Computers in Human Behavior 122 (2021) 106852

Contents lists available at ScienceDirect

Computers in Human Behavior

journal homepage: www.elsevier.com/locate/comphumbeh

What Happens When Humans Believe Their Teammate is an AI? An Investigation into Humans Teaming with Autonomy

Geoff Musick [a,*], Thomas A. O'Neill [b], Beau G. Schelble [a], Nathan J. McNeese [a], Jonn B. Henke [b]

[a] Clemson University, School of Computing, 821 McMillan Road, Clemson, SC, 29631, USA
[b] University of Calgary, Department of Psychology, 2500 University Dr NW, Calgary, AB, T2N 1N4, Canada

ARTICLE INFO

Keywords:
Human-autonomy teaming
Human-agent interaction
Team cognition
Implicit coordination
Teamwork

ABSTRACT

As artificial intelligence (AI) continues to grow in proficiency, the potential for AI to be used as team members rather than tools is becoming closer to realization. This advancement is driving new research investigations into the applicability of human-human teamwork knowledge to the context of human-autonomy teaming. In the current study, we apply qualitative methods to explore how the perceived composition of a team (how many humans and how many agents on the team) affects sentiments toward teammates, team processes, cognitive states, and the emergence of a system of team cognition. A total of 46 teams completed a teamwork simulation task and were interviewed afterwards regarding their teamwork experience. All of the teams were comprised of only humans; however, two conditions were led to believe that their teammate(s) were autonomous agents. Interviews were analyzed using grounded theory and the Gioia methodology, which revealed thematic differences between the team compositions. In light of our results, we offer a new model that describes how early-stage action teams achieve effective team processes and emergent cognitive states.

1. Introduction

Over the years, a great deal of literature has examined factors leading to high performance in work teams (Mathieu et al., 2018; O'Neill and Salas, 2018; Salas et al., 2008). Recently, rapid advances in computational science, artificial intelligence (AI), human-automation interaction (HAI), and human-computer interaction (HCI) have gradually led to scenarios in which "synthetic agents" are being used as fully-fledged team members (Chen, 2018). Indeed, a recent systematic review (O'Neill, McNeese, Barron, & Schelble, 2020) focused on *human–autonomy teams* (HATs; McNeese et al., 2018) found 76 empirical studies, demonstrating the conceptualization and advancement of the topic. A HAT involves at least one human working interdependently with at least one form of autonomy, where an autonomous agent represents a computational sub-system partially or completely self-governed with respect to team activity (e.g., decision making, adaptation, task behaviors, communication; Demir et al., 2016). Given the current pace of technological development, the use of HATs is expected to proliferate in the future (Endsley, 2017). Therefore, determining the factors that lead to high performance in HATs is of critical importance (Larson & DeChurch, 2020).

Rich and Sidner (1997, p. 284) wrote: "We take the position that agents, when they interact with people, should be governed by the same principles that underlie human collaboration." One of the critical factors potentially underscoring performance in HATs, however, may involve human perceptions of autonomous agents. On the one hand, landmark research by Nass and colleagues found that when a human was tasked with working interdependently with a computer, the human anthropomorphized the computer by being polite to it, viewing it as cooperative, and experiencing the computer as friendly (Nass et al., 1994, 1995, 1996; see also; Burgoon et al., 2016; Walliser et al., 2017). This line of research suggests that, in a team context, humans may interact with autonomous agents as they would with other humans.

Despite the aforementioned research, other work clearly suggests that humans may not trust AI (Glikson & Woolley, 2020). Indeed, Nomura and colleagues (Nomura et al., 2004) developed a Negative Attitude toward Robots Scale (NARS) to assess people's general negative attitudes toward artificially-intelligent robots. Factors contributing to negative perceptions of autonomous agents include low reliability and transparency (Mercado et al., 2016), and a lack of independence and agency (Wynne & Lyons, 2018). Hong and Curran (2019) found that artwork developed by humans versus AI was judged to be of equivalent

* Corresponding author.
E-mail address: gmusick@clemson.edu (G. Musick).

https://doi.org/10.1016/j.chb.2021.106852
Received 28 January 2021; Received in revised form 6 April 2021; Accepted 30 April 2021
Available online 4 May 2021

quality, unless the participants were made aware that the art that was AI-generated. When humans knew whether the art was AI-generated, humans holding negative cognitive schemas rated the AI art as lower in quality than the human art. This suggests that humans hold a general, over-arching social-cognitive bias against autonomous agents, and this may explain why HATs rarely perform as well as human-human (i.e., all human) teams (e.g., Demir et al., 2016; Fan et al., 2005; McNeese et al., 2018).

Therefore, there appear to be two competing perspectives regarding how humans treat autonomous agents in a team context. Humans may anthropomorphize autonomous agents and work with them as they work with other humans, which should be conducive for high-performing HATs. On the other hand, they may be predisposed to viewing autonomous agents as tools lacking in agency, self-determination, and intelligence (Lyons et al., 2019; Wynne & Lyons, 2018), which might prevent them from engaging in collaboration and coordination, mutual learning, and collective goal pursuit. Notably, the implications of these two potential possibilities are profound. If humans treat other autonomous agents as they would human teammates, then the challenge in developing high-performing HATs is one of design (Rich & Sidner, 1997). Specifically, the design would focus on attempting to mold the autonomous agents after attributes and behaviors of effective team members (e.g., Rousseau et al., 2006) as well as critical member role task requirements and interdependencies (Johnson et al., 2011). On the other hand, if humans behave in a manner that reflects a general bias against engaging in teaming activity with the autonomous agents, the problem is one of social cognition and human perceptions of AI. This issue may need to be addressed through various interventions, such as enhanced experience, better education on AI capabilities, and training on how to work with autonomous agents in a team context (e.g., Cohen & Imada, 2005; Nikolaidis et al., 2015).

In the current research, we set out to qualitatively study HATs using a unique research design involving the participant confederate approach. We used a laboratory task that involved three team member roles. We considered all human teams, as well as humans purportedly playing with either one or two autonomous agent team members. However, all of the purported autonomous agents in this research were actually humans. What was unique about this was that it controlled for differences in autonomous agents actual behaviors by actually making the autonomous agents identical in behavior to human team players. Therefore, if differences involving teams knowing they were all human versus teams believing they were part of HATs were observed and reported on by participants, it must mean that *perceptions* of working with autonomous agents led to different experiences in the HATs and not the capabilities of the autonomy itself. We approached the study using qualitative inquiry, given the novelty of this research and the uncertainty to what we might observe. Qualitative methods allowed us to investigate the emergence of themes based on the experience and perceptions of the participants, which offered a richness that a quantitative study would not offer.

1.1. Previous HAT research

Several studies have compared human-human teams versus HATs. McNeese et al. (2018) investigated the role of a synthetic agent developed using ACT-R cognitive modeling architecture capable of serving as a full-fledged teammate. The testbed involved the CERTT UAS-STE, involving three-member teams, with the agent playing the role of photographer along with humans playing the roles of navigator and pilot. The findings indicated that human-human teams outperformed the HATs. Specifically, HATs were not as effective in processing targets to be photographed at waypoints or using situation awareness (a team cognitive state) to adapt to the spontaneous introduction of new waypoints. The authors interpreted their findings as reflecting poor communication in the HAT condition, as analyses revealed that human-human teams shared more high-quality information than HATs. Other studies reported similar trends (Cooke, Demir, & McNeese, 2016; Demir et al., 2016; Myers et al., 2018) and collectively find that humans and autonomous agents have difficulty engaging in effective team processes such as explicit and implicit coordination, information exchange, and adaptability (cf. Salas et al., 2005). An exception included research by Fan and colleagues, who found that HATs performed better than human-human teams (Fan et al., 2010) but, again here, this was attributed to more robust team processes involving efficiency in distributing and managing workload across team members (Fan & Yen, 2010). Therefore, for HATs to perform well, members must engage in effective team processes involving communication and implicit coordination, which likely requires shared mental models and the emergence of team cognition.

One might wonder why the majority of research comparing human-human teams to HATs finds that HATs perform worse and have deficient team processes. One possibility is the lower capabilities of autonomous agent teammates used in those studies. For example, McNeese et al. s (2018) synthetic agent had a limited range of text that it could use and understand for communication, whereas the human-human teams in those studies were free to communicate through text using the language of their choosing. The synthetic agent may also not have been effective in adapting to other team member needs or anticipating behaviors, explaining the HATs overuse of pulling messages (requests for actions that should already have been done), and underuse of pushing messages (anticipating actions that team members would need and requesting confirmation). If the autonomous agents used in previous studies have weaker capabilities for performing the role and interacting with teammates, it is not surprising that HATs underperform human-human teams.

However, other studies have used a Wizard of Oz design to compare human-human teams versus HATs. Using the CERTT UAS-STE testbed, Demir et al. (2018) informed participants in the HAT condition that the photographer was an autonomous agent when, in fact, the photographer was a human confederate. In the human-human team, two participants were located in one room, and the third member was a human located in another room. The isolated human s role was the same as the purported autonomous agent in the HAT condition (i.e., pilot). The results suggested that even though the only difference in each condition was the description of the nature of the pilot teammate, and thus the beliefs of the navigator and photographer about the pilot (p. 260), HATs performed worse and exhibited fewer planning ahead behaviors than did human-human teams. Using a different testbed but similar Wizard of Oz design, Walliser et al. (2017) found that HATs performed worse than human-human teams in a non-team condition. Specifically, in the non-team condition, human participants were told that their performance would be evaluated independently of the autonomous agent. In the team condition, participants were told that their performance was interdependent with the autonomous agent. In this latter team condition, the differences between HATs and human-human teams vanished. These studies tell us that (a) the mere perception that a teammate is an autonomous agent affects team processes and performance independent of the actual behavior of the autonomous agent, and (b) that priming team interdependence may minimize the observed differences among human-human teams and HATs. Collectively, this reveals the potentially powerful effects of human perceptions and social cognitions involving teamwork with autonomous agents.

1.2. Fundamental team processes and states in HAT action teams

1.2.1. HAT action teams

As reported in O Neill et al. s (2020) review, current research on HATs has primarily been conducted in laboratory environments involving simulation-based command and control, emergency rescue, and other computer games that require cooperation and communication among team members to complete tasks (e.g., B4WT [blocks for world teams]; Harbers et al., 2011). The USARSim is a search-and-rescue game in which the team explores an unknown environment and identifies as many positions of victims as possible (Lewis et al., 2011). The Cognitive Engineering Research on Team Tasks Unmanned Aerial System-Synthetic Task Environment (CERTT- UAS-STE) is based on the United States Air Force Predator UAS ground control station and requires three interdependent teammates in distinct roles (pilot, navigator, and photographer) to efficiently take photographs of waypoints (Cooke et al., 2016; McNeese et al., 2018). Notably, all empirical studies of HATs focus on simulated team tasks such as these.

In light of the above, it appears that existing investigations and applications of autonomous agents tend to be within the action phases of teamwork (cf. Marks et al., 2001). Action phases are periods of intensive interdependent task execution activity (as opposed to transition phases such as formal strategic planning, goal selection, pre-task planning, and conflict management). Similarly, extant research appears to be conducted within a specific type of team referred to as action teams by Sundstrom (1999, pp. 20 21). Prototypical action teams involve surgery, military, and expedition teams (Driskell et al., 2018) which represent the current and future anticipated usage of many HATs (Schelble, Flathmann, & McNeese, 2020). This focused use of HATs is not surprising given that AI and autonomous agents current capabilities are centered around executing task-related functions rather than dealing with unstructured and abstract problems (e.g., creating a team mission statement or team values framework).

1.2.2. *Implicit team coordination*

A fundamental team process underscoring success in an action-related team activity is implicit team coordination. Team coordination is generally defined as involving the orchestrating the sequence and timing of interdependent actions (Marks et al., 2001), thereby ensuring integration and alignment in actions (DeChurch & Marks, 2006). Action teams must make decisions quickly, execute interdependent tasks seamlessly, and communicate essential information efficiently (Butchibabu et al., 2016). Therefore, they must go beyond a more explicit, pre-planned coordination approach which takes time to develop and update during transition phases occurring before and between action phases and focus on achieving implicit team coordination (Blinkensderfer et al., 2010; Mohammed et al., 2010). Rico and colleagues (Rico et al., 2008) defined implicit coordination as the ability of a team to act in concert by predicting the needs of the task and the team members and adjusting behavior accordingly, without the need for overt communication (p. 165). During action phases, teams rely on implicit coordination in a dynamic and ongoing mutual adjustment process involving the anticipation of others future actions and engaging in subsequent actions informed by predictions of others behavior (Fisher et al., 2012).

1.2.3. *Shared mental models*

There is substantial literature investigating how to maximize implicit team coordination in human action teams, and much of it points to the critical role of shared mental models (Mohammed et al., 2010; Fisher et al., 2012; Rico et al., 2019). The shared mental models construct resides within the broader concept of team cognition, which focuses on aspects of knowledge location and sharedness within the team (Salas & Fiore, 2004). Importantly, shared mental models were invoked to explain how teams were able to demonstrate coordinated activity with very little communication during intensive task activity (Mohammed et al., 2010). A team with strong shared mental models demonstrates a common understanding of how member roles need to be integrated to coordinate task execution as efficiently as possible (Rentsch et al., 1994). Not surprisingly, in the literature on human teams, shared mental models have been strongly related to team coordination and, as a result, high performance in action tasks (Fisher et al., 2012; Gabelica et al., 2016; Smith-Jentsch et al., 2005, 2008). Mental models that do not converge leave the team in a scattered, discombobulated state that results in duplication of work, suboptimal decisions, and disconnected activity among team members.

1.2.4. *Communication*

Classic multilevel theoretical treatments view communication and interaction as the cornerstone for the emergence of team-level phenomena (e.g., Kozlowski & Klein, 2000; Morgeson & Hofmann, 1999; Rousseau, 1985). Members interactions in teams, such as various forms of communication and task behaviors, signal intentions, assumptions, and preferences (Mckinney, Barker, Davis, & Smith, 2005). At team formation, members are unfamiliar with each other and how the team will work as a cohesive unit (Kozlowski & Bell, 2008, pp. 15 44). It is through repeated cycles of communication and interaction that a team may come to develop effective team processes and emergent states (O Neill & McLarnon, 2018). Therefore, high-quality communication is a team process needed to encourage the development of highly efficient coordination implicit coordination that is supported by an intact shared mental model (Entin & Serfaty, 1999; McNeese et al., 2018). Accordingly, communication must be examined in the context of team processes and emergent states in HAT research.

1.3. *The current study*

In order to disentangle the effects of the limited capabilities of autonomous agents with human social cognitive perceptions of those agents on the emergence of team cognition involving team processes (i.e., communication, implicit coordination) and cognitive states (i.e., shared mental models), we implemented an elaborated version of the Demir et al. (2018) and Walliser et al. (2017) research designs. To achieve this, in three-member teams, participants were led to believe that their team members were autonomous agents when those team members were really other human participants who were *also* led to believe that their peers were autonomous agents (referred to as the Multi-Agent HAT condition). If the (human) participants display a bias toward working with their counterparts, it suggests that *there is a bias unique to human perceptions of working with agents* that is entirely separate from how the autonomy functions or behaves (because, in fact, there are only humans in the team). In addition, we consider another HAT condition wherein two humans are informed that they are working with another agent (that is really just another human, referred to as the Multi-Human HAT condition). This design allows us also to examine whether having a human teammate while working with an autonomous agent would prime participants to approach their peers as legitimate teammates and positively affect the development of healthy team cognition. Finally, we considered a baseline, traditional team of three-member that were aware that they were all humans playing the game together (referred to as a Traditional Human Team).

We look at action teams specifically, as well as their underlying critical processes of communication and implicit coordination, and the critical emergent cognitive state of shared mental models. Our design methodology and qualitative analyses focused on addressing the following general research question (RQ):

> *RQ: How does the perceived composition of a team affect sentiment toward (or perceptions of) teammates, team processes (i.e., communication, implicit coordination), cognitive states (i.e., shared mental models), and the emergence (or lack thereof) of a system of team cognition?*

2. Method

2.1. *Participants*

One hundred undergraduate participants (67% female) were recruited from a large southeastern university psychology subject pool. This sample resulted in forty-six teams completing the experiment. The Traditional Human Team and Multi-Human HAT conditions had 15 teams each complete the experiment, and the Multi-Agent HAT condition had 16 teams. Each participant that completed the experiment received course credit.

2.2. IIHAT task

A team task simulation called "IIHAT" (Implicit Interaction for Human-Autonomy Teams) was explicitly developed for the current experiment. The IIHAT simulation did not allow for communication to specifically isolate the effects of implicit coordination in team cognition and HATs (Entin & Serfaty, 1999; Hanna & Richards, 2014, 2015; Shively et al., 2017). Additionally, this controlled for text-based communications that may have revealed that the purported autonomous agents were actually other human participants.

The IIHAT simulation goal is for players to escape a fictitious island by collecting six objectives with players located in the top left, bottom left, and bottom-right corners of the map in as few team member moves as possible. This task was made interdependent by assigning each player unique abilities that enabled them to overcome specific terrain to help collect objectives located within that type of terrain. The simulation involved three players and displayed the current map and an information panel on the right displaying the teams' current number of moves (Fig. 1). The number of moves the team has taken increases with each team member move but does not increase if a team member chooses to skip their turn by pressing T. Moves were taken sequentially in order of Player 1, to Player 2, to Player 3, and then back to Player 1.

Every map contained six objectives that teams were required to collect to advance to the next map. Each restricted terrain area contained one objective for the team member with the appropriate ability to collect it, whereas the three other objectives were located on neutral terrain accessible to everyone. The three neutral objectives were strategically placed in brown areas of the map accessible to any player to encourage interaction and coordination among team members (Entin & Serfaty, 1999; Hanna & Richards, 2014, 2015). Specifically, these neutral objectives were placed equidistant between each pair of players. After pilot testing the IIHAT simulation, 3-s pauses to game-play were implemented after every third turn, and a 5-s pause at the beginning of each new map, in order to give players a chance to assess past actions and further plan their strategy for the new state of the simulation. Finally, players were given the ability to skip their turn for any reason (e.g., to observe what other members do and adjust accordingly).

The IIHAT simulation began with five "easy" difficulty maps, followed by five "medium" difficulty maps, and ended with five "hard" difficulty maps. Maps of increasing difficulty required more interaction, interdependence, and coordination for team efficiency and performance (seen in Fig. 2). The IIHAT simulation was designed to create high levels of team interdependence (Saavedra et al., 1993; Van de Ven et al., 1976) by presenting all team members with a joint task to complete through the non-verbal (i.e., implicit) communication needed to execute a coordinated approach and develop shared mental models over time.

2.3. Materials and equipment

A JavaScipt development toolkit known as Phaser HTML was leveraged to develop the IIHAT simulation. The simulation's multiplayer facets were handled using web services like socket.io, Node.js, and an Amazon Web Services EC2 instance and data recorded in a relational database. Participants were all separated by dividers, ensuring they could not see one another. Players used desktop computers to complete the experiment, using the WASD set of keys to move around the simulation and the T key to skip turns during the simulation.

2.4. Autonomous agent

This experiment used a unique variation of the Wizard of Oz approach to simulate an autonomous agent's appearance and actions completing the IIHAT task simulation in order to isolate the effects of human *perception* of working with an autonomous teammate. Standard Wizard of Oz uses a trained human confederate acting as technology (computer feature or autonomy), ensuring the participant believes the actions/responses they see are genuinely the technology's actions/responses (Kelley, 1983, 2018). This technique was modified in the current study by also deploying what is known as the "participant confederate approach", which has been shown to improve the face validity of studies (Leis & Reinerman-Jones, 2015; White et al., 2005). Specifically, in the Multi-Agent HAT condition, we had each human player unknowingly serve as the confederate for other Multi-Agent HAT participants. To achieve this, each of the three participants in a given session was told they were playing with other autonomous agents when they were really playing with each other. However, to effectively ensure

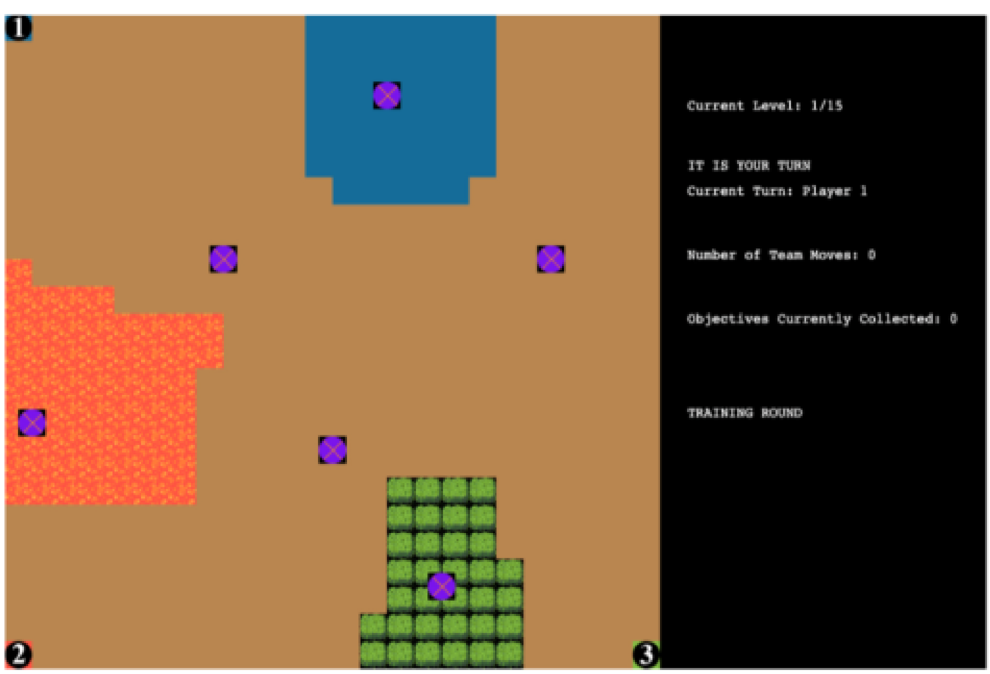

Fig. 1. IIHAT task simulation.

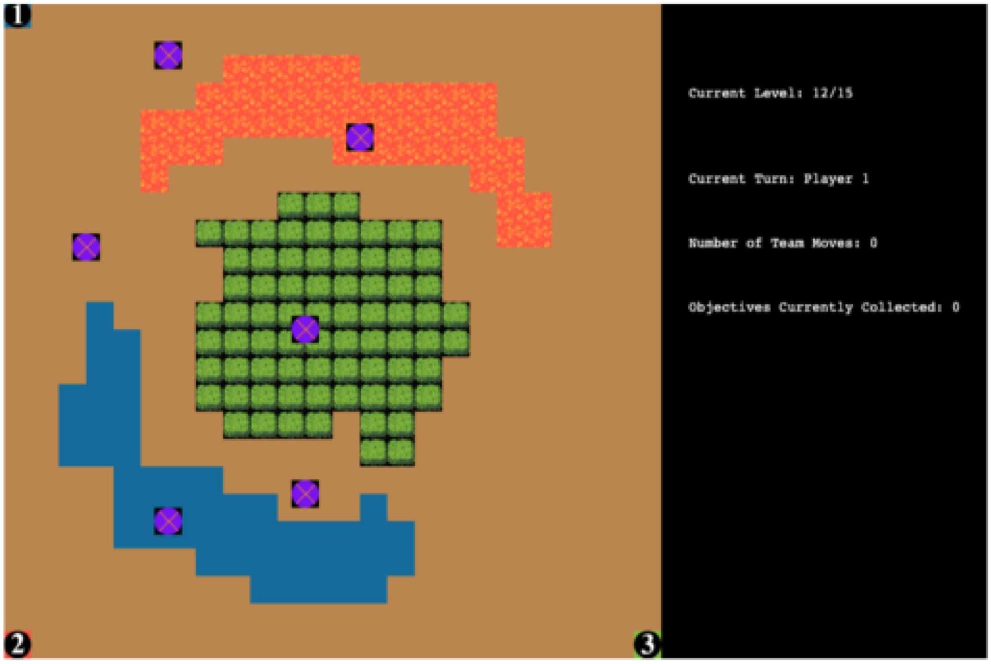

Fig. 2. A more complex map of the IIHAT task requiring more interactions between players.

the two conditions were comparable to one another, the Wizard of Oz agent in the Multi-Human HAT condition had to act the same as a typical human player, which was achieved by observing players' behaviors in the Traditional Human Team condition. From those observations, general principles were created to emulate the typical human player, which was executed by two trained confederates. A manipulation check was implemented after the focus group interview to ensure that all participants truly believed they were working with AI as their teammates, which only four individuals failed, with two in the Multi-Human HAT condition and two in the Multi-Agent HAT condition.

2.5. Procedure

Informed consent was obtained from each participant along with collecting demographic information in a pre-task survey. This pre-task survey was not for measurement or comparison purposes, and was conducted in order to collect demographic data for descriptive purposes. After the pre-task survey, participants were given an instruction sheet describing the IIHAT simulation. Participants were told they would be unable to verbally or textually communicate with one another throughout their time completing the task but that their moves may signal intentions. Each team completed the IIHAT task simulation followed by a post-task semi-structured focus group interview as a team that averaged 6 min and 43 s. The experiment took roughly an hour for each participant to complete in its entirety. The Traditional Human Team condition was given specific instructions regarding the task and that the three individuals would be working together as a team to escape each island by collecting the objectives as efficiently as possible. The Multi-Human HAT individuals were given these same instructions but were also repeatedly told that they would be completing the task with an autonomous teammate as their third teammate (Player 3). Multi-Human HAT participants were given no other information regarding their autonomous teammate's ability other than that it had been trained to complete the task as efficiently as possible. Lastly, the Multi-Agent HAT condition participants were given the same instructions as the Multi-Human HAT condition except for being repeatedly told that each participant would be working with two autonomous teammates each. Both of the conditions that included autonomous agents conveyed no outward evidence indicating a failure of the experiment's deception. A researcher was physically present throughout the task to ensure directions were followed and to observe for indications of potential deception failure.

2.6. Focus group interview

The focus group interview was conducted by the same two authors overseeing the experiment sessions. The interviews took place in the lab face to face after teams completed the IIHAT task simulation and were semi-structured with eight predefined questions targeting aspects of coordination, team cognition, and team interaction. The interview questions can be seen in their entirety in Table 1.

Each of the interview questions was subject to additional probing follow-up questions seeking to have the participants identify specific experiences from the experimental task that led them to provide the answers they did. An example probe question that would typically follow question four is as follows: "Can you provide an example of a time when everyone on your team was on the same page when gathering

Table 1
Focus group interview questions.

Number	Question
1.	Was your first thought about working with the team and adapting to their movements or about how you would do it individually and then wait for your teammates to finish?
2.	How did you think about approaching cooperatively working with your team initially?
3.	How do you think your team went about gathering objectives? Did you each collect objectives without regard to the movements of your teammates/by making a plan before you started based on what you assumed your teammates would do/depended on the board?
4.	Do you feel that everyone on your team thought about cooperating and gathering objectives the same? If not, why?
5.	Cognitively, how do you think your team determined which objectives they were responsible for and which they were not?
6.	How did your team overcome potential differences in team members island escape strategies?
7.	Do you feel team cognition was established within your team? If needed describe team cognition.
8.	Did the model/common ground happen in the early games or later games?
9.	What do you feel led to having/or not having team cognition?

objectives?" After the experiment was completed, all interviews were voice recorded and transcribed by a team of three researchers with the help of Otter.AI.

2.7. Analytic approach

The qualitative analysis used to analyze the interview data was based on grounded theory (Glaser et al., 1968) in order to investigate how human team members cooperate with and build team cognition with what they perceive as AI teammates. The interview data was coded according to the following steps: (1) two of the study's authors read through all of the interview transcripts to ascertain a basic understanding of the participants' perceptions of working with AI teammates; (2) the same two authors independently summarized the interview data into a set of themes based on participants' narratives of their experience working with their human and/or perceived AI teammates during the experimental task; (3) all the authors collectively discussed and refined the themes to ensure the participants' experiences were accurately and wholly understood; (4) the initial two authors re-read the interview transcript data to extract example quotations to represent the identified themes; (5) all authors collectively discussed, refined, and organized themes and aggregate dimensions to generate a comprehensive understanding of how the perception of working with an AI teammate affects individuals trust, team cognition, and cooperation, with this information coming together to culminate in a model of HAT development. This qualitative interview analysis methodology is common and was modeled after past research (McDonald et al., 2019; Zhang et al., 2021), and the Gioia methodology regarding themes (Gioia et al., 2012).

To ensure that qualitative rigor was applied to the organization and presentation of concepts and themes, the Gioia qualitative methodology was adopted (Gioia et al., 2012). This methodology is commonly used to capture "concepts relevant to the human organizational experience in terms that are adequate at the level of meaning of the people living the experience and adequate at the level of scientific theorizing about the experience" (Gioia et al., 2012, p. 16.) This analytic approach involved the first four of the five steps, described in the previous paragraph,

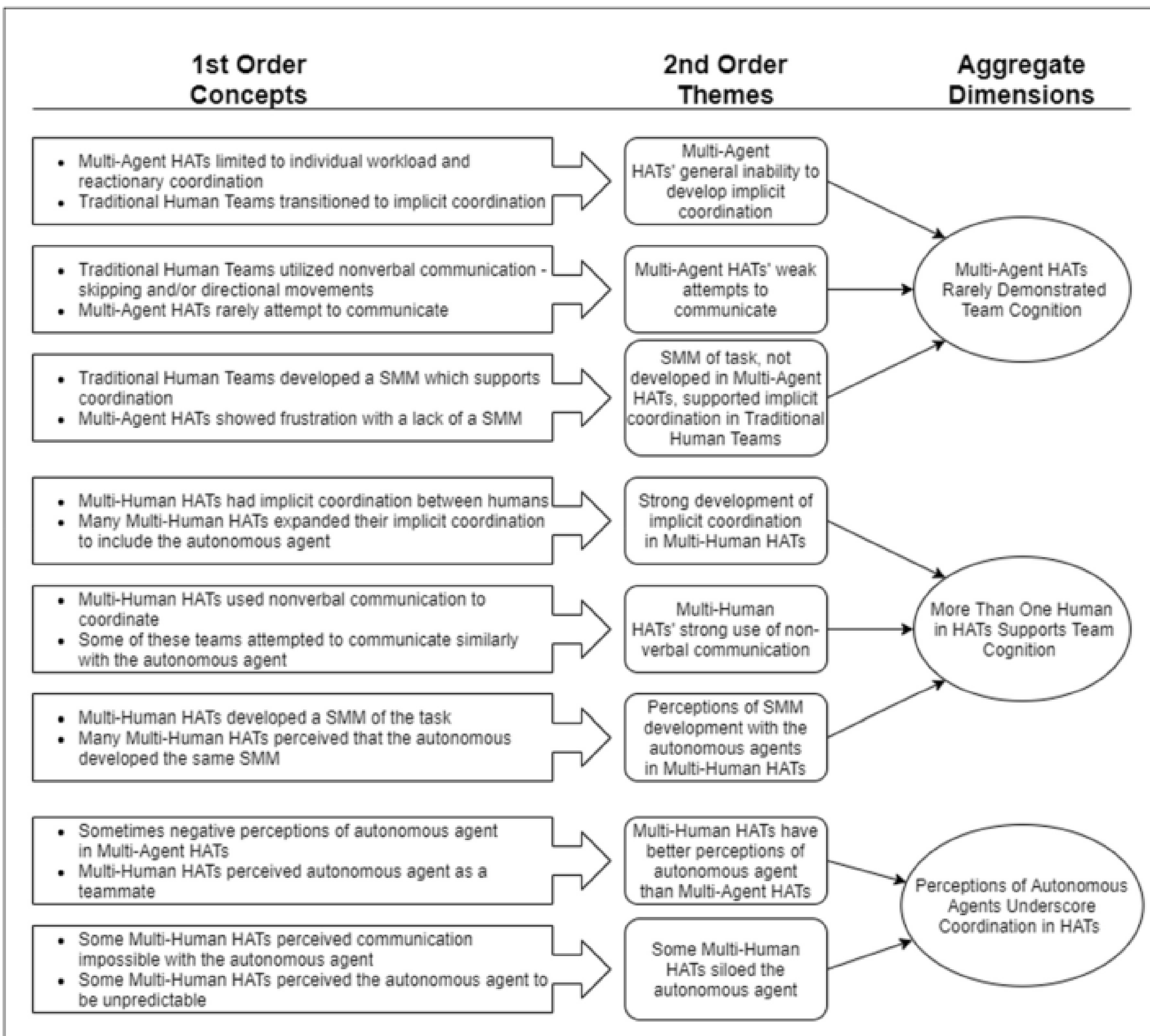

Fig. 3. Data structure of thematic analysis.

which resulted in the emergence of numerous 2nd-order themes that encapsulated various concepts. Step five involved the further distillation of 2nd-order themes into aggregate dimensions used to create a data structure (Fig. 3). This data structure serves as both a valuable visual aid and a graphical representation that communicates the progression from raw data to themes.

3. Results

Thematic analyses following the Gioia (2012) qualitative methodology revealed eight themes defined by common first-order concepts expressed in the participant quotes. The eight themes were sorted into three aggregate dimensions. We begin by reviewing the themes under each of the three aggregate dimensions and quotes that exemplify each of them. These aggregate dimensions include: 1) a comparison between processes in Traditional Human Teams and Multi-Agent HATs, 2) a comparison between Multi-Human HATs and Multi-Agent HATs in terms of processes, and 3) a comparison between Multi-Human HATs and Multi-Agent HATs in terms of perceptions. We then frame the Discussion around an integration of the themes and advance a proposed model of HAT development in the context of action tasks.

3.1. Multi-Agent HATs rarely demonstrated team cognition

The first aggregate dimension reflected a comparison of themes involving the Traditional Human Team versus Multi-Agent HAT conditions. Three themes emerged: Multi-Agent HATs general inability to develop implicit coordination, weak communication attempts, and lack of SMMs.

3.1.1. Multi-Agent HATs general inability to develop implicit coordination

As a baseline form of coordination, all conditions began by focusing on their individual workload and coordinating behavior in response to other team member actions:

> [I would] just see what route they took to see what route I needed to take. -Multi-Agent HAT 2-P1 (Condition Multi-Agent HAT; Team 2; Participant 1)

> I kind of made a route in my head but then if one of them moves one way or another I d adapt. -Traditional Human Team 11-P3

Both of these quotes described a behavior involving noticing the actions of teammates and then reacting as appropriate to individual behaviors. Importantly, this reactive coordination was the dominant approach to coordination reported by participants in Multi-Agent HATs. Most of the Traditional Human Team participants, on the other hand, reported advancing their coordinating behaviors beyond reactive coordination to include implicit coordination:

> I think in the beginning we all just went to the closest one to us, but sometimes that might have been closer to someone else. But towards the end, I think we figured that out and started to maybe going to one that was a little more out of our way [that was more] out of someone else s way. -Traditional Human Team 14-P2

Though implicit coordination took many forms, a common thread in comments from Traditional Human Team participants involved the use of explicit proactive and anticipatory behavior. As Traditional Human Team 14-P2 described, this coordination involved proactively determining what a teammate was likely to do and the movement requirements associated with that action. By thinking about this ahead of time and making comparisons to their own objectives, team members could implicitly coordinate as a team, which incorporated but exceeded individual, inward-focused and reactive strategies.

3.1.2. Multi-Agent HATs weak attempts to communicate

When Traditional Human Teams participated in implicit coordination, which involved watching their teammates, extrapolating their expected moves, and adjusting their own behavior, participants often reported the desire to improve and support this coordination through ongoing nonverbal communication, such as signalling intentions. Specifically, every turn, players had the option of moving in a direction or skipping their turn. Since participants were unable to communicate with each other verbally, Traditional Human Team players often made communication attempts through the skipping turns feature. This pattern is exemplified in the following:

> I would also check to see how far away I was compared to the objective and then compared to them and if I really didn t think that I should move because I was farther away then I just kept my spot, even if they weren t going to go for it just to show that I felt like that was a better move. -Traditional Human Team 10-P2

Described by Traditional Human Team 10-P2, this signalling not only supports coordination but communicates feelings of what might be a better move or a better approach. This communication, therefore, supported the development of implicit coordination but also SMMs (see below in section 3.1.3). Additionally, Traditional Human Teams also attempted nonverbal communication through particular directional movements:

> I could tell that I could get an objective that they maybe would have also gotten but I know I could ve gotten there faster so I started moving in that direction to make them see that it was easier. -Traditional Human Team 6-P3

Similar to the use of skipping turns, Traditional Human Team 6-P3 described communication that supported implicit coordination through the use of directional movements. Like skipping turns, this form of nonverbal communication was effective at conveying intentions, which participants found helped them anticipate others actions. Players would explicitly make a clear directional move in an early phase of a game or at a crucial decision point in an attempt to ensure their intentions were less ambiguous. Notably, both forms of nonverbal communication (skipping and directional movement) were rarely used in Multi-Agent HATs. Instead, Multi-Agent HATs were more likely to express frustration with their perceived inability to communicate:

> I mean, it was teamwork. But it also kind of felt like individual because we couldn t talk to them. -Multi-Agent HAT 9-P1

> You really can t. You can t communicate verbally, which is the frustrating part. -Multi-Agent HAT 18-P1

These quotes were common in the Multi-Agent HAT condition. In contrast, Multi-Agent HAT members reported being more likely to focus on how they could not communicate rather than capitalizing on the ways in which they could communicate nonverbally. Thus, Multi-Agent HAT members reported engaging in very limited attempts to communicate; and therefore, participants rarely felt they behaved as a team and/or progressed to a state involving a sense of shared cognition.

3.1.3. SMM of task, not developed in Multi-Agent HATs, supported implicit coordination in Traditional Human Teams

A common theme reported by members of Traditional Human Teams was their perception that their team had SMMs:

> So, I think with all of us having the same mindset, we flowed really well and then just got better as it went. -Traditional Human Team 6-P1

Traditional Human Team 6-P1 referred to having the same mindset, suggesting that their team s SMM of the task improved over time. This SMM, in turn, supported their ability for implicit coordination or flow as the game progressed. Similarly, Traditional Human Team 16-P3 articulated a SMM of the task:

> But then after a while, you d see when they pause and it was their way of saying okay, you get it. You see them going left or right or up and where they were.

This quote describes how the previously mentioned nonverbal communication used by Traditional Human Teams was perceived to be acknowledged and contributed to a common mutual understanding by each of the team members. Thus, a SMM of this aspect of the task (communication and its interpretation) was in alignment across teammates. In this context, these teams were able to develop a SMM for communication and its interpretation without any verbal communication or debriefing to confirm signals and their meanings. Further, Traditional Human Teams often referred to a gradual improvement in a SMM with their teammates over time:

> [Team cognition] happened as we understood more where the other people would go. -Traditional Human Team 4-P2

> Also, from the previous maps you can kind of see a pattern so you can assume where they might go. -Traditional Human Team 5-P2

As they gained experience with these teammates, it appears that Traditional Human Team members coalesced around a SMM. Their mental model of their teammates supported their ability to predict what objectives particular teammates would go after. However, Multi-Agent HATs were much less likely to describe a process pertaining to SMM development. Instead, Multi-Agent HATs were much more likely to discuss their frustration with a lack of SMMs than Traditional Human Teams.

> There is like a few times where one of the [AI] would just grab like an extra one and then it was like something that I was on the path going to but then I just be like, okay, I ll just chill here. -Multi-Agent HAT 7-P1

> I was kind of like yeah dude, like he wasted so many extra turns doing that but like, it s okay. -Multi-Agent HAT 9-P1

> I was kind of stuck here. I was like thinking they were going to go this way. And then they didn t. -Multi-Agent HAT 12-P1

As suggested by these quotes, Multi-Agent HAT participants expressed frustration with their autonomous agent teammates for not having a SMM of the task and refusing to make attempts to develop one. Multi-Agent HAT 17-P1 shared one such view of the autonomous agent in the following quote:

> I didn t think it was adapting all to what I was doing. I thought it was just, I was just kind of going off what they were doing, and they weren t changing.

Multi-Agent HAT 17-P1 appeared to think that the autonomous agents were unable to adapt and change. This perception would likely undermine any motivation to try to develop levels of coordination, such as the implicit communication described by many Traditional Human Team participants.

3.2. More than one human in HATs supports team cognition

The second aggregate dimension reflected a focus on first-order concepts and themes applicable to Multi-Human HAT members, relative to those in Multi-Agent HATs. Three themes emerged: Multi-Human HATs relatively strong development of implicit coordination, Multi-Human HATs strong use of non-verbal communication, and the perceived SMM that included the autonomous agent on Multi-Human HATs.

3.2.1. Strong development of implicit coordination in Multi-Human HATs

As was reported in the other two conditions, teams in the Multi-Human HAT condition began by focusing on individual workload and by using reactionary coordination:

> Well, I guess if we both tried to go for an objective, if you or the AI kept going, then I would just stop because I realized you were going to get it. -Multi-Human HAT 9-P1

However, a distinguishing factor between the Multi-Agent HAT and Multi-Human HAT conditions is that participants on Multi-Human HATs thought they were teaming with at least one human. Quotes suggested these teams were more likely to proactively observe their teammates and coordinate actions like those in the Traditional Human Team condition:

> I looked at my character first, and then I looked at what was on my path. And then I looked to see their area and what was on their path to see if there was like an outlier, that maybe I was closer to get than they were. -Multi-Human HAT 11-P2

> If you re [taking] your path to the objective and I only have to go two moves this way, and my teammate would have to go like five moves this way it was kind of just like, watching their movement and then watching your movement and like, just kind of the unspoken communication of like, I ll grab this one for you. -Multi-Human HAT 16-P1

As suggested in these quotes, Multi-Human HAT 11-P2 and 16-P1 both developed a proactive and anticipatory approach that deliberately considered their teammates likely movements and requirements. In this way, Multi-Human HATs perceived teamwork and participated in coordinating behaviors in a similar way to Traditional Human Teams rather than the other HAT condition (Multi-Agent HATs). Interestingly, Multi-Human HATs did not solely limit their implicit coordination to their human teammate:

> Yeah, I think it was also important to know which ones, if your teammates had to go a certain way to get to the ones that they can only get to going through that path, which ones they would also pick up. -Multi-Human HAT 4-P1

> It s like at first it was - what you need to do to get your square and like how you re going to play in your path and then later was more like, okay, well, how can somebody else probably get it before me and kind of seeing the strength in your other partners. -Multi-Human HAT 15-P1

Multi-Human HAT 4-P1 and 15-P1 both described the same proactive mindset and used plurality to describe their teammates and the strength in your other partners. In this way, Multi-Human HAT participants extended their implicit coordinating behaviors to encompass the entire team including the autonomous agent rather than only the human teammate. This ability for Multi-Human HATs to develop implicit coordination stands in stark contrast to Multi-Agent HATs which were primarily limited to reactive coordination as described in the first aggregate dimension.

3.2.2. Multi-Human HATs strong use of non-verbal communication

Similar to Traditional Human Teams, but unlike Multi-Agent HATs, Multi-Human HATs reported strong evidence of nonverbal communication:

> If I would skip a turn, that would try to communicate that someone else should go for an objective if I skipped my turn. -Multi-Human HAT 1-P1

> Yeah, instead like going up and just keep going up. Doing one way or the other. -Multi-Human HAT 11-P1

This non-verbal communication reported by Multi-Human HAT members was similar, or indistinguishable, from Traditional Human Teams given the active use of skipping (described by Multi-Human HAT

1-P1) and directional movements that communicated intention (described by Multi-Human HAT 11-P1). In this way, Multi-Human HATs were able to communicate to their teammates what they thought the best course of action should be. Like with implicit coordination, Multi-Human HAT members included the autonomous agent in this communication:

> Me and the AI would be in the same path and then when I would go up or down, like rather than keep going towards each other, like, okay, well now we know what we re doing. -Multi-Human HAT 15-P1

In this quote, Multi-Human HAT 15-P1 described the use of directional movements, which were utilized by both Traditional Human Teams and between humans on Multi-Human HATs, to signal intent or strategy. However, this description is specific to an interaction between a human and agent. The participant described a reciprocated interaction with the agent that involved going in a specific direction to avoid ambiguity regarding their intention. Similarly, skipping turns was described to include the autonomous agent:

> When I saw teammates skipping turns, I knew that that was a thing for me that developed more trust because I knew then that I was supposed to go towards another objective. -Multi-Human HAT 4-P1

In this quote, Multi-Human HAT 4-P1 described how seeing teammates (both human and autonomous) skip turns improved trust and coordination. Unlike the Multi-Agent HATs who were frustrated with the perceived impossibility of communication, Multi-Human HATs described both forms of nonverbal communication, skipping and directional movement signalling, which supported their coordination similar to Traditional Human Teams.

3.2.3. Perceptions of SMM development with the autonomous agents in Multi-Human HATs

In addition to Multi-Human HAT members engaging in communication and developing implicit coordination, these teams also had stronger perceptions of SMMs:

> It [the AI] modified to the moves that we made and the one where she got three, it modified to that. -Multi-Human HAT 1-P2

> I think [the AI] had the same understanding, they re basing their decisions off of the decisions of others. -Multi-Human HAT 4-P1

> I feel like it was kind of predicting what we were doing. It was going off with that. -Multi-Human HAT 10-P2

Multi-Human HAT 1-P2, 4-P1, and 10-P2 all discussed how they perceived their teammate to be on the same page regarding strategy and decision making. Importantly, all three of these quotes described their perception of the autonomous agent, which indicated their belief that the autonomous agent had similar SMMs and participated in team cognition. The belief that the autonomous agent had SMMs was also reciprocated as Multi-Human HATs described a SMM of the task and coordinating behaviors that included the autonomous agent:

> We got the hang of it more for sure. And like, understood where to go. And like, what objectives we were supposed to get. So I guess that came with knowing what they were going to do. Yeah, I definitely knew what we were going to do. -Multi-Human HAT 17-P1

> I kind of just assumed what my teammates were going to do. -Multi-Human HAT 7-P1

As discussed previously, these Multi-Human HATs often described these perceptions using plural language to include their AI teammate (e.g., they and teammates). In these teams, the perception of SMMs extended beyond the human teammate to include the autonomous agent as well.

3.3. Perceptions of autonomous agents underscore coordination in HATs

The third aggregate dimension reflected a comparison of themes involving the Multi-Human HAT versus Multi-Agent HAT conditions. Two themes emerged: Perceptions of autonomous agents were positive in Multi-Human HATs and negative in Multi-Agent HATs, and Multi-Agent HATs tended to silo their autonomous agent teammates.

3.3.1. Multi-Human HATs have better perceptions of autonomous agents than Multi-Agent HATs

A strong contrasting theme emerged regarding opposing perceptions of autonomous agents in Multi-Agent HAT versus Multi-Human HAT conditions. First, Multi-Agent HAT members reported negatively-valenced perceptions of the autonomous agent:

> Well, sort of the beginning it was all right. And then it started to kind of break off and they started doing their own thing after a bit. -Multi-Agent HAT 12-P1

> I thought it was just like, literally just the computer. -Multi-Agent HAT 17-P1

At first, I thought the jungle one was [intellectually disabled]. It wasn t moving the right way. -Multi-Agent HAT 18-P1.

These quotes reflect what many Multi-Agent HATs thought of the autonomous agent, specifically that it was doing its own thing and unable to adapt. The quotes also clearly suggest a lack of perceived agency of the autonomous agent (literally just the computer and intellectually disabled). Though not all Multi-Agent HATs thought this negatively of the autonomous agent, this sentiment was pervasive enough to emerge as a clear theme.

Second, Multi-Human HATs were far more likely to positively perceive the autonomous agent, often going as far as referring to it as a team member:

> It [the AI] was just kind of like another person. -Multi-Human HAT 2-P1

> It was just like anyone else. -Multi-Human HAT 7-P1

> Yeah, I just based what I was going to do based off where my teammates were going to go. -Multi-Human HAT 4-P1

> I didn t really notice much of a difference. I just mostly just adapted to what was going on with the AI teammate and that was pretty much that. -Multi-Human HAT 10-P2

These participants often perceived their AI teammate to be just like their human teammate. This idea often came across in quotes subtly as Multi-Human HAT 4-P1 among many other Multi-Human HAT participants referred to their teammates instead of just describing their singular human teammate. This perception was fairly consistent in this condition as almost all Multi-Human HATs (1, 2, 3, 4, 5, 6, 7, 8, 10, 13, and 16) mentioned that they treated the autonomous agent the same as their human teammate. Thus, the increased likelihood of Multi-Human HATs to perceive the autonomous agent as a teammate along with a decreased chance to perceive it negatively likely played a pivotal role in HAT development.

3.3.2. Some Multi-Human HATs siloed the autonomous agent

We found a theme shedding some negative light on Multi-Human HATs. Specifically, not all Multi-Human HAT participants chose to engage with the autonomous agent. When this happened, the autonomous agent was effectively sidelined out of the teamwork. One way siloing was reported was through the form of simply ignoring the agent:

> I kind of felt like it wasn t really there. I think it knew what it was going to do. I feel like I was working with her more than the AI or

whatever. So, I didn t really pay attention to what it was doing. -Multi-Human HAT 3-P1

When this siloing of the autonomous agent in teamwork occurred, Multi-Human HAT participants articulated one of two perceptions behind it. First, they reported siloing the autonomous agent due to a perception that communication with the agent was pointless:

> I didn t think I was going to be able to communicate with the AI anyway, so I just kind of let it do what it s going to do. -Multi-Human HAT 4-P2

As Multi-Human HAT 4-P2 described, if participants thought that nonverbal communication was impossible with the autonomous agent, they did not attempt to communicate or coordinate with it. This limitation resulted in coordination between only humans and leaving the agent out of teamwork.

Second, some participants described the perception that the autonomous agent was too unpredictable:

> The only way I thought it was different is sometimes I didn t know the AI one was going to like, continue to go after one or not. Like I felt more confident in what she [the human teammate] was going to do than what the AI was going to do. -Multi-Human HAT 14-P2

Multi-Human HAT 14-P2 described the perception that the autonomous agent was unpredictable. Even though both teammates were actually human, participants may have brought their preconceived notions about autonomous agents into the experiment, assuming they could not interpret and respond meaningfully to communication attempts.

In sum, although having an additional human on the team in the Multi-Human HAT condition appeared to be helpful, some humans still reported limimted engagement with the agent.

4. Discussion

In the current study we investigated HATs using a unique research design that leverages the participant confederate approach (Leis & Reinerman-Jones, 2015; White et al., 2005). Specifically, all of the purported autonomous agents in this research were actually humans. However, in the Multi-Agent HAT condition, each human believed they were the only human playing with two autonomous agent team members. In reality, all were humans, and all were believing this same storyline. A critical feature of this design is that it held constant, at the sample level, differences in autonomous agents behaviors and capabilities given that the so-called autonomous agents were literally typical human team players. Furthermore, the autonomous agent behaviors were enacted by a researcher emulating typical human team players in the Multi-Human HAT condition. Accordingly, the differences we observed among human-human and the two HAT variants must be related to perceptions of working with autonomous agents, rather than limited capabilities or different behavioral patterns used by artificial intelligence.

In the Introduction, we asked whether humans would be more likely to anthropomorphize the autonomous agents or treat them as tools rather than teammates. If the former were true, we would expect HAT experiences to be relatively similar to experiences in all human teams. If the latter were true, we would expect team development to be impaired as human roles are replaced by autonomous agents, since the human(s) would not attempt to build a team that coordinates, communicates, and shares a common mental model. Importantly, our findings suggest support for both possibilities. Where there was only a single autonomous agent and another human, we found that participants often reported attempts to communicate by signalling intentions. As the team developed, coordination became gradually more implicit and members demonstrated a perceived shared understanding of the team s approach. This coordination did not happen in teams with one human led to believe they were working with two other autonomous agents, however. In this case (i.e., the Multi-Human HAT condition), participants did not report teamwork-related behaviors or attempts to build a common understanding of the task. Thus, we find somewhat different activity patterns and team development depending on whether participants were assigned to the Traditional Human Team, Multi-Human HAT, or Multi-Agent HAT conditions.

Based on our data, these differences are summarized in Table 2 which reflect the themes described in Fig. 3. In reflecting on our original research question, we found that the perception of team composition did affect sentiments toward teammates, team processes, cognitive states and the emergence of a system of team cognition. Though Multi-Human HAT members did not coordinate with their AI teammate as well as their human teammate, these teams did describe higher levels of implicit coordination, communication, and perceptions of a SMM compared to Multi-Agent HATs. These sentiment and process differences culminated in the inclusion of the agent in team cognition for Multi-Human HATs.

Below, we advance three submodels describing the process of teamwork that participants reports suggest for each of the three conditions. We begin in section 4.1 by describing the emergence and self-reinforcing cycle of what we label team cognition, based on what our participant reports suggest as a collective. Then we continue in 4.2 with a model describing the emergence of team cognition for the three conditions.

4.1. The emergence of team cognition in HATs

Based on thematic analysis using the reports from participants, team cognition was experienced in the Traditional Human Team and Multi-Human HAT conditions and involved a feedback loop outlined at the bottom of Fig. 4. Though team cognition did not emerge in all teams, the processes and interactions were relatively uniform when it did emerge. Teams displayed evidence of SMMs about the task and team, which supported implicit coordination. In this context, implicit coordination involved team members being able to proactively predict their teammates actions and strategies and adjust their behavior accordingly. Adjusting behavior took the form of teammates changing which objective(s) they were pursuing. Behavior adjustment also coincided with and took the form of team members attempting to communicate nonverbally by signalling to their teammates which objective they thought they were or were not responsible to collect. This nonverbal communication further reinforced or improved the SMMs present on these teams as teammates gained a better understanding of their teammates intentions and strategies. Importantly, some HATs were able to participate in this team cognition feedback loop (i.e., Multi-Human HATs), while others were not (i.e., Multi-Agent HATs).

4.2. A model of team cognition and its emergence in HAT action phases

Drawing from our data, we advance a model that outlines the supporting perceptions and processes necessary for team cognition to be supported and emerge in HATs performing action tasks (Fig. 4). Submodels apply to the three team conditions examined in this study, Multi-Agent HATs, Traditional Human Teams, and Multi-Human HATs. Beginning with the sub-model for Multi-Agent HATs, the model notes that these teams began with an individual mindset and relied on reactionary coordination to accomplish their goals. They gained experience in the task environment over time, however. This experience also involved making observations regarding their teammates, which influenced their perceptions of their perceived agent teammates. As participants came into the task with pre-existing attitudes toward AI, these attitudes, in conjunction with their current experience and observations in the simulation, informed whether they perceived if team cognition was even possible with the agent. For instance, Multi-Agent HATs were far less likely to participate in non-verbal communication with the agent than Multi-Human HATs. It is possible that training humans on Multi-

Table 2
Themes organized by condition and Process/Perception.

	Implicit Coordination		Non-Verbal Communication		Perception of SMM Development		Positive Perceptions of Agent(s)	Siloing Agent
	w/ Human(s)	w/ Agent(s)	w/ Human(s)	w/ Agent(s)	w/ Human(s)	w/ Agent(s)		
Traditional Human Teams	High		High		High			
Multi-Human HATs	High	Moderate	High	Moderate	High	Moderate	High	Moderate
Multi-Agent HATs		Low		Low		Low	Low	

Note. High/moderate/low refers to whether the the process was common in the condition, or the level that something was perceived (e.g., Multi-Agent HATs had low levels of non-verbal communication with the agent; Multi-Human HATs commonly had positive perceptions of the agent). Gray cells are not applicable to the condition based on team composition.

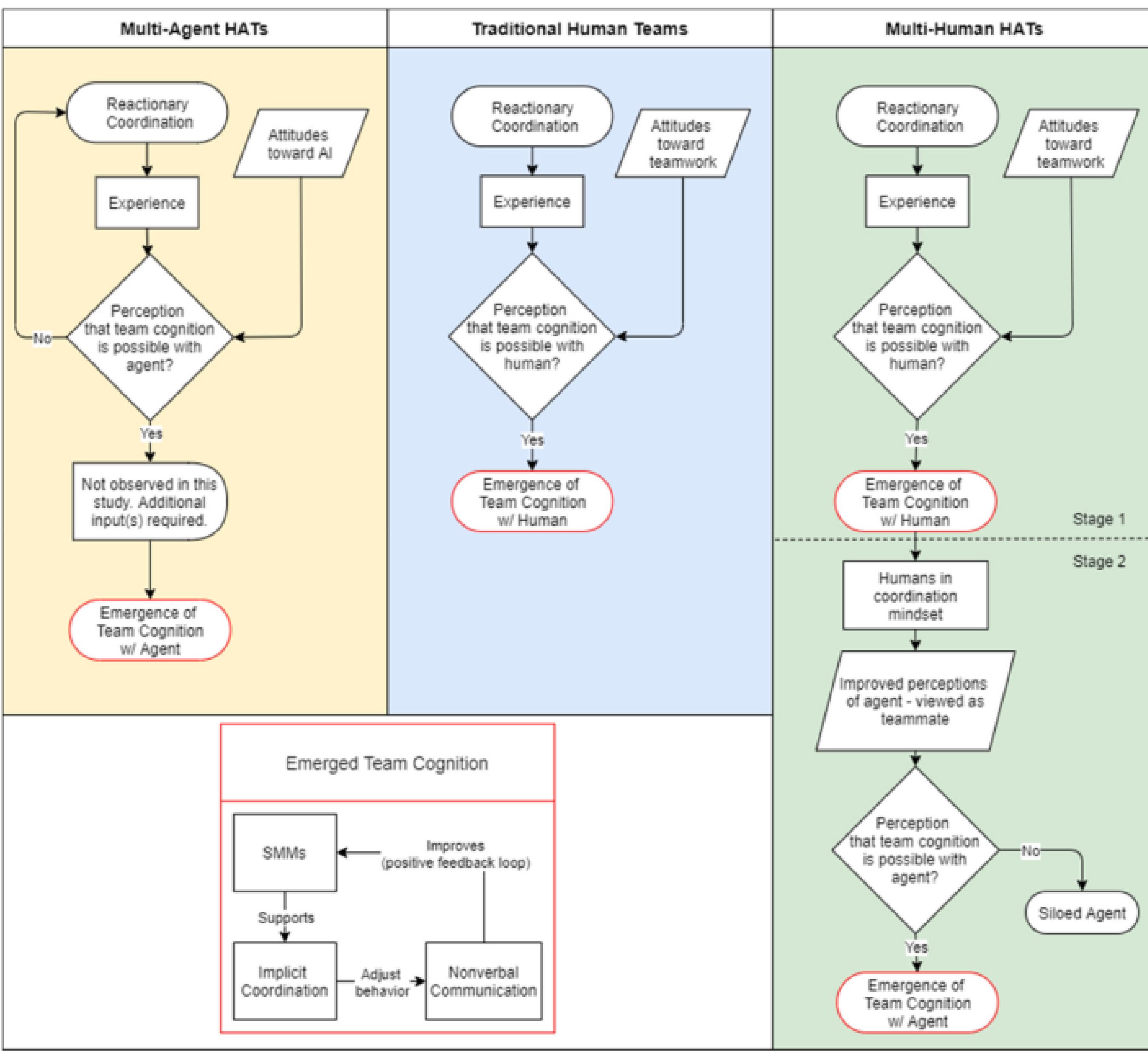

Fig. 4. A Model of Team Cognition and its Emergence in HAT Action Teams.

Agent HATs regarding the AI s communication capabilities would have alleviated this deficiency. However, it is worth noting that Multi-Human HATs were able to perceive communication possible with the AI when another human was present (more discussion on this later in this section). Due to preconceived notions of AI on Multi-Agent HATs, team cognition was not viewed as possible; and therefore, no attempts were made to develop it. Therefore, Multi-Agent HATs rarely progressed beyond purely reactionary coordination with a focus on the individual.

The middle of Fig. 4 illustrates the emergence sub-model of team cognition on Traditional Human Teams. Like Multi-Agent HATs, these teams began by relying on reactionary coordination and observation of teammates, prior to having gained experience in the team. However, over repeated iterations and cycles of work together, they progressed beyond purely reactionary coordination and individualistic tendencies. Specifically, since these were traditional all-human teams, members naturally perceived team cognition to be possible with their human counterparts given pre-existing teamwork experience, observations, and attitudes. Not surprisingly, team cognition emerged in all of the Traditional Human Teams.

Finally, the sub-model of the emergence of team cognition in Multi-Human HATs involves two stages: one for interactions among the two humans and one for human-agent interactions. Given that Multi-Human HATs understood they were working with one other human team member, this team composition afforded some of the same perceptions and processes involved in team cognition emerging between humans. Indeed, an identical pathway to that observed in Traditional Human Teams occurred in the Multi-Human HAT sub-model, thereby leading to the initial emergence of team cognition between the humans only (Stage 1). This emergence of team cognition between humans is a vital distinction in explaining why team cognition emerged in Multi-Human HATs but not Multi-Agent HATs. This team cognition between humans served as a foundation for subsequent processes and changes in perceptions necessary for the emergence of team cognition between humans and agents. Specifically, participating in team cognition with another human may have put participants in a mindset that drew attention to team coordination opportunities.

Compared to Multi-Agent HATs, Multi-Human HATs were able to break out of their individual focus because having another human team member may have primed them to seek out and attend to teamwork behaviors from others, including the agent member. As they gained experience and made observations in this mindset, team cognition was encouraged to develop. Further, participants were able to make direct comparisons between their human and agent teammates. These comparisons allowed participants to observe similarities and differences in teamwork behaviors. With these observations, participant perceptions of the agent tended to improve, and most Multi-Human HATs began viewing the agent as a teammate. These revised perceptions resulted in Multi-Human HATs perceiving that team cognition was possible and more commonly supported the emergence of team cognition with the agent in this condition. However, it is worth noting that not all Multi-Human HATs extended the team cognition to include the agent. Those that perceived team cognition to be impossible with the agent focused on team cognition with their human teammate and siloed the agent instead (i.e., never made it past Stage 1). From participant quotes, it was clear that their perceptions of the agent did not improve enough or at all (e.g., still thought the agent s movements were unpredictable). The humans on these Multi-Human HATs did their best to use their emerged team cognition to advance the team task, but this was likely suboptimal since it excluded and underutilized the agent team member.

This model contributes as a building block for researchers studying how team composition affects team processes on HATs. First, as all previous HAT team composition research has focused on differences between all human teams and HATs (O Neill et al., 2020), this model provides insight regarding deliniations in the number of agents/humans that are on the team. Though one opportunity is to add one additional human to HATs with only one human on them, the number of humans and agents on a team is often pre-determined by the requirements of the teaming context. Further, it is unclear whether subsequent additions of humans on the team (e.g., three humans and one agent) would hurt or help team outcomes. Rather, an important takeaway is that team cognition can emerge between humans and agents on HATs by either altering the initial attitudes toward the agent or through enhancing the experience gained with the agent. This feeds into our second point, which is that the model outlines opportunities for supporting the emergence of team cognition on HATs. With this model, future research can examine how it might be possible to move HATs through a pathway more similar to humans that allows for shared cognition, or help identify how to extent shared cognition from just including humans to also including the agents. Specifically, the perception that team cognition is possible with the agent can be improved by either: (1) training humans before the task begins regarding the level of autonomy or capabilities of the agent or (2) improving the transparency or communication of the agent so that humans understand the agent s capabilities through experience.

4.3. Limitations and future research

We explicitly developed this research to focus on HATs performing action-related activities, using three-member teams, and using a task that was short in duration and permitted little communication (and no verbal communication). Therefore, the generalizability of the current research findings may be called into question, and we agree that much more research is needed before we can confidently assume our results will hold in more complex, longer duration, HAT field settings. On the other hand, the HAT literature s current state requires that scholars focus on understanding fundamental processes with highly controlled research designs (O Neill et al., 2020). This focus on the fundamentals is not unlike the state of early research on small groups in the 1950s (Rosenberg, 1959) or early research on virtual teams in the 1980s (Siegel et al., 1986). Research on HATs is exceptionally nascent, and therefore, we believe highly controlled research examining basic processes is needed before introducing the complexities of field research.

Another related limitation of this research is that the development of our model should be viewed as preliminary. Given that the model was developed in a particular context, the specifics of the model s stages may not always apply. For example, would a five-member team comprising two humans and three autonomous agents engage in more similar processes to the Multi-Human HAT or Multi-Agent HAT processes of the current study? This dynamic is unclear and warrants future study. Additionally, this study did not focus on performance but rather investigated perceptions. Though perception is a logical starting point for examining team cognition, future studies will ultimately need to examine the interplay between perception and performance as it relates to HAT composition. It will also be essential to test our model quantitatively through the use of path analysis or structural equation modeling. Given the exploratory nature of the current research, it would not have been possible to measure the variables needed to test these linkages a priori, but future research can begin with a model such as ours in mind and examine how it and rival models fit newly-collected quantitative data.

One of the future directions that are sorely needed in HAT research, and further raised as a critical need by our findings, is the need for investigations of HAT training programs. Training may focus on developing more favorable perceptions of autonomous agents and enhancing teamwork involving autonomous agents. Interestingly, of the few studies that have considered training in the past, cross-training appears to be effective (e.g., Mercado et al., 2016). A plausible mechanism for this is that in HATs, cross-training may build empathy for the agent by better understanding its role. More broadly, research examining autonomous agents as teammates versus tools suggests that interventions that encourage collectively-shared outcomes and emphasizing mutual interdependence will support healthy teamwork in HATs (e.g., Nass et al., 1996; Walliser et al., 2017; Wynne & Lyons, 2018). Therefore, teamwork interventions that promote a team mindset among human team members, empathy, and an understanding that autonomous agents have a degree of agency may promote teamwork in interactions and the emergence of functional team dynamics. However, these interventions require development and testing.

Another future direction for HAT research involves the design of autonomous agents. If autonomous agents are to be treated as team members, they must be viewed as volitional, agentic, and independent of human control (to a degree). Research by Nass and colleagues (e.g., Nass et al., 1996), Wynne and Lyons (2018), O Neill et al. (2020), and several others make this clear. How to design an agent, however, that will be perceived this way requires attention. This will involve psychologists and computer scientists, because the critical levers related to perception as well as the technical design details need to be addressed holistically. For example, in the current study, for an agent to be perceived as agentic, it s movements should display a form of logic that humans *perceive* as conveying a meaningful message. Logical, or even perfect moves, that are not viewed as such by humans, would not work. Therefore, the context for which the agent is expected to function will be critical in its design (Bradshaw et al., 2003; Johnson et al., 2011).

Finally, longitudinal research is sorely needed on HATs. This long-term research is important because the literature is focused almost exclusively on short-term, ad-hoc teams, and it is very likely that the dynamic between humans and autonomous agents may evolve over time and experience (see related human teams literature, e.g., Larson, McLarnon, & O Neill, 2020). Over time and repeated cycles of interaction, learning takes place in teams, and a maturation process that aims to better utilize each member s strengths is given an opportunity to unfold (Kozlowski et al., 2016). Although HATs tend to underperform human-human teams currently (O Neill et al., 2020), it is unclear whether this advantage would persist in the long run.

5. Conclusion

In the current research, we find that HATs comprising of multiple agent members appeared to be the least effective. Given that we controlled for differences in human versus AI-related behaviors through our design, we can conclude that Multi-Agent HATs were impaired because of some social cognitive processes involved in the perception of autonomous agents as teammates. We find a different set of processes in Multi-Human HATs, which fared much better. We advance models of team development and the emergence of team cognition that attempt to explain the differences observed. Future research should aim to build on and revise these models as new data are collected and analyzed, we believe this is a useful step in advancing current knowledge on HATs.

Credit author statement

Geoff Musick: Conceptualization, Validation, Formal analysis, Investigation, Data curation, Writing original draft, Visualization. **Thomas A. O'Neill**: Writing original draft, Visualization. **Beau G. Schelble**: Conceptualization, Methodology, Software, Investigation, Writing original draft. **Nathan J. McNeese**: Conceptualization, Writing Reviewing and Editing. **Jonn B. Henke**: Validation, Data curation, Writing Reviewing and Editing.

Acknowledgments

This research was partially funded by NSF Award #1829008, AFOSR Award FA9550-20-1-0342 (Program Manager: Laura Steckman), and a grant from the Social Sciences and Humanities Research Council of Canada.

References

Blickensderfer, E. L., Reynolds, R., Salas, E., & Cannon-Bowers, J. A. (2010). Shared expectations and implicit coordination in tennis doubles teams. *Journal of Applied Sport Psychology, 22*(4), 486 499.

Bradshaw, J. M., Sierhuis, M., Acquisti, A., Feltovich, P., Hoffman, R., Jeffers, R., & Van Hoof, R. (2003). Adjustable autonomy and human-agent teamwork in practice: An interim report on space applications. In *Agent autonomy* (pp. 243 280). Boston, MA: Springer.

Burgoon, J. K., Bonito, J. A., Lowry, P. B., Humpherys, S. L., Moody, G. D., Gaskin, J. E., & Giboney, J. S. (2016). Application of expectancy violations theory to communication with and judgments about embodied agents during a decision-making task. *International Journal of Human-Computer Studies, 91*, 24 36.

Butchibabu, A., Sparano-Huiban, C., Sonenberg, L., & Shah, J. (2016). Implicit coordination strategies for effective team communication. *Human Factors, 58*(4), 595 610.

Chen, J. Y. (2018). Human-autonomy teaming in military settings. *Theoretical Issues in Ergonomics Science, 19*(3), 255 258.

Cohen, J., & Imada, A. (2005). Agent-based training of distributed command and control teams. In *Proceedings of the human factors and ergonomics society annual meeting* (Vol. 49, pp. 2164 2168). Sage CA: Los Angeles, CA: SAGE Publications. No. 25.

Cooke, N. J., Demir, M., & McNeese, N. J. (2016). *Synthetic teammates as team players: Coordination of human and synthetic teammates*. Cognitive Engineering Research Institute Mesa United States.

DeChurch, L. A., & Marks, M. A. (2006). Leadership in multiteam systems. *Journal of Applied Psychology, 91*(2), 311.

Demir, M., McNeese, N. J., & Cooke, N. J. (2016). Team communication behaviors of the human-automation teaming. In *2016 IEEE international multi-disciplinary conference on cognitive methods in situation awareness and decision support (CogSIMA)* (pp. 28 34). IEEE.

Demir, M., McNeese, N. J., & 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.

Driskell, T., Salas, E., & Driskell, J. E. (2018). Teams in extreme environments: Alterations in team development and teamwork. *Human Resource Management Review, 28*(4), 434 449.

Endsley, M. R. (2017). From here to autonomy: Lessons learned from human automation research. *Human Factors, 59*(1), 5 27.

Entin, E. E., & Serfaty, D. (1999). Adaptive team coordination. *Human Factors, 41*(2), 312 325.

Fan, X., McNeese, M., & Yen, J. (2010). NDM-based cognitive agents for supporting decision-making teams. *Human-Computer Interaction, 25*(3), 195 234.

Fan, X., Sun, S., McNeese, M., & Yen, J. (2005). Extending the recognition-primed decision model to support human-agent collaboration. In *Proceedings of the fourth international joint conference on Autonomous agents and multiagent systems* (pp. 945 952).

Fan, X., & Yen, J. (2010). Modeling cognitive loads for evolving shared mental models in human agent collaboration. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics), 41*(2), 354 367.

Fisher, D. M., Bell, S. T., Dierdorff, E. C., & Belohlav, J. A. (2012). Facet personality and surface-level diversity as team mental model antecedents: Implications for implicit coordination. *Journal of Applied Psychology, 97*(4), 825.

Gabelica, C., Van den Bossche, P., Fiore, S. M., Segers, M., & Gijselaers, W. H. (2016). Establishing team knowledge coordination from a learning perspective. *Human Performance, 29*(1), 33 53.

Gioia, D. A., Corley, K. G., & Hamilton, A. L. (2012). Seeking qualitative rigor in inductive research: Notes on the Gioia methodology. *Organizational Research Methods, 16*(1), 15 31.

Glikson, E., & Woolley, A. W. (2020). *Human trust in artificial intelligence: Review of empirical research*. Academy of Management Annals ((ja)).

Hanna, N., & Richards, D. (2014). The impact of communication on a human-agent shared mental model and team performance. In *Proceedings of the 2014 international conference on Autonomous agents and Multi-Agent systems* (pp. 1485 1486).

Hanna, N., & Richards, D. (2015). The impact of virtual agent personality on a shared mental model with humans during collaboration. In *Proceedings of the 2015 international conference on autonomous agents and multiagent systems* (pp. 1777 1778).

Harbers, M., Bradshaw, J. M., Johnson, M., Feltovich, P., van den Bosch, K., & Meyer, J. J. (2011). Explanation and coordination in human-agent teams: A study in the BW4T testbed. In *2011 IEEE/WIC/ACM international conferences on web intelligence and intelligent agent technology* (Vol. 3, pp. 17 20). IEEE.

Hong, J. W., & Curran, N. M. (2019). Artificial intelligence, artists, and art: Attitudes toward artwork produced by humans vs. artificial intelligence. *ACM Transactions on Multimedia Computing, Communications, and Applications, 15*(2s), 1 16.

Johnson, M., Bradshaw, J. M., Feltovich, P. J., Jonker, C. M., Van Riemsdijk, B., & Sierhuis, M. (2010). The fundamental principle of coactive design: Interdependence must shape autonomy. In *International Workshop on coordination, organizations, institutions, and norms in agent systems* (pp. 172 191). Berlin, Heidelberg: Springer.

Kelley, J. F. (1983). An empirical methodology for writing user-friendly natural language computer applications. In *Proceedings of the SIGCHI conference on human factors in computing systems* (pp. 193 196).

Kelley, J. F. (2018). Wizard of Oz (WoZ) a yellow brick journey. *Journal of Usability Studies, 13*(3), 119 124.

Kozlowski, S. W., & Bell, B. S. (2008). *Team learning, development, and adaptation*. Group learning.

Kozlowski, S. W., Chao, G. T., Grand, J. A., Braun, M. T., & Kuljanin, G. (2016). Capturing the multilevel dynamics of emergence: Computational modeling,

simulation, and virtual experimentation. *Organizational Psychology Review, 6*(1), 3 33.
Kozlowski, S. W., & Klein, K. J. (2000). *A multilevel approach to theory and research in organizations: Contextual, temporal, and emergent processes*.
Larson, L., & DeChurch, L. A. (2020). Leading teams in the digital age: Four perspectives on technology and what they mean for leading teams. *The Leadership Quarterly, 31* (1), 101377.
Larson, N. L., McLarnon, M. J., & O Neill, T. A. (2020). Challenging the static quo: Trajectories of engagement in team processes toward a deadline. *Journal of Applied Psychology, 105*(10), 1145.
Leis, R., & Reinerman-Jones, L. (2015). Methodological implications of confederate use for experimentation in safety-critical domains. *Procedia Manufacturing, 3*, 1233 1240.
Lewis, M., Wang, H., Chien, S. Y., Velagapudi, P., Scerri, P., & Sycara, K. (2011). Process and performance in human-robot teams. *Journal of Cognitive Engineering and Decision Making, 5*(2), 186 208.
Lyons, J. B., Wynne, K. T., Mahoney, S., & Roebke, M. A. (2019). Trust and human-machine teaming: A qualitative study. In *Artificial intelligence for the internet of everything* (pp. 101 116). Academic Press.
Marks, M. A., Mathieu, J. E., & Zaccaro, S. J. (2001). A temporally based framework and taxonomy of team processes. *Academy of Management Review, 26*(3), 356 376.
Mathieu, J. E., Wolfson, M. A., & Park, S. (2018). The evolution of work team research since Hawthorne. *American Psychologist, 73*(4), 308.
McDonald, N., Schoenebeck, S., & Forte, A. (2019). Reliability and inter-rater reliability in qualitative research: Norms and guidelines for CSCW and HCI practice. *Proceedings of the ACM on Human-Computer Interaction, 3*(CSCW), 1 23.
McNeese, N. J., Demir, M., Cooke, N. J., & Myers, C. (2018). Teaming with a synthetic teammate: Insights into human-autonomy teaming. *Human Factors, 60*(2), 262 273.
Mckinney, Barker, J. R., Davis, K. J., & Smith, D. (2005). How swift starting action teams get off the ground: What United flight 232 and airline flight crews can tell us about team communication. *Management Communication Quarterly, 19*(2), 198 237.
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.
Mohammed, S., Ferzandi, L., & Hamilton, K. (2010). Metaphor no more: A 15-year review of the team mental model construct. *Journal of Management, 36*(4), 876 910.
Morgeson, F. P., & Hofmann, D. A. (1999). The structure and function of collective constructs: Implications for multilevel research and theory development. *Academy of Management Review, 24*(2), 249 265.
Myers, C., Ball, J., Cooke, N., Freiman, M., Caisse, M., Rodgers, S., McNeese, N. J. (2018). Autonomous intelligent agents for team training. *IEEE Intelligent Systems, 34* (2), 3 14.
Nass, C., Fogg, B. J., & Moon, Y. (1996). Can computers be teammates? *International Journal of Human-Computer Studies, 45*(6), 669 678.
Nass, C., Moon, Y., Fogg, B. J., Reeves, B., & Dryer, C. (1995). Can computer personalities be human personalities?. In *Conference companion on Human factors in computing systems* (pp. 228 229).
Nass, C., Steuer, J., Henriksen, L., & Dryer, D. C. (1994). Machines, social attributions, and ethopoeia: Performance assessments of computers subsequent to" self-" or" other-" evaluations. *International Journal of Human-Computer Studies, 40*(3), 543 559.
Nikolaidis, S., Lasota, P., Ramakrishnan, R., & Shah, J. (2015). Improved human robot team performance through cross-training, an approach inspired by human team training practices. *The International Journal of Robotics Research, 34*(14), 1711 1730.
Nomura, T., Kanda, T., Suzuki, T., & Kato, K. (2004). Psychology in human-robot communication: An attempt through investigation of negative attitudes and anxiety toward robots. In *RO-MAN 2004. 13th IEEE international workshop on robot and human interactive communication (IEEE catalog No. 04TH8759)* (pp. 35 40). IEEE.
O Neill, T. A., & Salas, E. (2018). Creating high performance teamwork in organizations. *Human Resource Management Review, 28*(4), 325 331.
O Neill, T. A., McNeese, N. J., Barron, A., & Schelble, B. G. (2020). *Human autonomy teaming: A review and analysis of the empirical literature*. Human Factors, Article 0018720820960865.
Rentsch, J. R., Heffner, T. S., & Duffy, L. T. (1994). What you know is what you get from experience: Team experience related to teamwork schemas. *Group & Organization Management, 19*(4), 450 474.
Rich, C., & Sidner, C. L. (1997). COLLAGEN: When agents collaborate with people. In *Proceedings of the first international conference on Autonomous agents* (pp. 284 291).
Rico, R., Gibson, C. B., Sánchez-Manzanares, M., & Clark, M. A. (2019). Building team effectiveness through adaptation: Team knowledge and implicit and explicit coordination. *Organizational Psychology Review, 9*(2 3), 71 98.
Rico, R., Sánchez-Manzanares, M., Gil, F., & Gibson, C. (2008). Team implicit coordination processes: A team knowledge based approach. *Academy of Management Review, 33*(1), 163 184.
Rosenberg, S. (1959). A laboratory approach to interpersonal aspects of team performance. *Ergonomics, 2*(4), 335 348.
Rousseau, D. M. (1985). *Issues of level in organizational research: Multi-level and cross-level perspectives. Research in organizational behavior*.
Rousseau, V., Aubé, C., & Savoie, A. (2006). Teamwork behaviors: A review and an integration of frameworks. *Small Group Research, 37*(5), 540 570.
Saavedra, R., Earley, P. C., & Van Dyne, L. (1993). Complex interdependence in task-performing groups. *Journal of Applied Psychology, 78*(1), 61.
Salas, E., Cooke, N. J., & Rosen, M. A. (2008). On teams, teamwork, and team performance: Discoveries and developments. *Human Factors, 50*(3), 540 547.
Salas, E., & Fiore, S. M. (2004). *Team cognition: Understanding the factors that drive process and performance*. American Psychological Association.
Salas, E., Sims, D. E., & Burke, C. S. (2005). Is there a big five in teamwork? *Small Group Research, 36*(5), 555 599.
Schelble, B. G., Flathmann, C., & McNeese, N. J. (2020). Towards meaningfully integrating human-autonomy teaming in applied settings. In *Proceedings of the 8th international conference on human-agent interaction* (pp. 149 156).
Shively, R. J., Lachter, J., Brandt, S. L., Matessa, M., Battiste, V., & Johnson, W. W. (2017). Why human-autonomy teaming?. In *International conference on applied human factors and ergonomics* (pp. 3 11). Cham: Springer.
Siegel, J., Dubrovsky, V., Kiesler, S., & McGuire, T. W. (1986). Group processes in computer-mediated communication. *Organizational Behavior and Human Decision Processes, 37*(2), 157 187.
Smith-Jentsch, K. A., Cannon-Bowers, J. A., Tannenbaum, S. I., & Salas, E. (2008). Guided team self-correction: Impacts on team mental models, processes, and effectiveness. *Small Group Research, 39*(3), 303 327.
Smith-Jentsch, K. A., Mathieu, J. E., & Kraiger, K. (2005). Investigating linear and interactive effects of shared mental models on safety and efficiency in a field setting. *Journal of Applied Psychology, 90*(3), 523.
Sundstrom, E. (1999). The challenges of supporting work team effectiveness. *Supporting Work Team Effectiveness, 3*, 23.
Van de Ven, A. H., Delbecq, A. L., & Koenig, R., Jr. (1976). Determinants of coordination modes within organizations. *American Sociological Review*, 322 338.
Walliser, J. C., Mead, P. R., & Shaw, T. H. (2017). The perception of teamwork with an autonomous agent enhances affect and performance outcomes. In *Proceedings of the human factors and ergonomics society annual meeting* (Vol. 61, pp. 231 235). Sage CA: Los Angeles, CA: SAGE Publications. No. 1.
White, C. H., Sargent, J., & Manusov, V. L. (2005). Researcher choices and practices in the study of nonverbal communication. *The Sourcebook of Nonverbal Measures: Going Beyond Words*, 3 21.
Wynne, K. T., & Lyons, J. B. (2018). An integrative model of autonomous agent teammate-likeness. *Theoretical Issues in Ergonomics Science, 19*(3), 353 374.
Zhang, R., McNeese, N. J., Freeman, G., & Musick, G. (2021). An ideal human" expectations of AI teammates in human-AI teaming. *Proceedings of the ACM on Human-Computer Interaction, 4*(CSCW3), 1 25.