann, N. A., Garamszegi, L. Z., & Araya-Ajoy, Y. G. (2020). Robustness of linear mixed-effects models to violations of distributional assumptions. *Methods in Ecology and Evolution*, *11*(9), 1141–1152. https://doi.org/10.1111/2041-210X.13434

Schulman, J., Wolski, F., Dhariwal, P., Radford, A., & Klimov, O. (2017). Proximal policy optimization algorithms. *arXiv preprint arXiv:1707.06347*. https://doi.org/10.48550/arXiv.1707.06347

Sheridan, T. B., & Parasuraman, R. (2000). *Human vs. automation in responding to failures: An expected-value analysis* [Paper presentation]. Proceedings of the Human Factors and Ergonomics Society Annual Meeting (Vol. 44, pp. 1–4). https://doi.org/10.1177/154193120004401301

Shneiderman, B. (2020). Human-centered artificial intelligence: Reliable, safe & trustworthy. *International Journal of Human–Computer Interaction*, *36*(6), 495–504. https://doi.org/10.1080/10447318.2020.1741118

Skaug, H. A., & Busch, P. A. (2022). *AI for situational awareness in situations with high uncertainty: An explorative case study.* ECIS 2022 Research Papers. 62. https://aisel.aisnet.org/ecis2022_rp/62?utm_source=aisel.aisnet.org%2Fecis2022_rp%2F62&utm_medium=PDF&utm_campaign=PDFCoverPages

Stephanidis, C., Salvendy, G., Antona, M., Duffy, V. G., Gao, Q., Karwowski, W., Konomi, S., Nah, F., Ntoa, S., Rau, P.-L. P., Siau, K., & Zhou, J. (2025). Seven hci grand challenges revisited: Five-year progress. *International Journal of Human–Computer Interaction*, 1–49. https://doi.org/10.1080/10447318.2025.2450411

Stout, R. J., Cannon-Bowers, J. A., Salas, E., & Milanovich, D. M. (1999). Planning, shared mental models, and coordinated performance: An empirical link is established. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *41*(1), 61–71. https://doi.org/10.1518/001872099779577273

Tannenbaum, S. I., & Cerasoli, C. P. (2013). Do team and individual debriefs enhance performance? a meta-analysis. *Human Factors*, *55*(1), 231–245. https://doi.org/10.1177/0018720812448394

Taylor, R. M. (1990). Situational Awareness Rating Technique (SART): The development of a tool for aircrew systems design. In *Situational awareness in aerospace operations* (AGARD-CP-478; pp. 3/1–3/17). NATO-AGARD.

Tuckman, B. W., & Jensen, M. A. C. (1977). Stages of small-group development revisited. *Group & Organization Studies*, *2*(4), 419–427. https://doi.org/10.1177/105960117700200404

Van Rensburg, J. J., Santos, C. M., de Jong, S. B., & Uitdewilligen, S. (2021). The five-factor perceived shared mental model scale: A consolidation of items across the contemporary literature. *Frontiers in Psychology*, *12*, 784200. https://doi.org/10.3389/fpsyg.2021.784200

Weger, K., Leder, S., Mesmer, B., Menon, V., & Schaub, H. (2022). How effectively do we communicate? an analysis of team reflexivity in transition and action phases of team collaboration. *IEEE Transactions on Professional Communication*, *65*(3), 392–410. https://doi.org/10.1109/TPC.2022.3186773

Wellens, A. R. (1993). Group situation awareness and distributed decision making: From military to civilian applications. In N. J. Castellan, Jr. (Ed.), *Individual and group decision making* (pp. 267–291). Erlbaum.

Wright, J. L., Chen, J. Y., Barnes, M. J., & Hancock, P. A. (2016). *The effect of agent reasoning transparency on automation bias: An analysis of response performance* [Paper presentation]. International Conference on Virtual, Augmented and Mixed Reality (pp. 465–477). https://doi.org/10.1007/978-3-319-39907-2_45

Yan, B., & Gurkan, N. (2023). *It depends on the timing: The ripple effect of ai on team decision-making* [Paper presentation]. Proceedings of the 56th Hawaii International Conference on System Sciences (pp. 312–322). https://hdl.handle.net/10125/102667

Yano, E. T., de Abreu, W., Gustavsson, P. M., & Ahlfeldt, R.-M. (2015). *A framework to support the development of cyber resiliency with situational awareness capability* [Paper presentation]. 20th International Command and Control Research and Technology Symposium (pp. 16–19).

Zhang, R., Duan, W., Flathmann, C., McNeese, N., Freeman, G., & Williams, A. (2023). Investigating ai teammate communication strategies and their impact in human-ai teams for effective teamwork. *Proceedings of the ACM on Human-Computer Interaction*, *7*(CSCW2), 1–31. https://doi.org/10.1145/3610072

About the authors

Beau Schelble is currently an assistant professor of Industrial & Systems Engineering at the University of Tennessee, Knoxville leading the AI & Robotics for Collaborative Systems (ARCS) lab, studying AI security, information-sharing, and team cognition within human-AI teams. He received his PhD in Human-Centered Computing from Clemson University in 2023.

Rohit Mallick is currently a Human-Centered Computing Phd student in the Team Research Analytics in Computational Environments (TRACE) research group at Clemson University. He graduated with his BS from Purdue University in 2020. His research interests lie in understanding the role of emotions in human-AI teaming performance & satisfaction.

Allyson Hauptman is currently an officer in the United States Army & an assistant professor of Cyber Science at the United States Military Academy West Point. She earned her Phd in Human-Centered Computing from Clemson University in 2024. Her research interests include adaptive autonomy & AI usage on the battlefield.

Nathan McNeese, PhD, is currently the McQueen Quattlebaum Endowed Professor of Human-Centered Computing at Clemson University, Director of CU-CHAI, and the Founding Director of the TRACE Research Group. He received his Information Sciences & Technology PhD from The Pennsylvania State University in 2014. His research interests lie in human-AI teaming.