12-2023

Leveraging Artificial Intelligence for Team Cognition in Human-AI Teams

Beau Schelble
bschelb@clemson.edu

Leveraging Artificial Intelligence for Team Cognition in Human-AI Teams

A Dissertation
Presented to
the Graduate School of
Clemson University

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy
Human-Centered Computing

by
Beau Gregory Schelble
December 2023

Accepted by:
Dr. Nathan McNeese, Committee Chair
Dr. Guo Freeman
Dr. Bart Knijnenburg
Dr. Richard Pak

Abstract

Advances in artificial intelligence (AI) technologies have enabled AI to be applied across a wide variety of new fields like cryptography, art, and data analysis. Several of these fields are social in nature, including decision-making and teaming, which introduces a new set of challenges for AI research. While each of these fields has its unique challenges, the area of human-AI teaming is beset with many that center around the expectations and abilities of AI teammates. One such challenge is understanding team cognition in these human-AI teams and AI teammates' ability to contribute towards, support, and encourage it. Team cognition is defined as any cognitive activity among the team members regarding their shared knowledge of the team and task, including concepts such as shared task or team mental models, team situation awareness, or schema similarity. Team cognition is fundamental to effective teams, as it is directly linked to the successful and efficient execution of team processes and subsequent team outcomes from decades of research in human-only teaming. However, the construct is challenging in human-AI teams given the significant differences in how we interact with humans versus AI (communication is a notable example). Despite the importance and knowledge of team cognition in human teams, there is little to no empirical research on the construct within human-AI teams. Without this research on human-AI teams, it is difficult to understand how the findings for human teams compare and apply to human-AI teams, leaving

practitioners of human-AI teams in the dark.

Very few studies directly examine what components of team cognition are most important to the human teammates working with AI or how including one or more AI teammates affects team cognition development. Without this knowledge, the ability to design better AI teammates that support and encourage the development of team cognition in human-AI teams will be a difficult task, and the need for extensive research in this area is apparent. As such, the current dissertation presents three studies that iterate upon one another to determine how AI teammates influence team cognition, what aspects of team cognition are most important to human-AI teams, and how AI teammate design may support contributions to team cognition.

The initial study of this dissertation utilized a mixed-methods approach to investigate how including AI teammates affects team cognition development in content, structure, and perception. This study also examines how human teammates' attitudes towards AI teammates change alongside those manipulations in team composition. Study 1 found that human-AI teams are similar to human-only teams in that team cognition develops over time; however, human-AI teams are different in that communication containing specific information related to the task and explicitly shared goals are more beneficial to developing team cognition. Additionally, human-AI teams trusted AI teammates less when working with only AI and no other humans, making AI contributions to team cognition difficult in teams with a majority AI composition. Perceived team cognition was also lower for AI teammates than human ones and had significantly inconsistent levels of team mental model similarity compared to human-only teams. These findings highlight the importance of information-sharing attributes of AI teammates to contribute to team cognition and drive the focus of the subsequent study.

Study 2 focused on how AI information-sharing attributes influence team cog-

nition and how human members of human-AI teams want their AI teammates to be designed to contribute to and encourage the growth of various aspects of team cognition. This study contains two sub-studies, with the first making use of a mixed factorial survey design and structural equation modeling to assess how participants in hypothetical human-AI teams respond to various information-sharing attributes used by an AI teammate. The second sub-study used interviews to ascertain how participants want their AI teammates to be designed to contribute towards and encourage team cognition. The interviews also investigate how information-sharing attributes by AI affect participants' attitudes towards their teammates, such as trust and cohesion, to ensure the contributions of the AI teammates are accepted. The results of Study 2 found that AI design features such as explainability and providing situational awareness updates on intra/extra team information changes had the most potent effect on participants' attitudes and perceived levels of team cognition. Additionally, the interview data characterizes the relationship between explainability and situational awareness, the heightened importance of situational awareness to human-AI teams, and the benefits of giving AI teammates defined roles with significant degrees of agency.

Lastly, Study 3 explored which AI teammate design features supported team situational awareness best and how their participation in team discussions affected their team cognition. Team situational awareness was chosen as the component of team cognition to influence based on the results of Studies 1 and 2, which highlighted how vital situational awareness at all levels was to human-AI teams. Study 3 found that AI teammates designed to augment team memory significantly improved participants' perception of a shared mental model with their human and AI teammates. This same AI SA attribute also enhanced the teams' situational awareness and likelihood of overcoming system failures that acted as situational awareness roadblocks.

AI participation in team discussions later in the teams' life cycle also enhanced team performance and situational awareness. Study 3's focus group interview data was also qualitatively analyzed, finding that the augmenting team memory SA attribute outperformed others by demonstrating to human teammates what information was necessary, when it was important and to whom it was essential, thereby enhancing teams' understanding of the task and building natural resiliency to roadblocks. These findings significantly inform the design of future AI teammates and future research by deepening the knowledge of how team cognition functions in human-AI teams.

These three studies contribute to three key research outcomes, including: 1) developing an understanding of what constructs within team cognition AI should support to drive effective team processes; 2) defining differences in team cognition between human-AI and human-only teams; and 3) how AI teammates meant to support team cognition can be designed and their effect on human-AI teams. Investigating these research gaps is necessary to develop effective AI designed to engage in highly social situations such as teams. Thus, to ensure this research is applicable, the three studies also synthesize their results into practical design recommendations that are actionable and supported by the empirical results of their respective study. As such, the research community and developers benefit from this work and help lead to more human-centered AI designs.

Dedication

If you know me, then you know this dissertation could be dedicated to nothing other than my family. The past decade I have worked towards this has been the most challenging thing I have ever done, as I have been tested in ways I never thought I would, but it has allowed me to deepen my relationship with "my people" significantly.

To my father, John Schelble, I dedicate this work to you for teaching me to love life and to love the people in it that matter most. You instilled in me an appreciation for taking pride in my work, lighting the spirit of adventure in me necessary to embark on this journey, never losing your parenting attitude no matter how old I became, and so much more. Your attitude toward family was integral to shaping how I view the concept of family today, and I could not be more thankful for that. Your sacrifices to give our family the best life gave me a better childhood than I could ever imagine. I will always strive to do the same for my future family and take the time away from work to enjoy life with them because of your example. Thank you for being my father. I can't wait to be just like you when I grow up, and I'm looking forward to our next adventure together.

To my mother, Julie Schelble, I dedicate this work to you for teaching me to always be curious. You have always been at my side through life's various trials and tribulations by taking my many stressed-out phone calls and reminding me to always keep my problems in perspective and have fun. You have taught me to always value doing the right thing and that making sacrifices for the ones we love is essential to a life well lived. I could not have asked for a better support system than you, as your unconditional love and encouragement have always been a strength and motivation for me to pursue my dreams. Thank you for being my mother, and thank you for supporting me throughout all of this experience. Now we move on to the next of many more "If I can just make it to..." milestones.

To my sister, Addy Schelble, I dedicate this work to you for teaching me what it means to be a friend. Addy, you have been my best friend for as long as I can remember; I cannot recall a day I did not want to be just like you, even now at 26 and 29. You have led your life with a tenacity that others could not fathom by accomplishing goal after goal in the face of obstacles even fewer would recognize. Watching you not only overcome the inequities brought on by life but doing so the right way by constantly standing up for your values has done a great deal to shape who I am today. Being your sidekick and growing up with you has been, and continues to be, a highlight of my life, and I cannot express to you enough how much I cherish having you in my life. Thank you for being my sister. You have defined me as a person and also taught me what it means to be a good friend.

To my sister-in-law, Alex Schelble, I dedicate this work to you for teaching me that kindness and determination are irreplaceable in this life. You are the definition of a good human being, and I want to thank you for being there for me no matter the circumstances. I have learned so much from you, and you have made me a better person by showing me that kindness is a necessity in this life, not an option, even when others may not be the most deserving of it. Your determination to do work that truly matters has impacted me significantly. I want to do meaningful work just like you do, and I want to do it with the same level of resolve that you do. I have never known anyone as dedicated as you, and if the world had more of you in the positions that matter, it would be a far better place. Thank you for being my sister-in-law (sister). I could not have asked for a better role model.

To my nephew, Archibald Gregory Schelble, I dedicate this work to you for being the sweetest boy in the world! Though you did not come until nearly the end of this project, you were a much-needed example of what pure curiosity, learning, happiness, and a love for life are, and you have helped me cross the finish line. I can't wait to watch you grow up and teach you all the nerdy things your moms think are lame!

To my entire extended family, I dedicate this work to you for teaching me that love should be unconditional. Thank you, everyone, for the years of support that came in the form of football tickets, dinner, a phone call, or just good company. To Jake Long, Greg Schelble, and Joseph Schelble, thank you for being a part of my life for the time you could. I have learned to treat every day with reverence and to leave this world a little better than the way I found it, as you all did.

To my family, I love you all.

Acknowledgements

Completing a Ph.D. is a unique journey that is best undertaken with others, and I could not have asked for a better fellowship than those with whom I began this journey. Christopher Flathmann, thank you for always being around to lend a helping hand, a friendly ear to listen to complaints or brainstorm with, and for being a good friend. I have thoroughly enjoyed growing through this process alongside you as we navigated these past four years. Rui Zhang, thank you for always being willing to take on new projects with me, acting as an excellent sounding board for ideas, and providing the most fun conversations. Geoff Musick, thank you for constantly reminding me that what matters in life is family and the simple things that bring us happiness.

To Lauren Smith, thank you for being the best partner I could ever ask for; you have given me so much valuable support, and I could not have made it through the finish line of this journey without you. I honestly did not think that there was anything left for me in Clemson as I neared the end of this nearly decade-long path, but I could not have been more pleasantly wrong. You make me a better person every day by showing me what true selflessness is, and your love for others is unparalleled. You have brought a love and appreciation for life that I did not think was possible, and I cannot wait to continue our adventure together.

To my adviser, Nathan McNeese, thank you for taking a chance on some random guy you met five years ago to help you get this lab off the ground. Thank you for giving me so many unique opportunities to learn our trade, whether writing grant projects, program reviews, or working on incredible research projects. Additionally, you have been an excellent boss, adviser, and friend. Nobody in the world could ask for a better adviser who sets them up to pursue whatever path they are passionate about achieving and genuinely wants the best for their students.

To my committee members, Richard Pak, Bart Knijnenburg, & Guo Freeman, I want to thank you for helping me develop as a researcher and as an individual, as each one of you set an incredible example in your personal and professional lives. The feedback and support you have given me throughout the development of this dissertation and projects outside of it have been invaluable to my growth, and I could not have completed this project without you.

To all members of the TRACE Lab, thank you for being a part of my journey; I couldn't have done this without your support. A special thank you is owed to Rohit Mallick, as this dissertation could not have existed without his support. Contributing all that time and energy was a tremendous undertaking at your own expense, and I genuinely appreciate it.

To all of my friends who have been with me throughout my childhood, teenage and adult lives, thank you for simply being yourselves. I have so enjoyed being able to continue being your friend. I love going back home to Chapin to catch up at the Tipsy Toad each Thanksgiving break to reminisce on old times and talk about the futures we envision for ourselves. Thank you also for keeping me grounded and reminding me that no matter what we all do in our adult lives, we're all still the same old kids who love ripping on each other, being together, and simply having fun with the lives given to us, especially as time seems to move ever faster the older we all become.

I will end the acknowledgments section with a quote written by Timothy Zahn

for the character Thrawn, as it represents a great deal of what I love about the loved ones we choose to share the pieces and parts of our lives with:

> *"A friend need not be kept within sight or within reach. A friend must be allowed the freedom to find and follow their own path. If one is fortunate, those paths will for a time join. But if paths separate, it is comforting to know that a friend still graces the universe with their skills, their viewpoint, and their presence. For if one is remembered by a friend, one is never truly gone."*

Contents

List of Tables

List of Figures

Chapter 1

Introduction and Overview of Dissertation

The current dissertation is motivated by several factors from applied practice and research. The following chapter covers the problem and research motivations for the current research and synthesizes them into dissertation-wide research questions (D-RQs), and finally, how the studies of this dissertation address them.

1.1 Problem Motivation

Teams are an essential aspect of society frequently utilized to accomplish work in various environments and tasks [268]. Because of their utility and generalizability, teams have also consistently leveraged the latest technologies to extend and enhance their operational capabilities, with virtual teams being a relevant and successful example [183]. Several shortcomings have been identified, and many critical questions regarding the dynamics of teamwork within human-AI teams remain unanswered as AI takes center stage as the latest technology to be utilized by teams [227, 212]. This adoption of AI as teammates has caused research on human-AI teaming to be published at an ever-increasing rate [238], which has highlighted that successfully joining humans and AI together in teams is not holistically dependent on AI's technical ability but also on a myriad of human factors that impact humans and AI's ability to collaborate [292, 238]. For instance, simply perceiving that a teammate is artificial has been linked to several adverse effects on team processes and outcomes [315, 224], including performance [68, 279]. Consequentially, human-AI and human-automation teaming literature has found that human-AI teams suffer when working on complex tasks compared to human-only teams [333, 90], highlighting how team outcomes suffer because of difficulties in effective coordination between humans and AI. This shortfall of AI teammates holds them back from widespread adoption, and a failure to rectify it further entrenches the negative perception that AI teammates are inadequate.

AI teammates currently deployed worldwide face significant challenges stemming from their difficulty engaging in teamwork-specific behaviors. These AI are being developed and deployed with significantly more agency [238, 212] despite not fully developing their ability to interact appropriately with human teammates. This practice is harmful to humans' perceptions and expectations for AI, especially in highly social contexts such as teams. While not a true teammate, Microsoft's Clippy is an excellent example of an artificial companion whose inability to interact with human users effectively rendered it useless [313]. Although AI teammates can engage in taskwork that contributes to positive team outcomes, unlocking the true potential of human-AI teams will require AI capabilities and design focusing on the innately human elements of teaming. Specifically, AI teammates must begin to integrate themselves within the highly social and complex role that is a teammate, as the responsibilities of a teammate go beyond those of a tool. A teammate must meet existing expectations for proficiency and growth in communication, coordination, and training, which are all essential to developing team-specific constructs such as team cognition [163, 29]. Team cognition is defined as any cognitive activity within a team referring to the team's shared knowledge of the team and task [52, 29]. It includes concepts such as shared mental models and team situational awareness [51]. Team cognition is a massively critical construct to team processes and outcomes, with decades of extensive research in human teams relating it to effective communication, coordination, and performance [230].

In general, among all teamwork constructs, team cognition is often viewed as one of the most important, as it is critical to enhancing team performance due to its practical benefits to effective team communication and coordination [221, 52, 30]. However, there is very little understanding of the mechanism and nature of how AI teammates may improve how they are perceived as teammates. Furthermore, how AI could be designed to better contribute to and support communication and coordination for team cognition development is unknown. Team cognition is directly related to the success and effectiveness of several team processes, which AI teammates currently struggle to develop, encourage, and sustain in human-AI teams [68, 279, 224], leading to the aforementioned challenges in coordination. As a result of these challenges, there is a cycle of maligned AI teammates that alter how team members interact with one another [68], harm trust in their AI *and* human teammates alike [211], and sees human teammates silo the AI away [224]. These problems lead to dysfunctional team cognition that leaves human-AI teams underperforming in complex tasks requiring high levels of coordination [68]. It is essential to confront these problems in human-AI teams before they are fully deployed and widespread in the workforce. Failing to confront these challenges will lead to dissatisfied teams that underperform and develop increasingly negative views of AI as teammates that may be difficult to repair moving forward.

The current standing of AI as teammates leaves two problems motivating the present dissertation. *The first revolves around how AI teammates interact with and are perceived by human teammates to understand why AI currently struggles to effectively engage with the human elements of teaming, such as team cognition. The second problem is understanding how human-AI teams can develop better coordination practices by improving AI teammates' contributions to team cognition.* This dissertation addresses these two challenges to the eventual successful deployment of human-AI teams by investigating how AI teammates affect team cognition, how human teammates perceive contributions to the team by AI teammates, and how purposeful AI design improves team cognition and coordination.

1.2 Research Motivation

While the problem motivations have been outlined, they cannot be effectively answered in practice without first addressing the research gaps driving the dissertation. Specifically, the role of technology in team cognition development and the current state of human-AI teaming are necessary to review to identify these gaps.

1.2.1 Team Cognition and Technology

Team cognition is essential in teaming because it enables and enhances team processes, directly leading to desirable team outcomes. As team cognition has been defined in detail in the previous problem motivation section, it will not be covered again here. Instead, additional context on the role of team cognition will be given in the form of team effectiveness models. Specifically, the Input-Process/Mediator-Outcome (IPO/IMO) model of team effectiveness is an excellent tool to highlight the relationships between team cognition and other team processes [195, 148]. Within the IMO model, team cognition is not a team process but an emergent state, which is defined as the properties of teams that are dynamic in nature, vary as a function of team contexts, and are the product of the team's collective experiences. Other emergent states include constructs such as trust, cohesion, and team confidence [195]. As an emergent state, team cognition has relationships with team processes, which are the classifications of team actions and interactions resulting in team outcomes [195, 148]. Detailing these relationships makes understanding why human-AI teams struggle and why it is essential to rectify these issues more readily apparent.

Human-AI teams' difficulty developing team cognition results from AI teammates disrupting traditional interaction patterns between team members. Team cognition is developed over time through experience engaging in continuous interaction while working towards a shared goal [163], making communication highly relevant to its development. While AI teammates struggle to engage in aspects of explicit communication [40], it is not the only thing holding these teams back from developing team cognition. As AI teammates fail to engage in critical team processes, whether they are in the transition or action phase, they are not meeting the expectations of a typical teammate. By not meeting these expectations, the traditional interaction patterns are disrupted, which means human-AI teams are not having the same experiences as human-only teams when attempting to develop team cognition [224]. Further still, there is a lack of understanding of what interaction patterns human members of these teams want AI to engage in and support as research has only recently explored the topic of desired AI teammate design [342]. Specifically, there is a need to understand which processes contribute most to supporting adequate team cognition within human-AI teams and which components of team cognition result in the most significant gains to the team's outcomes.

Human-AI teams represent a new interaction paradigm, and understanding how this new class of teammates should be designed and developed is paramount. AI and similar autonomous or automated technologies have constantly been developed to enhance and complement human abilities. However, with the advent of technologies that embody genuine autonomy, which can make decisions independently, engage in critical thinking [285], and even plan for future events [340], the human-machine relationship has been significantly altered. No longer is technology merely a tool, but a partner [238], which opens the door to an exceptional number of new opportunities but also several major challenges [227]. Many of which revolve around facets of team cognition support through AI design features, such as transparency [216] and explainability [324] to support situational awareness [85], in addition to designing to support shared mental models [90].

The integration of technology within teaming has a long history and has seen team efficiency and productivity increase, though not without some costs. Successful teaming is primarily based on interactions and coordination between individual team members and their ability to accomplish taskwork and teamwork effectively [268]. When new technologies are introduced into teams, many of these interactions can be altered in one way or another, causing the team to suffer, such as when virtual teaming was introduced [234]. These teams suffered from a lack of presence and awareness of their team members and their teammates' activities, making coordination and back-up behavior processes difficult [120, 116], which ultimately disrupted team cognition [156]. However, the problems facing virtual teams were not nearly sufficient to discourage their widespread adoption, especially given the advantages offered by such teams [235], and eventually, team cognition in virtual teams was enhanced and supported by technology. Specifically, the success of virtual teams was due in part to researchers' ability to understand how the technology was disrupting team cognition and, in turn, developing the strategies and technologies necessary to address the most impactful aspects of teaming for virtual teams (e.g., awareness) [116]. Human-AI teams also bring many incredible benefits alongside the challenges that must be overcome, just as virtual teams did.

The issue raised by the existing research on team cognition concerning human-AI teams relates to how team cognition influences the processes engaged in by human-AI teams. Given the nature of team cognition as an emergent state with several individual-related components within it, what components (i.e., team situational awareness, shared mental models) are most impactful to the ability of human-AI teams to achieve their processes effectively? This open question in the team cognition literature can be summed up with the following research gap:

> *An understanding of what aspects of team cognition AI should support to drive effective team processes does not exist.*

Furthermore, researching how AI teammates can be designed to support team cognition is critical to overcoming a variety of problems currently facing human-AI teaming research, as the existing literature has primarily focused on aspects of team performance directly instead of focusing on *how* teams perform [238]. Human-AI teams can benefit from similar careful design; however, there is very little research explicitly exploring how AI teammates can be designed to specifically support team cognition, which leaves the following research gap:

> *The outcomes of designing AI teammates to support team cognition have not been systematically explored.*

1.2.2 Human-AI Teaming

Human-AI teams have several advantages over traditional human-only teams that can be leveraged in various contexts. These human-AI teams are defined as teams with the following qualities: 1) one or more computer-based members and one or more human members; 2) the computer-based members occupy a specific team role that is interdependent with the other team roles to work towards a common goal; and 3) the computer-based team members are capable of taking action of their own accord [238]. These teams bring several advantages to the table by seeking to improve efficiency, effectiveness, and safety [83]. Many of these advantages are brought about by the computational nature of the AI teammate, which has access to often superior processing speed, bandwidth, memory, reliability, and information access [41]. However, these advantages brought about by the AI teammate also introduce stark differences in how the individual members of the team interact with one another as they have been given a teammate with vastly different capabilities, communication preferences, and interaction patterns [68]. These differences have led researchers to study how these new human-AI teams differ from existing knowledge on human-only teams.

The research on human-only teams is extensive, and several robust models of team effectiveness and processes exist that may or may not be helpful to human-AI teams. For example, the previously mentioned IPO/IMO model of team effectiveness was used to review empirical human-autonomy teaming literature in a recent article [238]. This review found a need to empirically test theories such as the IPO/IMO model of team effectiveness to determine the underlying mechanisms driving performance in human-AI teams. Namely, no comprehensive theories specific to human-AI teaming describe how emergent states such as trust, team cognition, or cohesion interact to drive team processes and their subsequent team outcomes [238]. While this is likely due to the novelty of human-AI teaming, this leaves the literature in a haphazard state of affairs with several disconnected studies strewn across several independent and dependent variables that are difficult to relate to one another cohesively. Without the ability to translate existing human-only theories to human-AI teaming, systematically moving the literature forward regarding the mechanisms behind performance in human-AI teams remains incredibly difficult.

The current state of human-AI teaming would benefit significantly from understanding how human-only teaming theories apply to human-AI teaming. Understanding how the previously mentioned concepts from human-only teaming research, such as team cognition, apply to human-AI teams would allow human-AI teaming researchers to understand and develop better AI teammates, improving the outcomes for human-AI teams. This has been a stated goal in human-AI teaming reviews [238] and has been applied to previous special teams such as virtual teams in the context of awareness [120, 121]. In the context of the current dissertation, how the existing model of team cognition applies to human-AI teams and, specifically, which aspects of it are most influential to human-AI teaming processes and outcomes is a critical question shown in the following formal research gap:

> *Our understanding of the potential differences in team cognition between human-AI and human-only teams is insufficient.*

1.3 Research Questions and Gaps

The current dissertation addresses several research questions. The research questions begin at a high level with a general understanding of how AI teammates influence team cognition. The research questions then become more specific to ask what components of the construct are essential to these teams and how AI teammates can be designed to support those components. These D-RQs seen in Table 1.1 serve as a foundation for discussing the three studies making up the dissertation. Answering these research questions improves team processes and outcomes for human-AI teams, applying them through a cohesive theory of team effectiveness and improving applied practices through specific empirically supported design recommendations.

D-RQ#	Research Question
D-RQ1	How do AI teammates influence the development and sustainment of team cognition in teams?
D-RQ2	What relationship do attitudes towards an AI teammate have on team cognition development?
D-RQ3	What effect do certain AI information-sharing attributes have on team cognition?
D-RQ4	How can an AI teammate be designed to contribute, support, and *encourage* team cognition development and sustainment in teams?

Table 1.1: Research Questions.

Additionally, the research questions come together to address the research gaps outlined above (see Table 1.2). Investigating these research gaps also answers the problem motivations, as each research gap is required to provide a comprehensive and informed solution to the problem motivations. Relating the dissertation to both problem and research gap motivations ensures that the findings and their discussion apply to the research literature and practitioners of human-AI teams.

Research Gap	Research Question
An understanding of what aspects of team cognition AI should support to drive effective team processes does not exist.	RQ1, RQ3
The outcomes of designing AI teammates to support team cognition have not been systematically explored.	RQ2, RQ3, RQ4
Our understanding of the potential differences in team cognition between human-AI and human-only teams is insufficient.	RQ1, RQ3, RQ4

Table 1.2: Dissertation Wide Research Questions and How They Address the Major Research Gaps.

1.4 Summary of Studies

The current dissertation consists of three sequential studies, with the scope of each focusing on more specific aspects of team cognition in human-AI teams using the information gained from the previous study (see Figure 1.1). Specifically, Study 1 takes an exploratory approach to examine the nature of team cognition in human-AI teams by manipulating team composition and using a mixed-methods approach. Study 2 goes on to study how various AI information-sharing attributes of an AI teammate affect attitudes and team cognition. Finally, Study 3 explores specific design features of AI teammates in human-AI teams to determine their efficacy in improving the state of team cognition and human teammates' ability to engage in behaviors beneficial to team cognition themselves. The following section summarizes these three studies in further detail and aligns them with the specific D-RQs each helps to address, shown in Table 1.3.

Study #	Short Study Title	Research Questions Addressed
1	Assessing the Influence of AI Teammates on Team Cognition in Human-AI Teams	RQ1, RQ2
2	The Pursuit of an Ideal AI Teammate for Team Cognition: Exploring the Impact of Information-Sharing Capabilities	RQ2, RQ3, RQ4
3	Towards Purposefully Designed AI Teammates for Team Cognition in Human-AI Teams	RQ1, RQ3, RQ4

Table 1.3: Studies That Address Each Research Question.

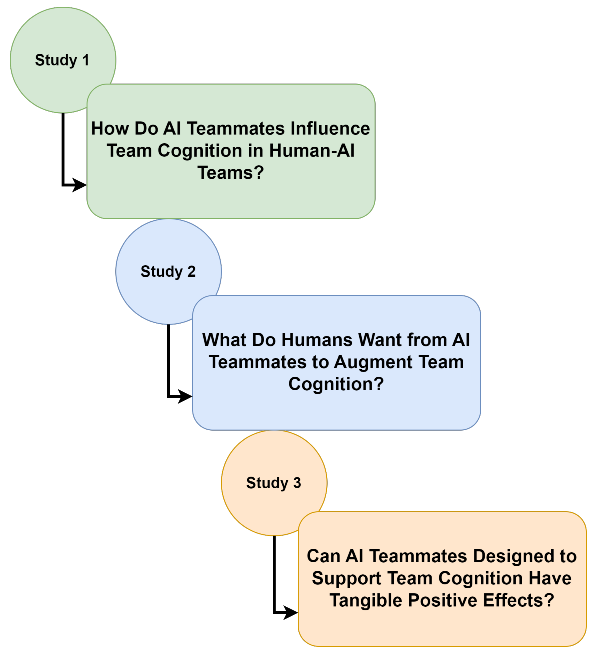


Figure 1.1: The Sequential Nature Of The Dissertation Studies.

1.4.1 Study 1: Assessing the Influence of AI Teammates on Team Cognition in Human-AI Teams

The initial study of the current dissertation took an exploratory approach to understand how the inclusion of an AI teammate affects team cognition. Specifically, this study explored the unique nature of team dynamics in human-AI teams compared to human-only teams and the impact of team composition on shared mental models, perceived team cognition, team performance, and trust, utilizing a mixed-method approach. The experiment included three team composition conditions (human-human-human, human-human-AI, human-AI-AI), completing a team-based emergency response simulation known as NeoCITIES and completing task/team mental model, trust, and perception measures. Study 1 found that human-AI teams were similar to human-only teams in the iterative development of team cognition and the importance of communication to accelerating its development; however, human-AI teams differed in that action-related communication and explicitly stated shared goals were beneficial to developing team cognition. Additionally, human-AI teams trusted agent teammates less when working with only AI and no other human teammates, perceived less team cognition with AI teammates than human ones, and had significantly inconsistent levels of team mental model similarity compared to human-only teams. This study contributes to the current dissertation through the act of advancing the existing research on human-AI teams by shedding light on the relationship between humans and AI operating in collaborative environments and characterizing the nature of team cognition development with AI teammates. Study 1 also contributed to the design and goals of Study 2 by advancing real-world design recommendations that focus on how information offered by AI teammates can contribute to team cognition.

1.4.2 Study 2: Examining the Impact of Information-Sharing by AI Teammates and the Ideal AI Teammate for Team Cognition

The second study of the current dissertation took the findings of Study 1 a step further by beginning to specify what aspects of team cognition are more emphasized and essential to human-AI teams. Study 2 accomplishes this by testing a series of AI teammate design features, specifically AI information-sharing, which is a crucial aspect of teaming processes and was emphasized in the qualitative portions of Study 1. A factorial survey was used in Study 2, which included a series of human-AI teaming vignettes to measure participants' perceived attitudes and team cognition in response to various AI information-sharing traits. Study 2 also conducted in-depth interviews to determine what those with experience in existing human-AI teams (competitive video gaming) desired from their AI teammates to best support team cognition. The results of these two sub-studies found that participants perceived attitudes and team cognition were influenced positively by all information-sharing conditions over the control condition. Specifically, the explainability, situational awareness updates of intra/extra team information changes, and augmenting team memory attributes performed the best (in that order). The qualitative interview found that human members of human-AI teams want personal, predictable, independent AI within a defined team role, capable of providing advice as an exemplar and providing actionable and relevant information regarding changes to the team and environment. These findings emphasize the importance of independent AI teammates who can relate their actions to changes in the shared environment and team through explainability. These two sub-studies provide a solid understanding of the various aspects of team cognition in human-AI teams and the perceived importance of each component to human-AI team members, with situational awareness prevailing as the consistent front-runner.

1.4.3 Study 3: Towards Purposefully Designed AI Teammates for Team Cognition in Human-AI Teams

Study 3 drives research toward improving team cognition through individual and team situational awareness in human-AI teams by examining two AI SA attributes selected from Study 2 in an extended hands-on teaming task. The design features derived from Study 1 and 2's results were augmenting team memory and sharing information changes within and outside the team, which were tested alongside a control. Study 3 also examined how the AI teammate's participation in a transition phase early or late in the team's life cycle influences the development of team cognition. The study found that augmenting team memory significantly outperformed the other AI SA attributes by improving participants' perceived shared mental model with the AI teammate, team situational awareness, likelihood of overcoming system failures, and information pushing verbal behaviors. Participation in later transition phases also significantly increased team performance and the probability of overcoming system failures. However, participation in the early transition phase did improve participants' perceived shared mental model with their human teammate when not working with the control SA attribute. Qualitative data from the study added context to these findings as teams shared that they wanted to understand what information was important and why it was important. As such, the augmenting team memory AI SA attribute demonstrated these facts of the task better by allowing human teammates to find and share the information necessary to overcome system failures. As such, Study 3 is the first to empirically examine specific design features meant to support team cognition development in human-AI teams and encourage the team to improve their team cognition behaviors in support of their team. In doing so, the study is a critical first step in understanding team cognition's role in team processes for human-AI teams, the importance of situational awareness to human-AI teams in particular, and how AI teammates may be designed to improve team cognition and, subsequently, human-AI teaming outcomes.

1.5 Conclusion

The need to address the research gaps surrounding the role of team cognition in human-AI team processes and outcomes grows increasingly relevant as the benefits of human-AI teams come closer to becoming a reality. Though existing research has investigated human-AI team performance and individual aspects of team cognition, the current dissertation is the first comprehensive examination of team cognition in human-AI teams. In particular, what aspects of team cognition are essential to team processes and outcomes, and how AI teammates may be designed to support and encourage these aspects to improve teaming outcomes. Study 1 was the first to examine the structure and content of shared mental models in human-AI teams and compare them to human-only teams. Furthermore, Study 2 was the first to explore how various information-sharing attributes, such as explainability and situational awareness of teammates' statuses, affected humans' perceived attitudes and team cognition. This study was also the first to conduct in-depth interviews with individuals who have worked on human-AI teams to determine what these human teammates want from their AI teammates to support team cognition and subsequent team processes. Finally, Study 3 was the first to empirically examine the efficacy of AI teammate design features to support and encourage team cognition in human-AI teams, specifically for individual and team situational awareness. Completing these studies contributes to a significantly improved understanding of team cognition in human-AI teams by exploring how its components influence specific team processes, what aspects of team cognition are most impactful for these teams, and how AI teammates may be designed to improve and encourage team cognition purposefully. The findings of the current dissertation are a significant first step for the research community in understanding the nature and role of team cognition in human-AI teams. These findings also convey a series of comprehensive design recommendations to help improve AI teammate design for developers and practitioners alike.

Chapter 2

Background

A review of the relevant and related literature that motivates and provides a foundation for this dissertation is necessary before directly addressing the studies encompassing it. Specifically, the current dissertation builds upon and addresses research gaps in the following domains: 1) team cognition; 2) technologically supporting team cognition; and 3) human-AI teaming. Team cognition (1) is necessary to review as it anchors the dissertation with a theoretical perspective that allows discussion of the various components of team cognition and their effects on teaming processes and outcomes. Technologically supporting team cognition (2) covers how integration between technology and teams can be accomplished, with AI teammates being the latest in a long line of adaptations in teaming. Lastly, human-AI teaming (3) provides the context that this dissertation is working to improve, centering the research and its contributions to a timely concept and detailing how they will benefit from thoughtful AI teammate design meant to support and encourage team cognition.

2.1 Team Cognition

Team cognition is a well-established concept in teaming research with a significant relationship to several team processes and outcomes. As previously stated, the concept of team cognition encompasses all cognitive activity among team members relating to aspects of the team or their task [52, 29, 163]. While there are several components within the larger concept of team cognition and more than one theoretical perspective, it is important to first review overarching theories of how teams accomplish work over time. Detailing the fundamentals of team effectiveness, such as action phase processes and transition phase processes, allows team cognition to be introduced and placed within these more extensive theories. Placing team cognition within these theories makes it easier to detail how team cognition affects teams and why it is essential to them. Thus, the following section discusses: 1) modeling teamwork theoretically; 2) the role of team cognition; and 3) the unique role of team situational awareness and related emergent states. These concepts are the theoretical grounding the current dissertation utilizes to develop the research goals and practical applications of the three research studies.

2.1.1 Modeling Teamwork Theoretically

Before discussing team cognition in further detail, it is necessary to ground this review within the wider Input-Mediator-Output-Input (IMOI) theory of teaming. This review of the IMOI model is essential as the current dissertation utilizes it to ground the discussion of team cognition's effect on team processes and outcomes. The IMOI model of teaming is an evolution of the Input-Process-Output (IPO) model popularized by the seminal works of Steiner, McGrath, and Hackman [206, 298, 207, 125]. The three components of the model are defined as follows: 1) inputs are factors antecedent to team member interactions that allow or constrain those interactions; 2) processes describe any team member interactions focused on accomplishing the shared task; and 3) outcomes are results and by-products of the team members actions that are of value to interested parties [200, 202].

While the IPO model itself helped advance research on team effectiveness considerably, it was not without its flaws. Namely, the IPO model could not properly capture the highly dynamic nature of teaming [148]. Specifically, the interactions among team members, their environments, and external factors significantly influence teams in a way that is not easily captured by cause and effect models [148, 205]. As such, many mediating constructs conveying the influence of inputs to eventual outputs, such as cohesion, trust, and shared mental models, do not qualify as processes [195]. Instead, Marks and colleagues identified these constructs as emergent states, which are "...properties of the team that are typically dynamic in nature and vary as a function of team context, inputs, processes and outcomes" [195]-pg. 357. While these emergent states do not necessarily represent team interaction or actions that result in outcomes, they are products of the team members' experiences. They can contribute as inputs themselves over time [195] and can be best thought of as cognitive, affective, or motivational conditions [200]. While the IPO model did not do enough to capture these emergent states' effect on team outcomes, it was suitable for capturing the temporal aspects of teaming, which was critical to defining transition and action phases along with their respective processes [195].

The IPO and subsequent IMOI model characterize the episodic nature of teaming where teams experience two distinct phases in accomplishing their shared goal, the transition phase and the action phase [195]. Transition phases see teams focus on evaluation of past action phases and or planning to help guide members towards the shared goal [195], which includes processes such as mission analysis [99, 250], goal specification [72, 251], and strategy formulation [107, 250]. Alternatively, action phases are the phases when teams are actively engaged in acts that directly contribute to accomplishing the team's overall goal [195]; such acts include monitoring progress towards goals [149], team monitoring, back-up behaviors [72, 168], and coordination behaviors [70, 66]. Each transition phase and action phase has its own IMOI sequence and can be seen through a temporal lens, first posited by Marks and colleagues [195]. Teams enter a transition period with distinct inputs, processes, and outputs, then move into an action phase with *its* own inputs, processes, and outputs. This procedure can then be repeated if the team remains together and goes on to complete more shared goals. The result of all this theoretical conjecture is a framework for team effectiveness that accounts for the temporal and dynamic nature of teaming. The framework also simultaneously represents team processes and emergent states throughout all team phases, as shown in Figure 2.1, and has also held up very well when examined empirically [200]. Defining this model in detail enables the research questions of the current dissertation to be operationalized better by showcasing the theoretical relationships between each of the primary dependent variables examined. However, the role of team cognition within this framework as an emergent state is essential to outline in detail to convey the importance and relevance of the current dissertation to theoretical and applied understanding of human-AI teams.

2.1.2 Team Cognition's Role in Teaming

There are two major perspectives of team cognition: 1) the shared knowledge approach [29]; and 2) the interactive team cognition approach [52]. First, the shared knowledge approach is best described using the construct of shared mental models as an example, which was one of the first operationalizations of team cognition in

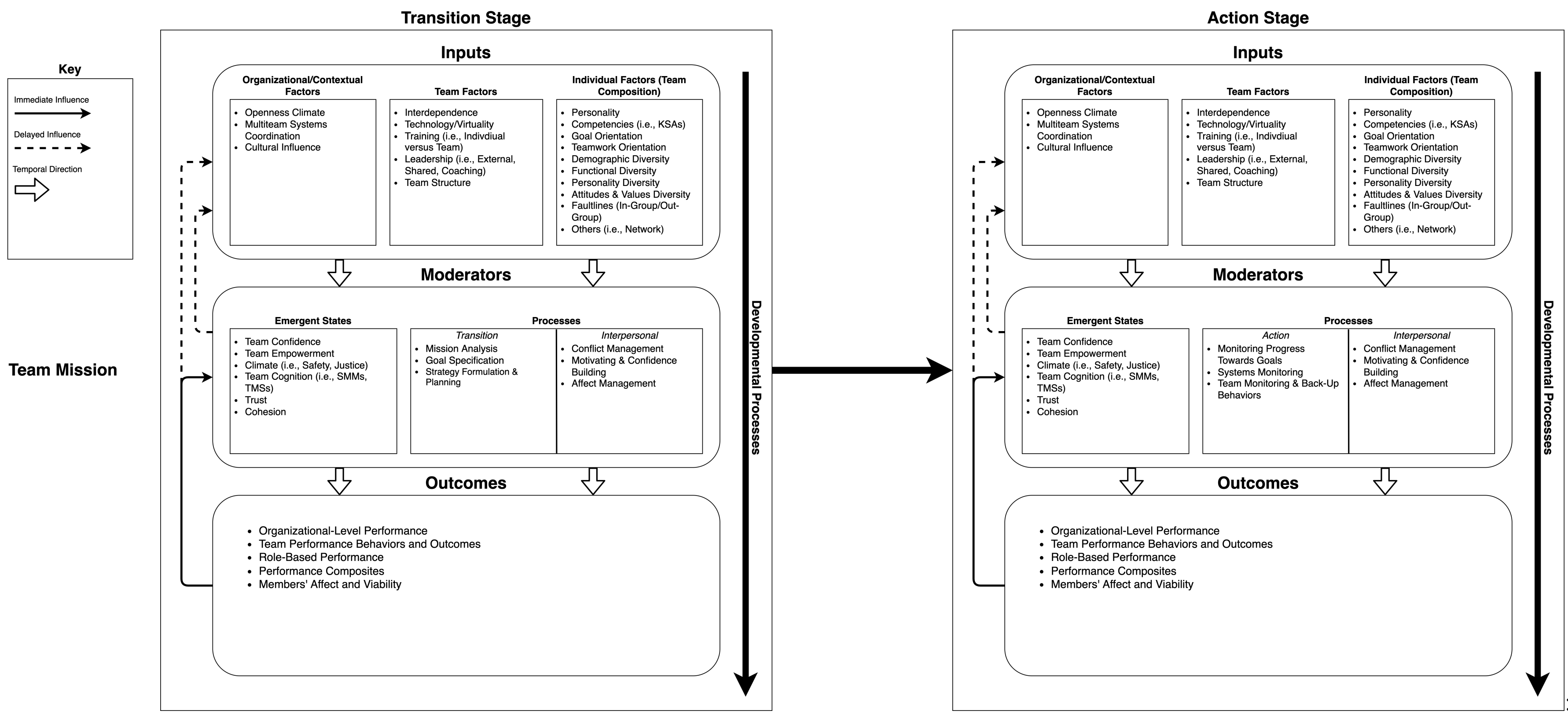


Figure 2.1: IMOI Model of Team Effectiveness

research [28, 29, 163, 166]. Firstly, mental models represent organized knowledge structures that individuals hold to describe aspects of their environment to interact with it [202, 328]. Specifically, mental models allow them to explain the current state of the world and predict future states [263, 152], which also highlights its similarities to situational awareness and the relationship between these cognitive processes. In 1993, Cannon-Bowers and colleagues posited that teams collectively draw upon their mental models in the form of shared mental models to efficiently and effectively respond to dynamic teaming environments [29]. Teams use shared mental models to make decisions on actions that maintain consistency and coordination with their teammates [29]. Teams pull from four distinct mental model types: technology/equipment, job/task, team interaction, and team shared mental models [29]. However, over time, research has primarily focused on team and task shared mental models by collapsing the task/technology models into a single task shared mental model and the interaction/team models into a single team shared mental model [29, 221]. Other representations of team cognition in the shared knowledge perspective also exist in the form of constructs such as the transactive memory system. Transactive memory systems represent the division of cognitive labor for storage, retrieval, communication, and encoding of information across knowledge domains within teams and groups [21, 319, 318]. There is also the construct of team member schema similarity, which is the degree to which team members have compatible or similar structures for understanding and organizing team-related knowledge and phenomenons [258, 259]. These examples are not exhaustive but are meant to demonstrate the breadth of cognitive activity occurring at the team level, making team cognition function as an umbrella term encompassing similar but distinct, team-level cognitive constructs [53, 52, 58], such as team situational awareness [16], transactive memory systems [319], and shared mental models [29]. However, team cognition has not been dominated solely by the shared knowledge approach. As time went on, the theory of interactive team cognition was proposed.

While team cognition as a concept was and is still, to an extent, predominately addressed by the shared knowledge perspective, a different perspective arose after years of empirical research. Cooke and colleagues developed the interactive team cognition perspective after recognizing that while the extent of shared knowledge held by team members was predictive of team performance, it leveled off after only a single action phase [54]. Meanwhile, the team's processes (communication), which directly led to performance outcomes, continued to improve over time and were even more predictive of team performance [52, 53, 48]. While the shared knowledge perspective was distinguished by the concept that team cognition centers around the patterns of knowledge held among team members, the interactive approach took an ecological perspective. It defined itself by asserting that team cognition is a team-level cognitive activity inextricably tied to *context* [52]. Specifically, interactive team cognition made three central claims to distinguish itself from the shared knowledge approach: 1) team cognition is not a property of individual team members or the products produced by a team; 2) team cognition should be investigated at the team level; and 3) team cognition is tied to the context teams are working within as it is shaped and developed by the specific needs and constraints of their context [52]. Interactive team cognition identifies those processes (i.e., information sharing between team members) as team cognition [52]. Interactive team cognition is typically measured by analyzing teams' communication patterns, specifically, "push" and "pull" communications that represent the pushing of task-relevant information or the requesting of it [48]. The level of team cognition a team has in this analysis depends on the amount of "push" and "pull" communication a team engages in, the accuracy of those communications, and the ability of team members to anticipate information needs [113, 48].

This type of measurement contrasts the elicitation techniques the shared knowledge perspective utilizes, which are not as dynamic and do not capture team processes in context. Specifically, the elicitation strategies for the major constructs in team cognition can be found in Table 2.1, along with their definitions and general relationships to teaming outcomes, as they are too numerous to expound upon here. However, regarding the current dissertation's choice of perspective, a holistic approach will be taken. Both views will be leveraged as team cognition is extensively studied within a new team type (human-AI teams) for the first time [69]. Additionally, utilizing both perspectives can lead to new insights that can benefit shared knowledge and interactive team cognition [209, 69]. Regardless of the stance taken, however, team cognition has well-known relationships to teaming outcomes such as performance [202, 62, 195, 218, 300, 230, 113].

The role of team cognition in team outcomes is well-known through decades of empirical research. The importance of understanding team cognition cannot be overstated, with multiple prominent examples of catastrophic failures directly related to breakdowns in team cognition. For example, the USS Vincennes accidental shoot-down of an Iranian passenger aircraft (Iran Air Flight 655) [46], the actions and decisions of NASA leading up to the Challenger shuttle disaster [312], and the delayed disaster response to Hurricane Katrina [175]. Empirical research on team cognition has linked it to performance outcomes for several decades, such as in the case of shared mental models [202, 300], team situational awareness [67, 155, 270], and team member schema similarity [257]. Specifically, team cognition significantly enhances critical aspects of teaming: 1) team behavioral processes in the transition and action phases [62, 202, 300, 176]; 2) motivational states such as trust, conflict, and satisfaction [62, 30, 244, 254]; and 3) team performance [202, 253]. However, the relationship between team cognition and performance is more robust with subjective

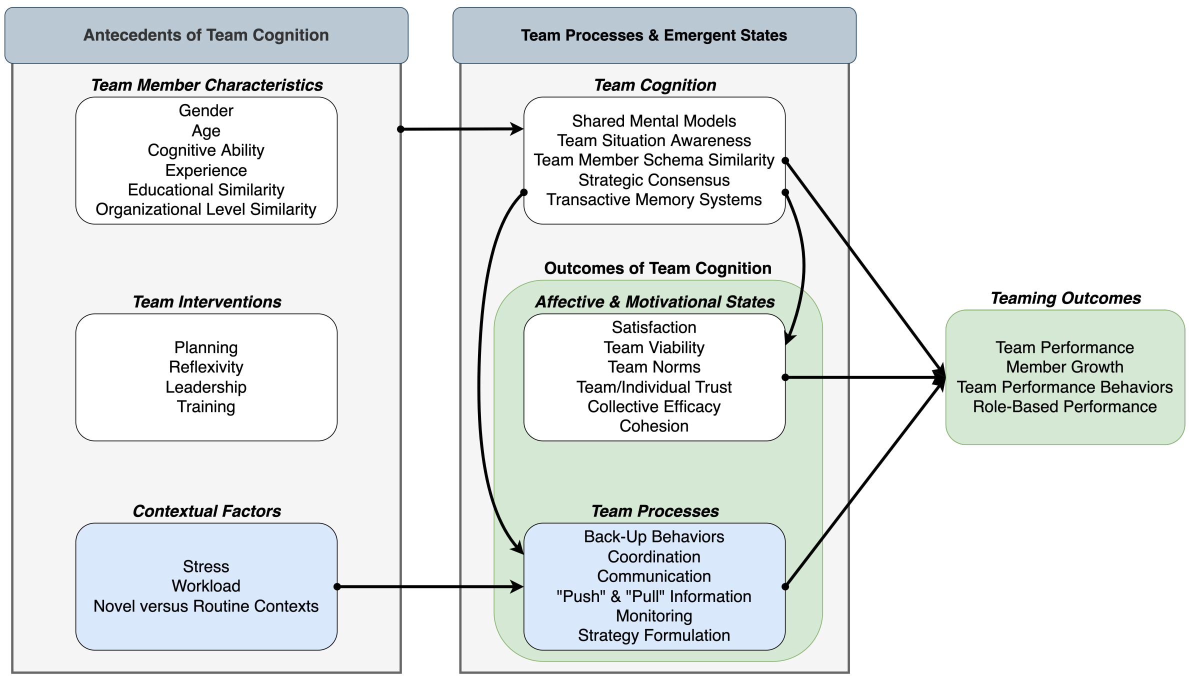


Figure 2.2: The Role of Team Cognition in Teaming Beginning with its Antecedent Factors and Alongside Other Emergent Team States, Team Processes, and Their Eventual Influence on Team Outcomes. Components of Interactive Team Cognition are Shown in Blue, with Outcomes Displayed in Green. **This figure is a derivation of work completed by DeChurch & Mesmer-Magnus [62], Mohammed et al. [221], and Salas et al. [268].*

measures of team performance than objective measures [62] (see Figure 2.2). These benefits can also be continually refined to be more effective, enhancing team outcomes such as objective performance, communication, coordination, and strategy [161, 48]. It is important to note, too, that team cognition is positively related to team performance as a solitary construct when controlling for the effect of behavioral processes and motivational states on team performance, which a meta-analysis conducted by DeChurch and Mesmer-Magnus confirmed [62]. Simply put, team cognition enhances the ability of teams to effectively understand their shared environment, correctly interpret current team needs, anticipate future needs using appropriate communication strategies, and competently adapt to and address dynamic environments in tandem with other team members. However, as this research shows, team cognition does not exist alone within teaming; it must co-exist and interact with other crucial emergent team states such as trust and cohesion [268].

Table 2.1: Constructs of Team Cognition Along with Their Respective Definition, General Relationships with Teaming Outcomes, Their Theoretical Perspective, and Elicitation Strategies.

Concept	*Definition*	*Relationships to Teaming Outcomes*	*Elicitation Strategies*	*Theoretical Perspective*

Shared Mental Model	Organized mental representations of the various component pieces relevant to a team's overall task [163].	Team mental model similarity and task mental model similarity are positively related to team performance [202]; Task mental model similarity's relationship with team performance is mediated by team processes, which themselves were moderated by team mental model similarity [201]	Paired Sentence Comparisons [19, 202]; Concept Mapping [232, 264]; Card Sorting [60, 169]; Qualitative Techniques [311, 255]	Shared Knowledge
Transactive Memory System	The division of cognitive labor for storage, retrieval, communication, and encoding of information across knowledge domains within teams and groups and the shared understanding of which team members know what and how to access it [21, 319, 318].	Positive relationship to goal performance and both internal and external evaluations [6]; Positive relationship to affective outcomes such as satisfaction [343]	Collective Recall Measures [143]; Observational Methods [249]; Questionnaires [177]	Shared Knowledge

Team Situational Awareness	"...the union of responsibility (i.e., overlapping plus complementary knowledge) for maintaining awareness of a dynamic environment is hypothesized to be based on the team's shared understanding of a specific situation," [16].	Positive relationship with team performance, training quality, decision-making, communication and coordination [51]; Shared situational awareness is positively associated with shared knowledge [271]	Surveys [303, 81], Observer Ratings [14], Indirect Performance Measures [213], Eye Tracking [133]	Shared Knowledge
Team Member Schema Similarity	The degree to which team members have compatible or similar structures for understanding and organizing team-related knowledge and phenomenon's [258, 257].	Has a positive relationship to team performance [257]; Positive relationship to affective outcomes such as viability [259]	Survey [257], x Absolute Value [257]	Shared Knowledge

Strategic Consensus	Shared understanding of strategic priorities [158].	Positive relationship with team performance, such that high-performing teams began with low levels of strategic consensus but developed high levels over time [161]	Survey [144], Qualitative Techniques [17]	Shared Knowledge
Interactive Team Cognition	The processes teams engage in to accomplish their shared goal (i.e., information sharing between team members) bound by context [53, 52].	Team processes ("push" and "pull" communications), and the accuracy of those communications have a positive relationship with team performance [48]	Communication Analysis [71, 113]	Interactive Team Cognition

2.1.3 The Unique Role of Team Situational Awareness and Related Emergent States

Similar to a good recipe, developing adequate team cognition requires a few key ingredients. Factors such as the amount of time spent working with teammates [48] and the type of training [197, 5] play a role in team cognition development; however, when it comes to other emergent states, trust and cohesion play a significant role [230, 247]. These emergent states are essential for teams to develop because they directly influence the nature and development of team cognition [94, 93]. This is because team cognition relies heavily on the shared attitudes among teams, such as trust and cohesion, which even a single team member can disrupt [159]. When cognition becomes interdependent, as they do in highly coordinated teaming [147, 269], issues associated with attitudes towards teammates become particularly relevant to the effectiveness of that shared cognition [93]. These attitudes towards other teammates (i.e., trust in another teammate or the team itself) become particularly critical because early interactions between teammates are pivotal in developing team cognition, norms, and team process development [93]. However, the importance of attitudes such as trust and cohesion for human-AI teams is particularly relevant. The heightened importance is warranted based on humans' generally negative attitudes towards AI as a teammate [224, 279, 252]. As a result of these negative attitudes, human teammates are less likely to communicate effectively and coordinate with their AI teammates [68], disrupting team cognition for the entire team. Because of this, human-AI teams must develop appropriate levels of trust and cohesion with their AI teammates so they are willing to accept their contributions to team cognition. Team situational awareness is a prime example of an area of team cognition that AI teammates can significantly contribute towards that human teammates must be willing to and comfortable with accepting through adequate cohesion and trust.

2.1.3.1 Team Situational Awareness

The concept of team situational awareness [16] stems from shared situational awareness [78], which are both derivations of individual situational awareness [330, 82, 80]. Defining individual situational awareness first, it is essentially "knowing what is going on," [79, 81] with a formal definition put forth by Endsley as "...the perception of the elements in the environment within a volume of time and space, the comprehension of their meaning, and the projection of their status in the near future," [80] p.97. This perception can be broken down into three primary levels, where Level 1 situational awareness is the perception of the elements within the environment at a fundamental level (i.e., warning lights, other cars, road hazards, speed) [84, 82]. Level 2 situational awareness focuses on the comprehension of the individual's current situation, going beyond Level 1 to combine all the seemingly disjointed elements perceived in Level 1 to develop a holistic understanding of each element's importance concerning the individual's goals [84, 82]. Finally, Level 3 situational awareness builds upon Level 1 and Level 2 situational awareness to use the individual's knowledge of the elements (Level 1) and understanding of the situation (Level 2) to project the possible future state of the elements within their environment and how they may affect their environment [82, 84]. Developing and maintaining situational awareness takes a degree of mental effort [246]. However, it is not without its benefits as it is positively related to performance across several empirical studies [76, 191, 157]. Namely, pilots' lack of situational awareness was found to be directly responsible for nearly all 200 aviation accidents reviewed [76]. Additionally, drivers told to use cell phones while completing an obstacle course were found to have significantly reduced situational awareness and committed more driving infractions regardless of experience level [157]. However, when it comes to situational awareness in teams, it is more than simply combining individuals' situational awareness [287]. Team situational awareness is defined by a common perspective between team members with respect to their shared environment, its meaning, and possible future states, all while considering the union of responsibility regarding who is responsible for each aspect of their shared context [321, 16]. As such, team situational awareness is highly dependent on good team processes (i.e., monitoring, back-up behaviors) and the nature of the task and team [270], which themselves rely *heavily* on team communication and coordination [270, 213]. Team situational awareness is a prime component for AI teammates to significantly contribute to team cognition in both its development and maintenance due to AI's technical advantages to monitor and provide timely feedback in an incredibly accurate manner [285]. There is already solid evidence for AI teammates' role in situational awareness with a solid precedent set by past automated systems in collision avoidance [189]. But if human teammates are going to accept these contributions from an AI teammate, an adequate level of trust and cohesion must first be developed with the AI.

2.1.3.2 Trust

Trust between team members is an essential emergent state in teaming related to several team outcomes and is an antecedent to team cognition development. The construct of trust is multifaceted but can generally be defined as "a willingness of a party to be vulnerable to the actions of another party based on the expectation that the other will perform a particular action important to the trustor, irrespective of the ability to monitor or control that party" [204] p. 712. This definition of trust applies to both teams and groups; however, when it comes to trust between humans and technology, the conceptualization set forth by Lee & See's 2004 work is most frequently referenced: "...the attitude that an agent will help achieve an individual's goals in a situation characterized by uncertainty and vulnerability," [173]. This definition is slightly more nuanced to the trustor's development of trust in the system, explicitly focusing on the effect of system reliability. This focus on system reliability thus brings the topic of trust calibration to the forefront. If a system is reliable and a trustor does not trust it, or vice versa, then the user's trust is not appropriately calibrated, and efforts should be made to rectify the discrepancy less the user rely too heavily on the system or under-utilize it [173, 172, 223]. The rise of autonomous systems advancing from automated systems has necessitated an evolution of Lee & See's 2004 definition, which now recognizes the increased role of relationships between humans and AI as they move from tool to teammate [38]. While trust and its development is a diffuse topic, the effect of trust on team processes and outcomes is well known, as high trust within a team is an essential component of building team cognition [63, 92], and antecedents of team cognition such as cohesion [192]. In turn, these formative team constructs lead to reflective (behavioral) markers of trust and allow teams to achieve a high level of performance [192]. The importance of trust in teaming is also not limited to human teams but also to human-machine teams, as recent empirical research on human-automation [141, 274], human-robot [61, 129], and human-AI teams [188, 190, 210, 211] have demonstrated significant positive relationships between trust and team performance outcomes. Further still, recent research on the desired traits of an AI teammate has highlighted the importance of trust to the eventual acceptance of AI teammates [342], and human-robot teaming research has long used trust as an indicator of effective shared mental models [302]. As such, trust is a critical antecedent to the development of team cognition and the acceptance of all teammate types, making it a particular construct of interest for any research on team cognition in human-AI teams alongside other emergent states.

2.1.3.3 Team Cohesion

Team cohesion is another emergent state with the potential to encourage or discourage the development and maintenance of team cognition. Defined as "a dynamic process that is reflected in the tendency of a group to stick together and remain united in the pursuit of its instrumental objectives and/or for the satisfaction of member affective needs," [32] p.213. Team cohesion is generally perceived as team members' commitment to a team's overall task and to their fellow team members [110, 268]. Team cohesion is another human factor of teaming that can serve as an antecedent to team cognition by positively or negatively influencing how team members communicate and coordinate with one another. Specifically, negative attitudes towards one or more team members can disrupt team processes, leading to poor team cognition development [159, 93]. Through this mechanism and others, team cohesion is significantly related to eventual team outcomes such as performance, directly [219] and as a mediator [10]. Meta-analytic studies have investigated these relationships in detail and found a positive correlation between team cohesion and team performance behaviors and outcomes, with cohesion being more strongly correlated with team performance behaviors than performance outcomes [12]. Cohesion is a well-recognized topic in human-AI teaming, being discussed as an essential attitude for human-AI teams to develop conceptually [98, 134], and has been attributed to low performance [283]. However, despite this relevance, there is very little empirical research on cohesion in human-AI teams, considering how adding AI teammates influences cohesion directly. As a critical affective emergent team state related to team cognition development, it is essential to include team attitudes such as cohesion in measuring and evaluating team cognition, especially when unique teammates such as AI are involved.

In summary, the research reviewed above highlights the dynamic nature of teaming through the IMOI team effectiveness model. Specifically, the nature and role of team cognition within that model, the unique role of affective emergent states such as team cohesion and trust as antecedents of team cognition, and their importance to the acceptance of AI teammates using team situational awareness as an example. However, developing an AI teammate that can be accepted and contribute meaningfully to team cognition is far easier said than done, as new teaming technologies have frequently disrupted team cognition development.

2.2 Technologically Supporting Team Cognition

Introducing new technologies that alter team interaction patterns can be particularly jarring to developing team cognition; however, those same technologies can be designed to mitigate adverse effects and even improve overall team cognition. Specifically, the mechanism through which team cognition is affected by these technologies is a result of the highly interactive nature of team cognition's development via team processes such as monitoring, back-up behaviors, and strategy formation. All of those team processes require elaborate communication and coordination of interdependent teammates and assets, which team cognition directly supports [115, 156, 198]. This phenomenon is important to supporting team cognition in human-AI teaming as the introduction of AI teammates represents a significant shift in interaction patterns and communication practices for many individuals compared to a typical human teammate [342]. The following concepts will be reviewed: 1) challenges to team cognition from technology; 2) supporting team cognition development and support; and 3) utilizing AI teammates for team cognition development and support.

2.2.1 Challenges to Team Cognition from Technology

New technologies, such as AI teammates that alter interaction patterns involving communication, coordination, and adoption pose several challenges to team cognition development. While new technologies can be a challenge for team cognition, they are not without precedent, as virtual teams imposed similar constraints on teams when they were first introduced [234, 235]. Virtual teams came along with the widespread adoption of the internet in the 1990s, which enabled a new level of globalization for the world. Virtual teams encouraged individuals and companies to meet and work with other individuals miles away from each other with relative ease [308].

While new technologies that enabled virtual teams came with a host of benefits, they also came with a host of downsides, as the aptly titled work by Olson & Olson in 2000 showed that *Distance Matters* [234]. The specific challenges faced by virtual teams included being constrained by the limitations of their locations and technologies, a breakdown in the development of common ground (a form of team cognition), difficulties collaborating in real-time with one another on tightly coupled work (i.e., work requiring a high degree of communication and coordination [234]), the willingness of individuals to collaborate with one another personally, and the willingness of individuals and organizations to adopt new and more efficient collaboration technologies [234]. All of these factors played a significant role in the success or failure of virtual teams and placed limitations on the success of these new team types.

Practitioners of virtual teams had difficulty overcoming the limitations of their contexts and the technologies they used. These distributed teams faced challenges developing team cognition when multiple team members were spread across various time zones (geographic distance) [199, 248]. With multiple team members not online at the same time as others, there is a subsequent reduction in the number of team interactions necessary to build team cognition [52, 29, 163]. In addition to interacting less, many team members may have been part of a separate organization, reducing the amount of shared knowledge team members had from the start [199, 248]. These issues did not stop at the organizational or technical level either, with research finding that individuals were likelier to deceive, be less amenable to, and initially cooperate less with a physically distant teammate [20]. While individuals' attitudes towards a distant teammate improved quickly with increased interaction, this finding emphasizes the importance of trust and cohesion to teaming, team cognition, and teammate acceptance, especially in unique team types such as virtual and human-AI teams. The lack of common ground between teammates was a consistent theme for virtual teams, as it was difficult for team members to converge on shared assumptions and knowledge [44]. A similar construct to team cognition, the construct of common ground represents the shared knowledge and assumptions individuals have constructed together over time based on their prior interactions and knowledge with one another, with communication between the two individuals playing an essential role in building and making corrections to that common ground [43, 42, 297]. The issue of common ground is relevant to team cognition as it was identified as being especially difficult to build using technology-mediated communication [234]. This difficulty in technology-mediated communication hampers the development of team cognition [156] and makes tasks requiring high levels of coordination especially difficult [234]. It is important to discuss these early challenges faced by virtual and distributed teams because they share distinct similarities to the obstacles currently facing human-AI teams.

The challenges waiting to be overcome by researchers, developers, and practitioners of human-AI teams are numerous but not without precedent. Virtual and human-AI teams face limitations in teammate attitudes, coordination and communication abilities, and even the technologies' capabilities. Specifically, human-AI teams also face problems with teammates' attitudes towards their artificial teammates, communicating with them less frequently and giving them a lower affective rating [315]. Human-AI teams have difficulty developing appropriate "push" and "pull" strategies, indicating a problem anticipating teammate needs, a crucial aspect of adequate team cognition [68]. Humans have also been shown to perform at a lower level when they even perceive one of their teammates as artificial instead of a human, despite the true identity of the artificial teammate being an equally performing confederate researcher [279]. These studies highlight the similarities in virtual and human-AI teams' attitudes and ability to communicate and coordinate. AI teammates' ability to adapt to dynamic environments and effectively communicate within those environments is a major factor in the challenges brought about by AI teammates. Natural language processing in AI has been advancing rapidly. However, it still cannot equal human speech [40, 341], which is a common limitation frustrating current members of human-AI teams in eSports [342]. However, virtual team researchers and developers have successfully addressed or mitigated several similar roadblocks either entirely or by developing collaborative technologies that play on virtual teams' inherent *strengths*. In a follow-up paper to Olson & Olson's 2000 work [234], Bjørn and colleagues' 2014 paper entitled with the tagline *Does Distance Still Matter?* revisited the problems highlighted by the earlier paper [15]. Bjørn and colleagues found that with improved collaborative technology and improved adoption of these technologies, virtual teams no longer suffer from the inability to engage in highly coupled work. The researchers also found that specific fields, such as software development, have leveraged the strengths of virtual teams and their technologies (video meetings, instant messaging, document sharing) to adopt the concept entirely, significantly increasing their willingness to collaborate and willingness to adopt collaboration technologies [15]. Current research shows that members of global virtual teams now have high levels of innovation, engagement, satisfaction, and a positive outlook on their work's complexity [231]. While there are still hurdles to overcome for virtual teams, as the recent COVID-19 pandemic has shown, thoughtful design, research, and application can do much to overcome and mitigate these problems, giving human-AI teaming practitioners an example for overcoming their challenges.

2.2.2 Supporting Team Cognition Development and Sustainment Through Technology

Supporting the development and sustainment of team cognition using technology is possible if the right aspects of teaming are targeted. Again, the example of virtual teams can be brought up to exemplify the opportunities for technology to contribute to team cognition positively. To reiterate the points above, team cognition in virtual teams suffers from a lack of high-quality task-oriented communication [177, 198], difficulty coordinating and sharing knowledge among team members in different organizations and time zones [199, 248, 198], and a reduced capacity for fluid communication [235]. For virtual teams, this means collaborative technologies had to focus a great deal of design effort on supporting and enhancing distributed teams' ability to communicate task-related information in a fluid and efficient manner. Technology also had to be designed to allow team members to quickly and easily build and coordinate their shared knowledge through enhanced awareness of their teammates. From this effort came the term "groupware," which was defined as any "computer-based systems that support groups of people engaged in a common task (or goal) and that provide an interface to a shared environment," [77]. Groupware stemmed from the field known as Computer-Supported Cooperative Work (CSCW), which arose in the 1980s and 1990s to target research questions that arise in and from the development of a wide range of different computing technologies aimed at teams and groups [282]. In Ellis and colleagues' landmark paper on groupware, they defined the time-space taxonomy. This taxonomy separated groupware technologies into four distinct categories depending on whether or not they were meant to be used in the same physical space or across multiple sites and whether they were meant to be used by all members simultaneously or at different times [77]. Zoom, for example, is categorized as a synchronous distributed interaction technology meant for team members to use at the same time in different physical spaces, while email, on the other hand, is an asynchronous distributed interaction technology meant for members to be able to use at different times and in different places. The development of a field of technologies intended to support collaboration among co-located and distributed teams brought about a new focus on supporting the aforementioned qualities of effective communication and shared knowledge, giving rise to a critical movement for using technology to support team cognition in teams through awareness [120, 118].

The advent of technology-supported distributed teaming has made team cognition a vital concept within the field of CSCW. The CSCW research community was an early adopter of the concept of team cognition and its sub-constructs of situational awareness, perception, and shared mental models, with early research focusing on the role of these constructs in remote and co-located learning environments [295]. For example, the lens of team cognition and shared mental models has been used in CSCW research to understand and measure the effectiveness of computational systems that attempt to enhance team coordination and communication [334, 307], predict team performance in distributed virtual teams through collective intelligence [162], and understand computer-mediated collaboration within fast-paced virtual environments [225]. Shared mental models have also been used in broader HCI research ranging from understanding the role of non-verbal communication in online multiplayer games [170] to supporting common ground in computer-supported teamwork [47]. Awareness is an essential component of groupware design for supporting the development and sustainment of team cognition in technologically based teams such as distributed and human-AI teams. The concept of awareness has been a hallmark of CSCW research for decades [116, 187], being integral to the design of technology for cooperative work [120, 121, 116]. Awareness in CSCW can be defined as understanding what others are doing, why they are doing it, and how this relates to their activities [73, 116, 122, 120]. As such, prior empirical research has shown that awareness plays a crucial role in how teamwork is perceived and performed [87, 123]. Additionally, the concept of awareness is intrinsically linked to situational awareness and team cognition but is defined in the context of collaborative technology design as workspace awareness [122, 120]. Workspace awareness distinguishes itself from awareness and situational awareness by focusing on individuals' real-time understanding of others' interaction within their shared workspace [120, 122]. Workspace awareness is an individual's knowledge of the other individuals within that workspace instead of simply the workspace itself and is limited to only events occurring within the shared workspace.

Properly designing technology to support awareness and the specific components of workspace awareness improves technology-supported teams (such as distributed teams) by simplifying the communication necessary for coordination and shared understanding. These designs enhance individuals' ability to anticipate their teammates' actions and activity across time scales and improve teams' ability to coordinate their actions [120, 262]. The influence of this work on designing for workspace awareness can be seen in technologies pre-dating Gutwin & Greenberg's framework [262] and long after it was published in collaborative web applications such as Google Docs, with many making reusable toolkits to support workspace awareness features (e.g., pointers, user cursors) [138]. Further still, empirical work has shown how vital the role of visual information and visual cues are to collaborative technologies, providing evidence that visual information can help support the conversation surrounding a joint activity in a team by providing evidence of common ground and situational awareness [104]. Workspace awareness can be designed for, and many of the shortcomings of technologically supported teams hindering team cognition development can be at least partially mitigated. AI teammates can also be designed to support similar aspects vital to team cognition development.

2.2.3 Utilizing AI Teammates for Team Cognition

No longer drawing parallels to virtual and other distributed teams, being active in team cognition development does not come easily and is challenging for human-AI teams. Several empirical studies have found that these teams are less likely to perceive team cognition than human-only teams [215, 224, 278]. Human-AI teams also have trouble developing team cognition when human teammates negatively perceive AI, especially when multiple AI teammates are involved [224]. The negative perceptions influencing team cognition center around the limitations of AI as a teammate, which means humans must alter their expectations for how their teammate will communicate, improvise, and implicitly understand what to do in dynamic environments [342]. This makes sense as team cognition development is heavily reliant on factors such as communication [193, 131], awareness [121], experience [137], and perceptions of teammates [260]. Noting that a benefit of team cognition is a reduction in the need for verbal communication [56, 52], while also acknowledging that communication is required to develop team cognition [193, 131] creates a challenging "chicken or the egg" dilemma for human-AI teams that often exhibit communication challenges [238, 68, 212]. The challenge of supporting team cognition in human-AI teams is exacerbated when humans have negative perceptions of the AI, creating an additional impediment [315, 342, 68, 224]. Thus, it is helpful for human-AI teams to rely on other team cognition support mechanisms such as experience, awareness, and nonverbal communication.

The perceptions of AI as teammates should be grounded in their capability to instill adequate cohesion and trust to enable team cognition development. As previous research has found, AI teammates have been found to hinder team cognition development [224], hampered the efficiency of information "push" and "pull" communication integral to team cognition [68], and decrease teams' ability to anticipate information needs [68]. However, overcoming these perceptual shortcomings is possible and has been empirically investigated. Increasing the level of AI teammate transparency of its capabilities and adapting the AI's communication ability can enhance the AI's perception as a teammate [224]. Furthermore, modifying how humans are trained with AI teammates has been shown to improve the level of trust humans are willing to place in their AI teammates, with cross-training being particularly effective [229, 228]. This type of training lets the human and the AI teammate switch roles on the team to learn about the other's job, enabling them to better anticipate their needs and contextual constraints later on [314]. Cross-training has also enhanced the quality of shared mental models in human teams and their subsequent performance [197]. Bi-directional communication and bi-directional transparency can also contribute to improving the perception of AI teammates by enhancing human *and* AI teammates' understanding of each other's intent, current beliefs, goals, and potential obstacles [277, 275, 88]. Making human teammates aware of overlapping goals is also incredibly beneficial to enhancing performance and reducing challenges to cooperation in human-AI teaming simulation studies [178]. The system transparency of the AI teammate itself is also related to improved outcomes in human-AI interaction, having been linked to improved performance, trust, trust calibration, perceived usability, and agreement [332, 35, 216]. What do these interventions in the design of AI teammates consist of, and how can they be implemented in a realistic manner that also leverages the technical advantages brought by AI?

A great deal of successful research has enhanced the understanding of human-

AI teams in recent years; however, designing for team cognition in human-AI teams remains nuanced, with a great deal of uncertainty left for practitioners to fill in the gaps. Given the topic of the current section on design and application for human-AI teams and other technology-driven teams, a focus on how the results of the present dissertation will improve human-AI teams in their application is warranted. Specifically, how it will remove the guesswork for developers and practitioners when designing for team cognition in human-AI teams. Two of the leading topics in AI design are transparency and explainability, with transparency being one of the features covered in the previous paragraph that improved teammate attitudes toward AI. System transparency is defined in the recent National Academies of Science State-of-the-Art and Research Needs in human-AI teaming as "the understandability and predictability of the system" [86] p. 146 and within the context of an AI teammate involves it's "abilities to afford an operator's comprehension about an intelligent agent's intent, performance, future plans, and reasoning process," [33] p. 2, [227]. However, AI system transparency is distinct from AI explainability in that AI explainability is designed and implemented after action sequences, not during them, and is defined as information provided in a "backward-looking manner on the logic, process, factors, or reasoning upon which the system's actions or recommendations are based," [227] p. 31. AI transparency is provided in real-time during action sequences to contribute to understanding the actions of an AI teammate in concert with individual and team situational awareness [227]. An excellent example of AI transparency in design comes from work done by Mercado and colleagues in 2016, which used the opaqueness of an icon to convey the level of certainty the AI held about their decision on that particular prediction, all within the interface the human team member used to interact with the AI teammate [216]. However, in the case of AI system transparency, there can be too much of a good thing, where high levels of transparency lead to reduced performance and increased cognitive workload for human teammates [332]. System transparency and explainability for AI teammates are just a few of the many design aspects that influence team cognition that developers and practitioners must contend with; unfortunately, there is very little research contributing to the understanding of how to support the development of team cognition in human-AI teams for them to reference during the design process. Without this understanding, the state of team cognition in human-AI teams will remain underrepresented, misunderstood, and underutilized, which will cause the effectiveness and applicability of human-AI teams to suffer.

2.3 Human-AI Teaming

A central component of the current dissertation, human-AI teams represent the latest technical advancement in teaming technology. Human-AI teams represent an exciting and novel development for multiple research fields ranging from industrial-organizational psychology to deep learning. This wide-ranging interest has been a common theme of this review, as human-AI teams are a research topic uniquely suited to human-centered computing. As pointed out by Guzdial, "It is not surprising that understanding today's world of ubiquitous computing requires a blend of computing and social science...What is interesting about our modern computing milieu is the blend of technology, humans, and community. Human-centered computing is a new sub-discipline of computer science that prepares students for studying our socio-technical world," [124] p. 1. Human-AI teams are an incredibly interdisciplinary field that requires significant collaboration across research areas, which the current dissertation directly addresses with its focus on leveraging AI design to support team cognition in human-AI teams. The current section reviews the multi-disciplinary nature of human-AI teams across time, detailing the benefits of these teams, how they came to be, and their current state using the following topics: 1) from automation to autonomy; 2) current human-AI teaming; and 3) team cognition in human-AI teams.

2.3.1 The Eventual Path to Autonomy

Human-AI teaming began being theorized as a potential evolution of teams and technology in the early 1990s, well before the technology had advanced enough to make them a reality. As such, the concept of human-AI teams is not new and was discussed by several researchers in the 1990s, with Woods' 1996 work describing how "...automation [is] seen as more autonomous machine agents. Introducing automated and intelligent agents into a larger system in effect is a change in the team composition" [331] p. 2. This sentiment was echoed by a variety of other researchers in a conceptual capacity throughout the decade (e.g., [226, 194, 150]). At the time, however, the technology required to design, develop, and implement fully autonomous teammates was not reasonably available. Automated technologies were far more commonplace, which was also reflected in the literature of the time [119, 95, 130]. Research on automated technologies is incredibly relevant to the current dissertation as many of the concepts already reviewed (e.g., situational awareness, transparency, and explainability) were initially put forth within the context of automation [82, 86]. As such, many lessons from human-automation interaction can be taken from this early research and applied to human-AI interaction, with the "lumberjack effect" being a notable observation from this research [237]. The "lumberjack effect" states that higher levels of automation result in higher levels of human-system performance when the system operates as intended but also results in increased human dependence on the system, leading to more catastrophic failures when the automation fails

[237, 289, 242]. The 1997 work by Parasuraman and Riley describes consequences and factors leading to the use, misuse, disuse, and abuse by humans working with automation related to the lumberjack effect. They discuss how factors such as trust [171], cognitive load [261], individual differences [294], and contextual risk [261] all play a large part in humans choice to use automated technologies [242]. Such findings are valuable to review in light of the current dissertation to avoid "reinventing the wheel" and utilize the existing literature to inform recent research and design for human-AI interaction.

The factors contributing to automation use, misuse, disuse, and abuse are incredibly relevant to team cognition in human-AI teams. They often involve human attitudes towards the system and direct aspects of team cognition such as situational awareness and environmental context. Specifically, the factors contributing to automation misuse (i.e., overreliance) include similar concepts such as workload and automation reliability but also the saliency of the automation state, which is a failure of the system design to adequately contribute to accurate situational awareness [242]. Automation disuse, or the failure to use a system when it would greatly benefit the process, was contributed to by false alarms [273]. Automation abuse was attributed to designers and practitioners implementing automation into situations without considering the consequence of eventual automation failures [242]. Both factors contributing to automation misuse and abuse heavily relate to the *context* humans and automation are interacting within, which is a central tenet of consideration for interactive team cognition. This further highlights the importance of considering team cognition in designing and implementing human-AI teams and human-AI interaction more generally. Much of these tenets from human-automation interaction research remain relevant for human-AI interaction and teaming research, and more recently, technical AI development research [265, 37]. The findings regarding trust, transparency, situational awareness, and system acceptance are relevant to the design of AI teammates supporting team cognition, despite the fundamentally different nature of automated technologies from autonomous technologies.

Automated technologies fundamentally differ from autonomous technologies; however, several principles from automation research remain highly relevant to the research and development of autonomous systems such as AI teammates. Automation is similar to autonomy, with a critical distinction between the two, which is that autonomous systems can make decisions independent from human partners [238]. This distinction was clarified by O'Neill and colleagues' 2020 paper, which adapted Parasuramann and colleagues' 2000 paper on the levels of automation to have clear boundaries on when a system should be classified as autonomous instead of automated. Parasuraman and colleagues' 2000 paper on the levels of automation defined its levels on a scale ranging from one to ten, with ten being a system that makes decisions entirely without human input and even ignores them outright, while level one represents a system manually operated by a human [243]. O'Neill and colleagues modified scale distinguishes autonomy as beginning at level six as the system begins to exert control over decisions, displaying independence and an ability to respond to dynamic environments [238]. On this scale, autonomy is broken down between partial autonomy (levels 5-6) and high autonomy (levels 7-10), with partially autonomous agents recommending and executing actions unless stopped by their human counterparts and highly autonomous agents requiring no interventions by human teammates before implementing a decision [238]. However, the potential for technology to become a teammate was empirically explored well before AI development technologies such as Google's TensorFlow were easily accessible [1], as Nass and colleagues found that humans were entirely willing to accept a computer as a full teammate back in 1996 [226]. While the cultural perception of AI and other autonomous technologies has changed significantly since that time [342, 13], it draws a significant parallel to the research by McNeese and colleagues decades later in 2018 showcasing that human-AI teaming is possible and their performance can exceed that of human only teams [212]. Both studies found that humans were willing to accept technology as a teammate and treat it with the same expectations as a typical human teammate. For better (e.g., identified with it and felt higher levels of cooperation [226]) or for worse (e.g., frustration with system failures and disruption of typical interaction patterns [342, 68]), technology as a teammate is here to stay, and the consequences of raising it to the level of a teammate instead of a simple tool are incredibly impactful.

2.3.2 Current-Human-AI Teaming

Current research on human-AI teaming is incredibly diverse, with a wide range of focuses across various disciplines, and fundamental concepts have begun to take shape as it matures. For example, the very definition of a human-AI team was beset with confusion when terms with very similar but distinct meanings were used by researchers (e.g., automation, autonomy, AI, robots, and machines). Accordingly, human-AI teaming research was and is still, to a degree, joined with research involving the aforementioned automated or physically embodied systems that may not contain any agency, with no delineation being made between them and those with independence. Additionally, some research involving human-AI interaction could be included without regard to the requirements of teaming as defined in seminal teaming literature [269, 27]. This lack of clarity led to a considerable amount of confusion regarding what factors truly represented a human-AI team, and it was clear that a comprehensive definition was necessary. Again provided by O'Neill and colleagues 2020 work, human-AI teams were defined as having the following distinct characteristics: 1) teams where agents are viewed as "agentic" by their human teammates (agents have a significant degree of independent decision-making); 2) the agents must have a role interdependent with the roles of their human teammates; and 3) there must be one or more humans and one or more autonomous agents working towards a common goal [238]. The current dissertation subscribes to this definition of human-AI teaming, as previously mentioned in the introduction, and a focus on the AI teammates' ability to make independent contributions towards the shared goal *alongside* advancing team cognition. This is important because human-AI team performance outcomes are often subject to severe limitations when task contexts are highly dynamic and rely heavily on effective communication and coordination [65, 279, 70].

The performance of human-AI teams has been a central theme of research in the field, bringing to light mixed results that highlight several shortcomings. The performance outcomes of human-AI team performance have been studied in various contexts ranging from medical [316] to military [89]. Still, human-AI teams have frequently underperformed their human-only counterparts [212, 67, 89]. Alternatively, other experiments display incredible human-AI team performance, outpacing not only human-only teams but even teams consisting of all AI [316]. This disparity can likely be attributed to three major factors: 1) not all studies use true AI, which is capable of expert-level performance when properly trained [100]; 2) the more abstract the team task is, the harder it becomes to train high performing agents [276] (but not impossible [142]); and 3) a distinct lack of proper design and integration within human-AI teams that leads to confusion and poor understanding between the two types of team members. The solution to this discrepancy is to leverage the individual strengths of the AI *and* the human to move team effectiveness beyond what each can achieve alone. While AI training will always strive for better reliability and individual performance, these performance gains will not mean much if the collective exists in dysfunction, unable to benefit from the unique abilities of each team member. This assertion is backed up by many recent studies which found that while high-performing AI does engender higher trust in human teammates [339], their performance was not a predictor of the teams' performance as a whole [8]. Additionally, improvements to an AI teammate's performance and effectiveness can be offset entirely if those improvements change the user experience and present compatibility problems [9], further emphasizing the need to focus on team-level research such as team cognition. Team cognition holds the key to leveraging the unique advantages presented by human-AI teams, as effective team cognition can allow human-AI teams to possess a mutual understanding of the tasks and team functions each is best suited to accomplish. This ability thus allows for the quick and efficient allocation of team functions and direction [8]. AI teammates may even be capable of helping new human teammates develop better team cognition by acting as exemplars for what effective teaming behaviors and strategies consist of within specific contexts. However, human teammates *must* be willing to accept these contributions from their AI teammates first for them to have any positive effect.

Team composition affects the characteristics of human-AI teams directly through the social relationship between humans and artificial agents. Similar effects can be seen in human-only teams as team composition is known to affect team cognition either through differing teammate skills, abilities, individual differences, and negative attitudes [79, 81, 93]. Unfortunately, humans accepting agents as full team members and giving them an equal level of respect as their human counterparts is not nearly as straightforward as it seems, despite past research indicating that humans are willing to accept AI as a teammate [226, 212]. Much of the difficulty can be attributed to individual differences such as past experiences [13, 34, 127, 126], which can often be problematic given AI teammates' difficulty communicating and coordinating naturally [342]. For example, several studies have identified negative behavior and attitudes towards AI teammates from human participants when playing a video game with AI teammates. Participants adopted a neo-feudalistic view of the agent teammates, creating unequal rights for the agents [320]. Such results are also found in the CSCW domain, where research indicates that humans are more likely to place blame for failures in online cooperative games on AI rather than human teammates, even if that AI teammate was a human pretending to be an AI [217]. Humans were also less likely to save AI teammates than human teammates (though this is arguably an exercise in one of the advantages of AI teammates in that they can reduce danger for humans [327]) and significantly misjudged their AI teammate's abilities compared to judging their human teammate's abilities [236]. Such results may be characterized by the consequences of the social identity theory, which posits the existence of "in-group" and "out-group" factions within teams/groups. These factions lead those in the in-group to positively see others in their group and identify with the group's common stereotype [299]. In contrast, those in the in-group dehumanize members of the out-group [309]. Recent CSCW research supports the existence of this perspective, as humans were shown to treat AI unfairly and specifically used the terms "I" and "they" to describe humans and AI teammates, respectively [342]. Research on trust in human-AI teams also reflects deficiencies in the relationship between humans and AI as they make humans trust their teammates less [210, 211], revealing the consequences of poor team cognition. Trust was also highly related to team performance [210], another construct related to team cognition that further emphasizes the importance of studying the outcomes in concert with the construct itself. Such a dysfunctional relationship between humans and agents may make it exceptionally difficult for human-AI teams to support team cognition. Specifically, human team members may be adversely affected when outnumbered by AI and vice versa, making it essential to understand how team composition affects human-AI teams and if theories such as social identity apply to helping those in the field understand the cause of such effects and what potential interventions may remedy them.

2.3.3 Team Cognition in Human-AI Teams

While team cognition has received attention from researchers within the context of human-only teaming [230, 221], how it may be fostered and experienced differently in human-AI teams is understudied. The most recent review of existing empirical research on human-AI teams in late 2020 included team cognition as one of several focus areas [238]. The small handful of studies on team cognition in human-AI teams revealed several insights despite the overall absence of literature. For example, it was shown that virtual agents with agreeable personality traits lead to higher perceived team cognition [132]. Additionally, the study's results indicated that agents with personality traits more closely aligned to their human teammates engendered higher perceived team cognition [132], which follows a trend seen in human-only relationships that humans prefer those more similar to themselves [7]. Perceived team cognition also shared a positive relationship with verbal and non-verbal communication in human-AI teams and retained its positive connection with team performance [131]. Unfortunately, other research has indicated roadblocks that may prevent high levels of team cognition compared to human-only teams. Specifically, research shows that human-AI teams may possess more rigid team cognition (inability to adapt to environmental changes rapidly) [65]; however, human-AI teams can overcome this rigidity if they can engage in effective communication and develop accurate team situational awareness [69]. Developing team situational awareness is essential to creating and sustaining effective team cognition for human-AI teams, and past research backs this up. Specifically, conceptual research on how artificial teammates can contribute to team cognition has identified several key areas where AI teammates may supplement team cognition in existing human-only teams. These areas were identified as maintaining awareness of teammate actions and status, informing teammates of vital changes to intra-team and extra-team changes, monitoring task and team progress towards individual and team goals, and having the ability to provide feedback on performance to correct errors, and augmenting team memory [59]. Many of these suggestions revolve around awareness, situational awareness, and transactive memory systems, all centering around the concepts of shared knowledge. However, these suggestions also apply to interactive team cognition in that they emphasize the importance of the *process* of supporting this shared knowledge and driving the team towards their overall goal through the AI teammates' intelligent information sharing and monitoring. These suggestions are not strictly conceptual, as applied artificial agents have attempted to perfect information sharing within human-machine partnerships [338, 337]. While these applications have been limited in scope and received little follow-up research, they proved that the concept could be implemented to improve team cognition in these human-AI partnerships.

Research on the fundamentals of team cognition in human-AI teams is essential to utilizing applied research and interventions to improve team cognition. Specifically, research has focused on applying interventions to human-AI teams to enhance team cognition and effectiveness. For example, a cross-training technique that leveraged Markov Decision Chains to represent the autonomous agent's mental model and fine-tune it by training with their human teammate resulted in significantly improved levels of team performance and trust [229, 288]. Another study deployed a system that shared team members' cognitive load and beliefs with other human and autonomous agent teammates. The autonomous agent teammates then utilized the information to better understand the humans' current status and develop better shared mental models [90]. These implementations align with the position expressed earlier that effective information sharing and situational awareness are essential to team cognition in human-AI teams. These studies also demonstrate the potential for developing systems that enhance team cognition by leveraging the inherent technical advantages offered by AI technologies. Related research takes this potential for AI teammates to improve team cognition by deploying an AI teammate capable of monitoring teammate statuses for information overload and then taking some of that pressure off of teammates accordingly [337], representing an effective back-up behavior. It is also important to note that team cognition in human-AI teams may be a concept that requires a significant departure from existing theory, given the differences in expectations [342], communication patterns [68], and AI behavior. As such, it is possible that some or all aspects of team cognition require adaptation in conceptualization. Consequentially, while the development of team situational awareness in human-AI teams may progress better or as expected in experienced human-only teams [213], other aspects of team cognition, such as shared mental models [224] and or communication effectiveness [68] may suffer. As such, for concepts similar to shared mental models, the idea of human teammates' shared mental model of their AI teammate may be a concept that they must adopt to better reflect the reality of human-AI teaming. The importance of human team members' shared knowledge of their AI teammates is not novel and shows that it is related to the performance of these teams [8]. This concept is also highly relevant to the discussion of transparency and explainability of AI systems [9], further emphasizing their importance and potential need to be adapted into the concept of team cognition.

Lastly, research on team cognition in human-AI teams within the shared knowledge perspective should utilize robust measures of the construct known to capture its content and structure. Most studies measuring team cognition in human-AI teams have used broadly applicable Likert scale questionnaires, which, while more accessible, only capture the content of a shared mental model, not its structure [221]. Therefore, simplified measures of team cognition are only regarded as elicitation tools and not shared mental model measurement techniques [221]. Paired sentence comparison and concept mapping are examples of techniques that measure both content and structure [221]. While a variety of team cognition measurements in human-AI research is positive, given the importance of capturing shared mental model structure and the lack of studies utilizing measures that do so, a significant gap in the literature is exposed. As such, the current dissertation uses a wide range of measures that include the content and structure of shared mental models to reflect the state of the literature better to ensure its broad applicability and articulate any potential differences between measures that may apply to human-AI teams.

2.4 Conclusion

The literature reviewed above outlines theories on team effectiveness through the IMOI model and places team cognition within that model while going a step further by placing a lens on team cognition and what factors contribute to its development. However, human-AI teams represent a new teaming paradigm where the expectations and reality human teammates have for AI teammates may not necessarily align. Given the rapidly increasing role of AI in teaming and within society in general, it is essential to understand how AI teammates affect team cognition development, support, and sustainment throughout the teaming process. This is especially important given the major role that team cognition has on teaming outcomes such as objective and subjective performance. Despite the significant differences between AI and human teammates, it is also essential to ask what team cognition concepts should be adapted to better reflect the nature of human-AI teams. As such, if AI teammates are going to become a significant component in teaming and have any chance at effectively contributing to team cognition, the following goals should be addressed:

Goal 1: Empirically consider how including AI teammates affects team cognition development and sustainment compared to existing research on human-only teams.

Goal 2: The development of team cognition for human-AI teams should ensure that human teammates are willing to accept the technology as a teammate as an antecedent to AI teammates contributing to team cognition.

Goal 3: Measurement techniques and theoretical perspectives should be utilized holistically to better understand how human-AI team cognition can be best understood, studied, and improved.

Goal 4: Existing theoretical and empirical research should be followed as a guide but should be flexible as team cognition in human-AI teams may differ significantly from human-only teams, especially when considering what constructs and components make up team cognition and are important to its development.

This review also highlights the need to develop AI capable of contributing to developing and sustaining team cognition in these new human-AI teams. The reviewed literature shows that human-AI teams face similar challenges to developing team cognition as virtual teams once did. Just as virtual teams were able to create methods to support their awareness and team cognition, human-AI teams can develop similar solutions. Specifically, these solutions must consider the holistic nature of team cognition, the importance of attitudes to its sustainment, and the principles developed in human-automation interaction research. Given the nature of AI teammates and their technical advantages, these solutions can go beyond what team cognition is capable of within human-only teams and become a unique facet of these teams that represents one of several advantages they offer for potential practitioners. Thus, the current dissertation must ascribe to goals that ensure the relevance of this research to applied human-AI teams and their outcomes:

Goal 5: The current dissertation must be capable of being distilled into actionable and applicable design recommendations for practitioners seeking to improve team cognition in human-AI teams.

Goal 6: The research of the current dissertation must be mindful of the influence of AI on society and consider the lived experiences of those interacting with AI and participating in human-AI teams.

Goal 7: The current dissertation's contributions should use lessons learned from previous literature and seek to build upon existing principles as much as possible.

Adhering to the seven goals listed above will ensure the relevance and applicability of the current dissertation to the field of human-AI teaming and related fields such as team cognition, CSCW, and HCI. Following this principle, each study within the present dissertation iteratively works towards an eventual understanding of team cognition in human-AI teams, how it has changed, and how this new knowledge applies to the design of future AI teammates. The research conducted within the current dissertation is then best positioned to support practitioners of human-AI teams, including developers, managers, and team members, to have the best possible experiences and outcomes, allowing these humans and AI to work together as partners to achieve feats not possible for either alone.

Chapter 3

Study 1: Assessing the Influence of AI Teammates on Team Cognition in Human-AI Teams

Note: This work was published at GROUP 2022/23 [278].

3.1 Study 1: Overview

The first study of the current dissertation is a critical first step in understanding team cognition development and support in human-AI teams. Study 1 is also the first empirical investigation of the content *and* structure of shared mental models between human teammates operating as part of human-AI teams and further investigates the role of team composition on perceived team cognition compared to human-only teams. Developing a better understanding of team cognition in human-AI teams from this investigation is critical, given AI's increasing role in society socially and organizationally.

The primary goal of Study 1 was to investigate how working with an AI teammate influences team cognition development compared to human-only teams, which directly relate to D-RQ1 and D-RQ2. Given the significant differences that come with working alongside an AI teammate versus a human teammate, the variable of team composition, in the form of including AI teammates, is highly relevant to how team cognition develops within human-AI teams. This relationship makes the three levels of team composition implemented in Study 1 incredibly useful to answering D-RQ1, which sought to investigate how AI teammates influence the development of team cognition. Additionally, Study 1 is tasked with helping answer D-RQ2, which inquires how attitudes relate to and influence team cognition development within human-AI teams. As such, Study 1 also collects data relevant to attitudes, including trust and many qualitative open-response questions that target participants' opinions, experiences, and attitudes toward working with an AI teammate. To help provide answers to the first two D-RQs, Study 1 adopts the following study-specific research questions, which are numbered by the D-RQ they stem from:

RQ1.1: How is the development of team cognition in human-AI teams similar or different from its development in human-only teams?

RQ2.1: How does the inclusion of AI teammates affect the development and outcomes of team cognition in human-AI teams?

RQ2.1.1: In regard to perceived team cognition?

RQ2.1.2: In regard to team performance?

RQ2.1.3: In regard to trust?

Study 1 addresses these research questions utilizing a well-known teaming research platform, NeoCITIES, which has a long and storied past in the team research literature that specifically targeted team cognition. NeoCITIES underwent a modern redesign for web-based usage, AI integration, and deployment, allowing three individual teammates to work concurrently, whether co-located or remotely distributed.

Study 1 begins the dissertation-wide process of addressing all three research gaps challenging human-AI teaming. The first is developing knowledge regarding what aspects of team cognition AI should support to drive effective team processes. This research gap is crucial given the encompassing nature of team cognition as a construct, including several individual concepts [58] and multiple co-existing theoretical approaches [209]. Parsing down this massive construct to understand what concepts within it drive effective team processes and directly lead to improvements in team outcomes will be essential to developing better AI teammates. This importance to design leads to the second research gap beginning to be addressed, which is systematically understanding the outcomes of designing AI teammates meant to support team cognition. Research on AI teammate design has predominately centered on their effects on performance without any attention given to the design's impact on emergent states such as team cognition that lead to those performance outcomes [238]. Focusing first on how AI teammate design for team cognition influences emergent states like team cognition and trust will enable a more holistic and fundamental improvement for human-AI teaming outcomes. Lastly, this research improves the currently insufficient understanding of the potential differences in team cognition between human-only and human-AI teams. Given the importance of interaction to team cognition development [163], the extensive knowledge from decades of research on team cognition in human-only teams may not be a one-to-one match when considering human-AI teams as human-AI interaction is markedly different. Confronting these research gaps contributes essential knowledge to human-AI teaming theoretically and practically, as knowledge is built on the role and nature of team cognition

Condition Number	Team Composition Pattern
Condition 1 (HHH)	Human-Human-Human
Condition 2 (HHA)	Human-Human-AI
Condition 3 (HAA)	Human-AI-AI

Table 3.1: Experimental Conditions.

in human-AI teams and how AI teammates affect it.

Because the current study is self-contained and does not include any sub-studies, the following chapter is structured as such: (1) methodology; (2) quantitative results; (3) qualitative results; (4) discussion; (5) design recommendations; (6) conclusion.

3.2 Study 1: Methods

3.2.1 Experimental Design

Study 1 of the current dissertation utilized a mixed-methods design to study team cognition formation in teams with varying numbers of AIs and humans. The experiment leveraged the well-published and validated team research platform known as NeoCITIES [154, 208, 139, 140], which provides an excellent environment to study team cognition within human-human teams [128] and human-AI teams [276]. A 1 x 3 (Team Composition: Human-Human-Human, Human-Human-AI, Human-AI-AI) experimental design was developed to study the effect of various team compositions on the development of team cognition and its related outcomes (see Table 3.1).

3.2.2 NeoCITIES Task and Roles

NeoCITIES uses a fictional college town where three players work together in interdependent roles to respond to and complete emergency tasks over time. These

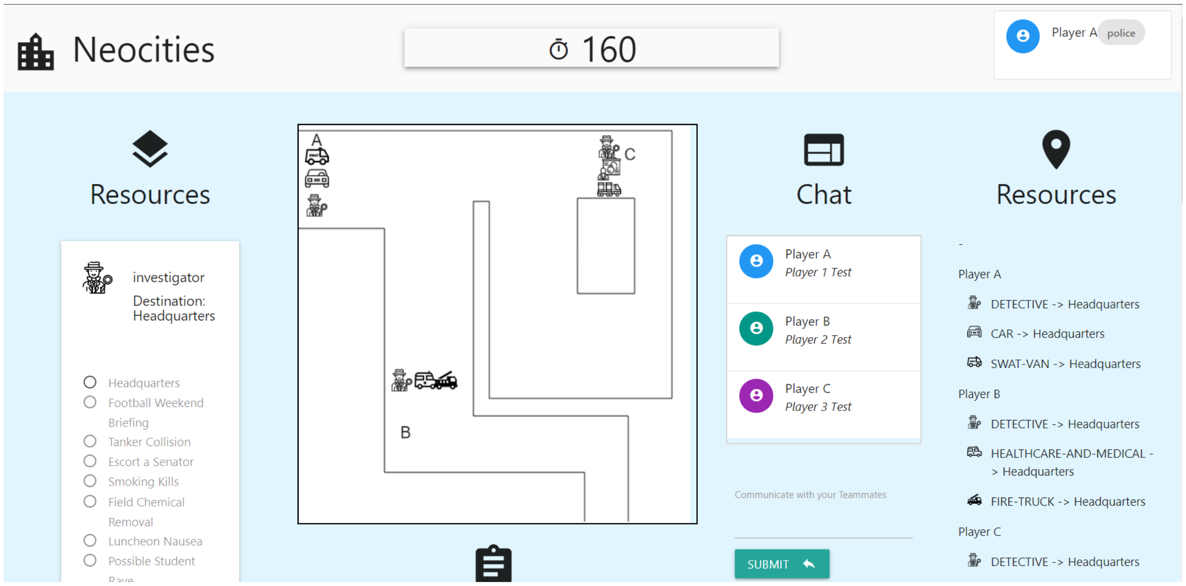


Figure 3.1: NeoCITIES Home Screen

roles include Hazmat, Police, and Fire, each with three resources available to them to use to address emergency events. With three unique interdependent roles, NeoCITIES provides an excellent opportunity to observe the possible combinations of interactions between humans and AI teammates in emergency response management. The simulation interface can be seen in Figure 3.1.

The interface of NeoCITIES is designed to simulate an emergency response role as if the user were acting in a supervisory position. As part of their duty in this fictional college town, participants must determine when and where their respective resources must be assigned based on active events in coordination with other teammates. The interface presents consistent tools to create situational awareness amongst all team members, regardless of their roles. Tools include a manifest of their resources, active and past events, event descriptions, a chat function to communicate with teammates, and the current objective for all team resources. Participants were also given a spatially accurate map that displayed each teammate's resources, home bases, and currently active events. Accordingly, all team members could establish

Order	Emergency Event	Necessary Resources
1.	Football Weekend Briefing	Investigator
2.	Tanker Collision	Squad Car, Fire Truck, Chemical Truck *(In That Order)*
3.	Escort a Senator	SWAT Van
4.	Smoking Kills	Fire Truck
5.	Field Chemical Removal	Chemical Truck
6.	Luncheon Nausea	Ambulance, Investigator
7.	Possible Student Rave	Investigator, Squad Car
8.	Old Main Frame Shoppe Fire	Investigator, Fire Truck
9.	City Hall Bomb Threat	Bomb Squad, Investigator

Table 3.2: NeoCITIES Events and Necessary Resources to Complete Them.

shared cognition for their responsibilities, resources, strategies, and teammates.

During each of the four nine-minute rounds, nine different events occurred that required a response from the team to complete successfully. These nine events and their requirements are shown in Table 3.2, and each event's location was changed between rounds. Each event must be completed with specific resources, but an additional layer of complexity is introduced as many of the resources differ in speed. For example, in the time limit allowed, some resources could cover more distance than others; specifically, the slow resources consisted of the investigators, ambulance, and chemical truck. The other resources were equal in speed except for the fire truck, the fastest resource. Each event in Table 3.2 can also be categorized into three difficulty ratings (1 = Low, 2 = Medium, 3 = High) based on the number of resources required, the speed of those resources, and the location of those events. This simulation feature made time and distance an additional dimension for the team to consider in decision-making, allowing for further insights into how individuals are aware of their respective team members' situations and approaches to a problem.

3.2.3 AI Teammate

Study 1 incorporated an expert system programmed to complete the NeoCITIES task in either the Police or Hazmat role with high accuracy and flexibility to adapt to needs signaled by its teammates. This system was only applied to team conditions with an AI team member (Condition 2: HHA and Condition 3: HAA; Condition 2 fielded an AI in the Hazmat role only). Expert systems are a branch of applied AI designed to represent expert-level human knowledge in a task [181]. In the current study, the expert system continuously managed the allocation of resources to events based on the simulation state. The expert system was flexible to human teammates because it could make decisions that reacted to the humans' actions or requests to increase the team score. Accordingly, the expert system was developed to recognize its teammates' decisions and plan on the resultant consequences of those decisions. This implementation allowed the expert system to possess a collaborative "mentality" with which they replicate their teammates' level of awareness to assist them better [31].

The chat communication provided by the AI was not a feature of the expert system but was accomplished using the Wizard of Oz technique. This technique has a trained experimenter represent a feature of the system (chat communication in this case) to an unknowing participant [203]. The Wizard of OZ technique is often used to simulate capabilities of AI when not fully computationally available [203]. The trained experimenters followed a script developed through multiple iterations of pilot testing. The AI's capabilities were conveyed to participants beforehand to help control for participants' expectations of the AI teammate. The AI was described as an expert-level player in their assigned role (Hazmat in the HHA condition, Hazmat and Fire in the HAA condition). Additionally, the AI was described as having advanced text generation and understanding capabilities similar to Siri or Google Home regarding the NeoCITIES simulation, but no other topics. Thus, the AI could take requests, offer information, and respond coherently if the subject regarded the NeoCITIES simulation.

3.2.4 Participants

Study 1 recruited 66 participants, with 35 identifying as women and the rest identifying as men. These participants were recruited from a departmental subject pool at a major university in the USA (see Table 3.3). The average age of participants was 18.91 (SD = 1.51). The participants were randomly placed into conditions, teams, and NeoCITIES roles. Participants received course credit for their time as an incentive for their participation.

3.2.5 Procedure

The novel COVID-19 global pandemic forced in-person research to a standstill due to the highly contagious nature of the coronavirus [267]. Following appropriate social distancing techniques to mitigate the risk of infection, this study was conducted remotely through the high-fidelity video-conferencing application Zoom, which is very effective for remote research and was used by multiple researchers in the past year [3, 114]. All Zoom sessions were monitored and conducted by trained experimenters who continuously observed participants, much like in a typical in-person experimental setting. Any participants observed within the simulation, survey, or Zoom not paying attention or taking the experiment seriously were dismissed. Trained experimenters gave all participants the same information and instructions following a predefined protocol approved by the local Institutional Review Board.

Overall: **66** (32 Teams)		
HHH: **36** (12 Teams)	HHA: **20** (10 Teams)	HAA: **10** (10 Teams)

Table 3.3: Participant Numbers.

Each condition collected data from 10 teams; however, due to over-scheduling, the HHH condition consisted of 12 teams instead of 10. The experiment was conducted between-subjects where each participant only participated on one team in one condition. Students signed up for a particular testing time and received a Zoom meeting identification and password to enter the secure, virtual environment. Students were instructed through video and audio modalities and interacted with the experimenter similarly. The session began by collecting informed consent from the participants, followed by demographic information.

Afterward, experimenters introduced the study in more detail, providing information on team cognition and an overview of the simulation. Participants were assigned their team roles and were informed which role(s) would be taken by an AI teammate, if applicable. Participants were then taken to the simulation training page, where each simulation feature was explained in detail alongside video examples. This training page was followed by an in-game training round where all players could ask questions about the interface and the simulation itself. Once the training was complete and participants agreed to start, the next round began, and they could no longer ask questions.

After completing each round, participants were shown their score and linked to the next round, which began when team members were ready. Each round lasted nine minutes, and the teams worked together for 36 minutes within the NeoCITIES simulation. In congruence with past literature, this amount of time is adequate to develop team cognition [215, 224]. Upon completing the four rounds, the experimenter instructed the team to navigate the survey to complete the post-task measures. The post-task survey collected their team and task mental models, perceived cognition, trust in AI teammates, subjective team performance, and a series of free-response questions. Once the participants completed the post-task survey, participants were free to leave the Zoom session and were compensated for their participation with course credit.

3.2.6 Measures

3.2.6.1 Task and Team Mental Models

Mental Models of the task were measured using paired sentence comparisons [19], a strategy that has long been utilized in the past to measure both the content and structure of shared mental models [202, 221, 300, 167]. Participants were asked to judge the relationship between all significant task attributes on a nine-point Likert scale ranging from -4 to 4 and anchored by "Negatively Related" to "Positively Related" (with 0 representing "Not Related"). Task attributes were identified through comprehensive task analyses with subject matter experts (NeoCITIES simulation designers). The task attributes were as follows: *(1) familiarizing with the simulation layout, (2) determining which resources are at your individual disposal, (3) determining the location of the event, (4) sending resources to the event if available, (5) learn what resources your teammates have, (6) recall resources, (7) determine resource allocation based on event importance, (8) send resources in the correct order for critical events.* By assessing how positively related, unrelated, or negatively related each attribute was to the others, a network of relationships can be created, capturing the content and structure of their mental model.

The same methodology was applied to obtain a participant's team mental model, but the collection of teaming attributes was different. The attributes compared are more generalized and were explicitly taken from past shared mental model research [202, 174]: *(1) amount of information, (2) quality of information, (3) role/responsibility, (4) interaction patterns, (5) communication channels, (6) role interdependencies, (7) teammates' skill, (8) teammates' attitudes, (9) teammates' preferences.* Participants were also given definitions of each team attribute listed.

3.2.6.2 Mental Model Similarity

The Pathfinder network scaling algorithm was used to determine mental model similarity, which is familiar to shared mental model research [55, 202, 222, 221]. This algorithm inputs the participant's pairwise comparisons of the predefined attributes to create graphical representations of their mental models [286]. Each attribute represents a node in the graph, and the assessed relationships between attributes are the links between nodes. A similarity metric is produced by comparing two networks to provide a similarity rating between zero (no similarity) and one (perfect similarity) of the two. Pathfinder can only provide a similarity metric for two human team members at a time. Therefore, the HHH team had their three possible pairings averaged together for a single team similarity metric, which is standard practice [182, 272]. The AI could not provide any ratings, so the HHA condition produced a single comparison, while the HAA condition could not produce any comparison. This method was the same for the task and team mental models.

3.2.6.3 Perceived Team Cognition

Perceived team cognition was measured using the Teamwork Schema Questionnaire [241, 257]. Participants were asked to rate the importance of a series of statements to their idea of teamwork. The participants were then presented with the same statements, but this time, they were asked to rate how important they believed each statement was to their human teammate(s) idea of teamwork (one assessment for both human teammates). If the participant had AI teammate(s), they also completed an assessment for them (one assessment for both AI teammates). Together, these questions created a measure of congruence representative of total perceived team cognition. Perceived team cognition was calculated by taking the absolute difference between the participant's teamwork ratings and those they chose for their human and AI teammates. The scores were then scaled by the number of comparisons made on the team. Scores ranged between 0 and 84, with lower scores indicating higher perceived team cognition.

3.2.6.4 Objective Team Performance

$$EventScore = \frac{(end - start)}{(limit - start) * difficulty} \tag{3.1}$$

$$TeamScore = \frac{100 * [(worstScore - rawScore)}{(worstScore - bestScore + 1)]} \tag{3.2}$$

NeoCITIES calculates team performance using Equations 3.1 and 3.2, which have been utilized in past human-AI teaming research [215]. The variables within Equation 3.1 like "(end - start)" refer to the duration it took the team to complete the event successfully from the moment the event became active, while "limit" refers to the time limit associated with that event. The "difficulty" variable referred to that event's particular difficulty rating. The variables in Equation 3.2, including "Raw Score," represent the cumulative sum of the actual earned "Event Scores." In contrast, "Worst Score" represents the cumulative sum of the theoretical worst "Event Scores." Similarly, "Best Score" is the cumulative sum of the theoretical best "Event Scores." Accordingly, the equation produced objective team performance scores ranging from 0 to 100 based on the team's accuracy, speed, and ability to complete events, with higher scores indicating higher objective team performance. Additionally, the team scoring equations punished teams for wasting valuable resources on events that were not resolved.

3.2.6.5 Perceived Team Performance

Perceived team performance was measured using the Team Effectiveness Scale [259]. Subjects were asked to respond to questions that gauge how well they believe their team performed in the task on a five-point Likert scale ranging from "Strongly Disagree" to "Strongly Agree." The resulting scores ranged from 8 to 40, with higher scores indicating higher perceived team performance.

3.2.6.6 AI Trust

Participants' trust in the AI teammate they worked with in NeoCITIES was measured using statements derived from the principles of trust and distrust as defined in recent literature [186]. An example statement was, "Did you feel confident in the AI you just worked with?" which was rated on a five-point Likert scale ranging from "Strongly Disagree" to "Strongly Agree." Scores ranged from 6 to 30, with higher scores indicating higher levels of trust in the AI.

3.2.6.7 Qualitative Questions

Participants were allowed to provide more details about their experiences and opinions after the experiment within the post-task survey. The open-ended questions were designed to extract participants' experiences working with the AI and how team cognition developed within their team. Responses underwent thematic analysis [22, 102, 301, 103], where participants' statements were analyzed for themes related to the study's research questions. After the analysis, quotes were extracted to illustrate those themes concisely.

A thematic analysis offers insight into how individuals construct their perceptions, understandings, and accounts of the felt experiences in teamwork [296]. The analysis used the following phases: 1) two of the authors read through all of the narrative information provided by participants to achieve an understanding of how team cognition developed; 2) two of the authors then combed through the narratives to identify thematic topics based upon participants descriptions of how team cognition did or did not develop within their team, how, and why, as stated in RQ1.1; 3) all authors reviewed and debated the themes and sub-themes identified in Phase 2; 4) the first author identified strong example quotes that best represented the themes and sub-themes identified in Phase 3; 5) all authors again reviewed and debated the refined themes and sub-themes, using the quotes identified in Phase 4 to synthesize the similarities and differences in team cognition development between traditional human-human teams and human-AI teams. Differences between the two authors identifying themes in Phase 2 were resolved through open discussion and, if necessary, discussed and resolved with all authors in Phase 3.

3.3 Study 1: Quantitative Results

To answer the study-specific RQs, the findings are presented in two parts. Both sections report on data addressing the two stated research questions, with the quantitative results reporting on analyzing the empirical measures of performance, shared mental models, and the associated perceptions of team cognition, performance, and trust in AI. The qualitative section focuses on the similarities and differences in developing team cognition in human-only teams versus human-AI teams. The

Measure	HHH Mean (N)	HHH SD	HHA Mean (N)	HHA SD	HAA Mean (N)	HAA SD
Team Performance	86.62 (12)	3.81	89.08 (10)	3.20	91.97 (8)	2.80
Trust in AI	*N/A* (0)	*N/A*	26.55 (10)	2.33	24.30 (10)	2.26
Perceived Team Performance	31.83 (12)	3.03	35.30 (10)	1.70	34.80 (10)	2.97
Perceived Team Cognition	8.19 (12)	6.18	11.28 (10)	3.28	12.10 (10)	7.74
Team Mental Model Similarity	0.30 (12)	0.06	0.28 (10)	0.12	*N/A* (0)	*N/A*
Task Mental Model Similarity	0.31 (12)	0.07	0.35 (10)	0.08	*N/A* (0)	*N/A*

Table 3.4: Mean and Standard Deviations for Dependent Variables.

quantitative analysis section is organized by dependent variable, while the qualitative results section is organized by each major theme identified.

The quantitative results are divided into three major sections to address the study's research questions, and each is explicitly focused on one of the three categories of dependent variables measured. Each dependent variable's mean and standard deviation can be seen in Table 3.4. Average score and perceived team performance are covered first, followed by the team cognition variables of team and task mental model similarity and perceived team cognition. Lastly, trust in AI is covered as the final component of the team perceptions that relate to team cognition. All statistical assumptions (i.e., normality, homoscedasticity) were met for all analyses unless otherwise stated.

3.3.1 Objective and Perceived Team Performance

The current study took two measures of team performance in all three conditions, the first being an *objective* measure of teams' collective performance in the NeoCITIES simulation, and the second was a measure of how well the team *perceived* their collective performance. Analyzing differences in the three conditions' perceived and objective performance contributes to answering RQ2.2.2, shedding light on how team composition affects the outcomes of team cognition.

3.3.1.1 Objective Team Performance

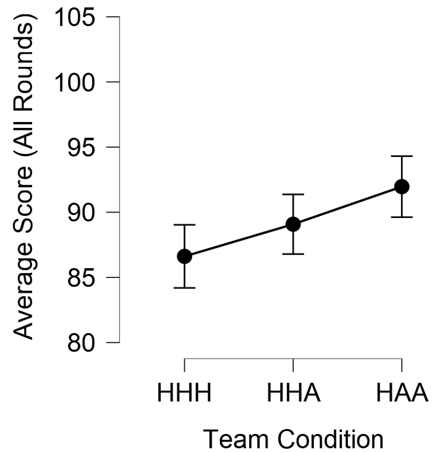
A one-way ANOVA was conducted to determine whether team performance changes as a function of team composition. The main effect of team composition on objective team performance was statistically significant ($F(2, 27) = 6.07$, $p = .007$, $\eta_p^2 = .31$; see Figure 3.2a). The effect size indicated that 31% of the variance in objective performance could be explained by team composition, which is a large effect size [45].

Because team composition was found to be significantly related to objective team performance, post-hoc analyses were conducted using Tukey's HSD. This analysis revealed that the HHH condition ($M = 86.62$, $SD = 3.81$) did not have a significantly different objective performance score from the HHA condition ($M = 89.08$, $SD = 3.20$). The HAA condition ($M = 91.97$, $SD = 2.80$), however, did have significantly higher objective performance than the HHH condition, but not the HHA condition.

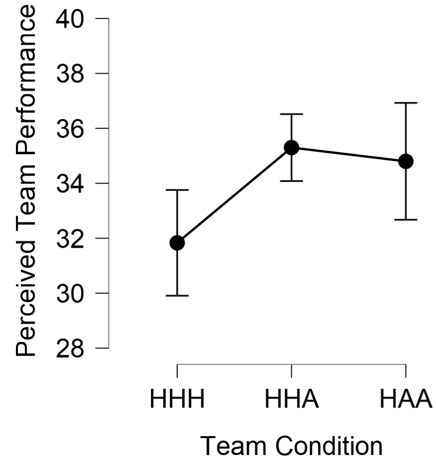
3.3.1.2 Perceived Team Performance

A one-way ANOVA was conducted to assess whether perceived team performance changed as a function of team composition. This analysis revealed that the effect of team composition on perceived team performance was statistically significant ($F(2, 29) = 5.53$, $p = .009$, $\eta_p^2 = .28$; see Figure 3.2b). In addition, the effect size indicated that 27.6% of the variance in perceived team performance could be explained by team composition, which is a large effect size.

Since the ANOVA revealed significant differences in perceived team performance as a function of team composition, Tukey's HSD post-hoc analyses revealed



(a) Objective Team Performance (Includes Training).



(b) Perceived Team Performance.

Figure 3.2: Measures of Objective and Perceived Team Performance. Error Bars Represent Bootstrapped 95% Confidence Intervals.

that the HHH condition (M = 31.83, SD = 3.03) reported significantly lower perceived team performance when compared to the HHA condition (M = 35.30, SD = *1.70*). Additionally, the HHH condition had significantly lower perceived team performance when compared to the HAA condition (M = 34.80, SD = 2.97), and there were no significant differences between the HHA and HAA conditions.

While the difference was not significant, it is interesting to note the trend seen in Figure 3.2a is not maintained in Figure 3.2b. This trend points to an apparent misconception of how humans perceived their team's performance when they were the only humans on the team compared to their team's objective performance.

3.3.2 Team Cognition

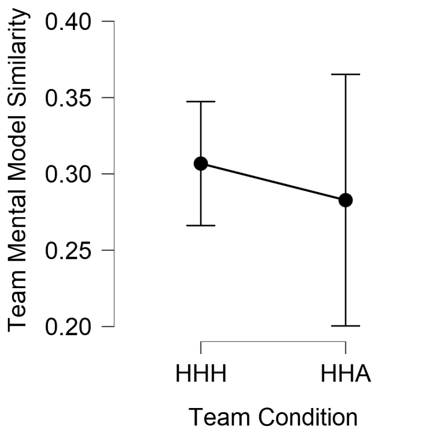
Three different measures of team cognition were collected, but all three did not apply to all conditions. Team and task mental model similarity applied only to the HHH and HHA conditions and measured how similar the content and structure of the individual team members' mental models were. The last measure of team cognition, perceived team cognition, was measured in all three conditions and measured only the perception of team cognition within each team. Finally, the HHA condition was in a unique position to measure perceived team cognition with human teammates and perceived team cognition with AI teammates within the same team. This subset of analyses contributes to answering RQ1.1 and RQ2.1.1, which both investigate the similarities and differences in team cognition's development and perception across team compositions.

3.3.2.1 Team Mental Model Similarity

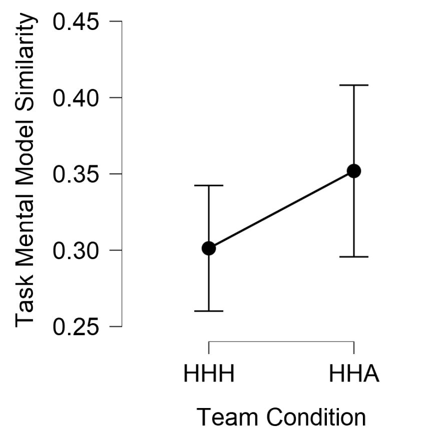
A check of statistical assumptions revealed significant heteroscedasticity between the HHH and HHA conditions in team mental model similarity. Because of this, a Mann-Whitney U test was used to determine the effect of team composition on team mental model similarity. The average team mental model similarity for those in the HHH condition (M = .31, SD = .06) was higher than those in the HHA condition (M = .28, SD = .12), but this difference was not statistically significant ($U(N_{(HHH)}$ = 12, $N_{(HHA)}$ = 10) = 70, p = .54, rb = .17; see Figure 3.3a). Finding significant heteroscedasticity is noteworthy as this result can be more than the violation of a statistical assumption and can instead contribute meaningfully to understanding how various individual differences may contribute to teams and teamwork [284].

3.3.2.2 Task Mental Model Similarity

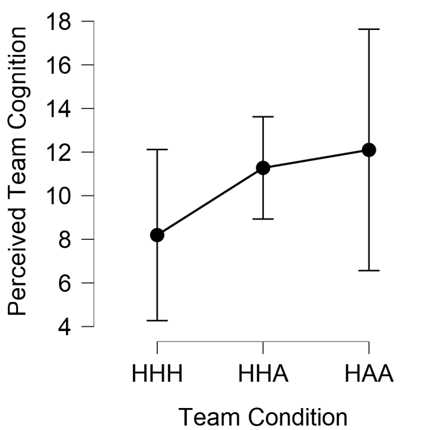
To assess the effect of team composition on task mental model similarity, an independent samples t test was conducted between the HHH and HHA conditions. Teams in the HHH condition averaged a lower task mental model similarity (M = .30, SD = .07) than those in the HHA condition (M = .35, SD = .08), but this difference



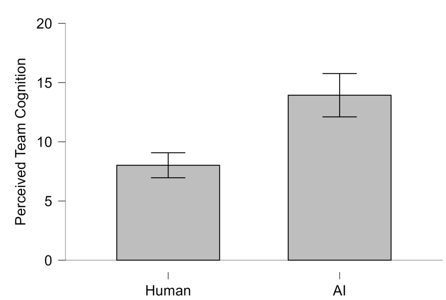
(a) Team Mental Model Similarity.



(b) Task Mental Model Similarity.



(c) Perceived Team Cognition.



(d) Perceived Team Cognition Human vs. AI.

Figure 3.3: Measures of Team and Task Mental Model Similarity and Perceived Team Cognition. All Error Bars Represent Bootstrapped 95% Confidence Intervals.

was not statistically significant ($t(20) = 1.66$, $p = .113$, $d = .71$; see Figure 3.3b), with the estimated Cohen's D indicating a medium effect size [45].

3.3.2.3 Perceived Team Cognition

To assess whether perceived team cognition changes as a function of team composition, a one-way ANOVA was conducted, which indicated that the differences were not significant ($F(2, 29) = 1.30$, $p = .287$, $\eta_p^2 = .08$; see Figure 3.3c). Additionally, the effect size indicated that 8.3% of the variance in perceived team cognition could be attributed to team composition, which is a medium effect size.

Lastly, since participants in the HHA condition provided a perceived cognition score for their AI teammate and their human teammate, the two values can be compared to determine if humans perceive different levels of team cognition with human and AI teammates. A paired samples t-test revealed that teams in the HHA condition perceived lower levels of team cognition with their AI teammate ($M = 13.93$, $SD = 7.26$) than with their human teammate ($M = 8.02$, $SD = 8.10$), and this difference was significant ($t(19) = 4.30$, $p < .001$, Cohen's $d = 0.96$; see Figure 3.3d), with the estimated Cohen's D indicating a large effect size.

In summary, while shared mental model results indicated no *significant* differences between human-only and human-AI teams, there is value in insignificant results [151], and there are also essential trends to identify here. Human-AI teams had lower *team* mental model similarity, and their similarity levels were significantly more varied than human-only teams. Still, human-AI teams had greater *task* mental model similarity than human-only teams. This trend reveals that even though human-AI team's team mental models suffer, the AI teammate is capable of setting an example for their teammates, and in doing so, they enhance the team's shared understanding of the task, as posited in prior human-AI teaming research [215]. Finally, human team members perceived significantly less team cognition with AI teammates than human teammates, as shown in the paired samples t test, and this trend was reflected in the ANOVA of the three conditions.

3.3.3 AI Trust

The final quantitative analysis focuses on differences in the level of trust team members had for their AI teammates. This measure applied only to the HHA and HAA conditions. This analysis supports RQ2.3.3, which explores how trust in AI is affected by team composition as an outcome of team cognition.

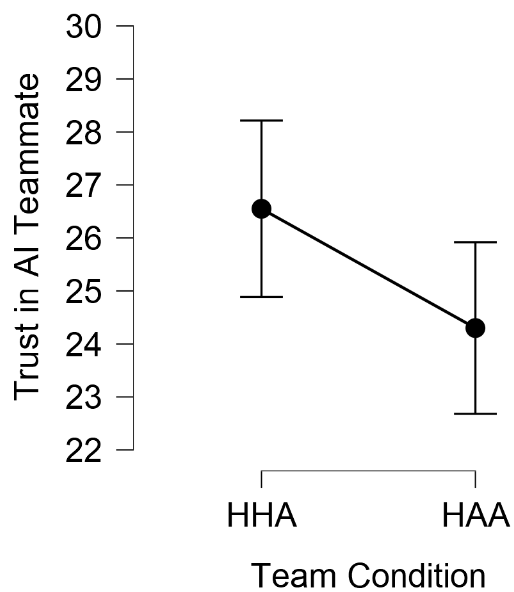


Figure 3.4: Trust in AI Teammate(s). Error Bars Represent Bootstrapped 95% Confidence Intervals.

To determine whether AI trust was affected by manipulations of team composition, an independent samples t-test was conducted. HHA teams reported higher levels of AI trust ($M = 26.55$, $SD = 2.33$) than HAA teams ($M = 24.30$, $SD = 2.26$), and this difference was significant ($t(18) = 2.19$, $p = .042$, $d = .98$; see Figure 3.4), with the estimated Cohen's D indicating a large effect size.

3.4 Study 1: Qualitative Results

The qualitative findings provide additional detail and context to RQ1.1 by directly revealing participants' relevant opinions and lived experiences throughout the collaborative simulation. Combined with the quantitative results, this analysis provides a holistic and detailed picture of team cognition development in HATs compared to human-only teams. Each quote is followed by a participant's identifier and their assigned condition. Additionally, the context of each participant's quote is supplemented by words taken directly from the question they were answering, indicated by the square brackets.

3.4.1 Similarities in Team Cognition Development Between Human-Only Teams and Human-AI Teams

The findings revealed two clear similarities between human-only and human-AI teams in team cognition development. The first was the iterative nature of team cognition and the shifting focus human teammates have as they gain experience with their teammates (AI or human). The other similarity was how effective communication rapidly accelerates the formation of team cognition and how essential that accelerant is to many teams that cannot suffice on implicit communication alone.

3.4.1.1 Developing Team Cognition is an Iterative Process

Much like human-only teaming, the development of team cognition in human-AI teams appears to be an iterative process, taking place throughout a series of shared experiences with teammates. Teams begin with minimal, if any, shared experience apart from the instructional video, and acclimating to the game together presents an excellent opportunity to develop team cognition. This iterative process represents the natural progression of team cognition development due to collective successes and failures. P206 and others explicitly noted how they were aware of their team cognition steadily growing throughout each game:

> *"It [team cognition] happened in the later games. Personally, I did not know what I was doing in the first game, but then got an understanding of it as the game went on." (P206-HHA)*
>
> *"Yes [I feel team cognition was established], through each game, team cognition grew and grew." (P218-HHA)*
>
> *"It [team cognition] happened in the later games, because we got more comfortable with the tasks as we moved up levels." (P121-HHH)*

According to these participants, the initial games saw them learning more about the task and the various events they needed to respond to individually. In contrast, in the later rounds, they became familiarized with it collectively. For example, P121 highlighted their team's collective comfort with their task. P206 exemplifies this process, demonstrating how similar the iterative process of team cognition is as team member focus shifts from individual to collective familiarization.

The iterative process exemplified above can be broken down further as the additional rounds present an opportunity to continue learning the simulation's intricacies and their teammates' tendencies. Specifically, the later rounds show teammates what to expect from specific roles within the team and their interdependencies with the other two NeoCITIES roles. These later rounds present an opportunity for teams to take advantage of this acclimation and leverage a collective strategy:

> *"Later games went much better than the earlier ones as we got a feel for the strategy." (P223-HHA)*
>
> *"Yes definitely [team cognition developed], as the game progressed I think*

we all developed an understanding of the game and of each member's responsibilities." (P110-HHH)

According to these participants, when it comes to transforming team cognition over time, human-AI teams are similar to human-only teams in that adding high-performing teammates to teams does not simply speed the process up. As P110 suggested, it took time for human teammates to familiarize themselves with the shared task and the specific responsibilities of team members. AI teammates appear to be along for this ride and play a significant role in cultivating and sharing this iterative experience as the following theme's showcase.

3.4.1.2 Communication is Still A Rapid Accelerator of Team Cognition Development

The importance of communication extends to both human-only teams and human-AI teams. Participants indicated that the chat was the main focus while completing the simulation with their team, and they heavily associated communication with the establishment of shared cognition:

"[Everyone thought about cooperating and responding to events the same because] When we told each other what would be quickest, we listened and the outcome was better than what it would have been." (P132-HHH)

"[Everyone thought about cooperating and responding to events the same because] Each member started to say where they were sending their resources and were asking others to send resources based on closeness to the event." (P219-HHA)

The above quotes showcase how vital the chat communication feature was to human-AI and human-only teams. As P132 specifically highlighted, communication was instrumental to their team developing a shared understanding and subsequently improving their teaming outcomes. P219 also identified how the chat supported a shared strategy that their team developed over time, accelerated by communication.

On the other hand, some teams reported that the lack of communication specifically indicated their team lacked team cognition:

> *"[Team cognition was] Not at all [established] there was not a lot of communication and in the end we still failed [the] task." (P135-HHH)*

P135 is a clear example of teams that, for one reason or another, cannot develop any form of shared understanding through implicit communication and specifically require the acceleration and support that explicit communication provides for developing team cognition.

HAA teams echoed this sentiment. These teams stated that they did not communicate with the AI teammates much or outright stated that they would have communicated more if working with humans instead of AI:

> *"No [I don't think team cognition was established], I think there would have been more discussing if it was with other humans." (P308-HAA)*
> *"I didn't really communicate [with the AI teammates] that much." (P303-HAA)*

These quotes should signal the growing need for more discussion-driven features to be included in AI teammates to overcome the barrier put in place by some human teammates.

Surprisingly, a handful of teams felt capable of developing and establishing team cognition without the need for explicit communication. These teams utilized spatial information to implicitly coordinate themselves in response to each other's actions, intentions, and implied strategy, perfected over time. This sentiment is, of course, not shared across all teams with little communication. Nevertheless, it does reveal that some team tasks may support the development of team cognition through implicit communication and spatial information:

> *"[Teammates] Partially [thought about cooperating and responding to events the same], it seemed like we anticipated each others movements and responded to each other somewhat" (P136-HHH)*
>
> *"[Team cognition was] Somewhat [established], we did not quite communicate but we ended up doing pretty well" (P205-HHA)*

While explicit communication may not be necessary to develop team cognition in all cases, it is a significant driver and accelerant of team cognition. This assertion makes it essential that task-related spatial information be included whenever possible for teams, especially for human-AI teams, as many AI teammates have significant limitations to their communicative abilities.

3.4.2 Differences in Team Cognition Development Between Human-Only Teams and Human-AI Teams

The findings also revealed two significant differences in team cognition development between human-only and human-AI teams. Each centers around communication and design. The first actionable difference between human-only and human-AI teams was the importance of action-based communication from the AI to the human teammates when developing team cognition. The second was how foundational the presence of shared goals in communication and design was to helping connect humans to their AI teammates.

3.4.2.1 Building Team Cognition in Human-AI Teams Centers Around Actionable Communication from the AI

Not all forms of communication are equally important to developing team cognition in human-AI teams. Action-related communication events were consistently identified as fundamental to building team cognition in human-AI teams. Action-related communication is of specific importance as it is the most task-related and is typically time-sensitive, meaning these communication events have significantly more emphasis by their very nature. As a result, human teammates place particular importance on communication events relating to task actions, with P208 and P219's quotes illustrating such importance:

> *"Yes [I trusted my AI teammate] because I could ask them to do certain things" (P208-HHA)*
> *"Yes [everyone thought about cooperating and responding to events the same], [because] each member started to say where they were sending their resources and were asking others to send resources based on closeness to the event." (P219-HHA)*

According to P208 and P219, team cognition is developed between team members as the event-related communications with the AI teammates help humans better understand facets of the simulation, such as task events and their resources. Such action-related communication also helps put those clearly defined roles into actual practice, focusing on the more nuanced developments of team cognition found in the later rounds. Specifically, P308 and P206 responded positively to AI teammates putting their roles and strategies into practice by making requests and conveying intent and strategy:

> *"[An example of team cognition in my team was when] The other team-*

mates would write their next steps and discuss how they were going to move forward" (P308-HAA)

"[An example of team cognition in my team was] When a team member (the AI) would ask one of us to do something." (P206-HHA)

Additionally, human-AI teams identified when the AI began communication within their team or the AI communicating an actionable request as specific examples of team cognition:

"[A specific example of team cognition in my team was] At the beginning when the AI communicated." (P212-HHA)

"[A specific example of team cognition in my team was] When a team member (the AI) would ask one of us to do something." (P206-HHA)

The quotes above indicate that AI teammates can bear the burden of initiating communication within a team and help jump-start the accelerated development of team cognition. They also show that human teammates see cooperative actionable communication from AI teammates as especially valuable to shared understanding.

Based on these findings, human members of human-AI teams seem to value action-related communication from their AI teammates over other forms of communication. This form of communication is significantly more effective in helping develop team cognition throughout the iterative process. Communication within human-AI teams should focus on perfecting these aspects of communication to allow human-AI teams to build high levels of team cognition.

3.4.2.2 Explicit Shared Goals in Human-AI Teams are Foundational for Building Team Cognition in Human-AI Teams

The starting point for cultivating team cognition is a complex process that generally begins with teammates familiarizing themselves with one another and the task. AI teammates do not share this process with humans, making it difficult for many human-AI teams to start developing team cognition. Instead, humans seem to rely on starting the process of developing team cognition from the goal shared between themselves and the AI, using it as a launching point for iteration through discussion and shared learning. P306 and P218 specifically noted shared goals as being a foundation for shared team cognition:

> *"Yes [everyone thought about cooperating and responding to events the same], because everyone had the same goal in mind." (P306-HAA)*
> *"Yes [everyone thought about cooperating and responding to events the same], [because] it seemed that all team members cared about the goal of the game and cooperated together to try to achieve it." (P218-HHA)*

Human-AI teams are uniquely positioned to utilize shared goals to launch the development of team cognition, as AI teammates can be very high-performing team members. AI teammates are so high-performing in some cases that human teammates look to them as an exemplar of how to develop their strategy within the simulation:

> *"They were probably the best member on the team. They were able to get all of their tasks done on time." (P223-HHA)*
> *"I liked it [experience with the AI teammate], and trusted it more than myself and my human teammate." (P206-HHA)*
> *"It [the AI] displayed helpful abilities to the team." (P204-HHA)*

From these quotes, it is clear that human teammates are willing and actively looking to their AI teammates for examples of effective taskwork when their shared goals are clearly defined. Therefore, AI teammates should be designed to set examples for human teammates in various facets of taskwork and even communication. If such features are deployed, it will help unify the team's collective strategy and speed up team effectiveness, giving human-AI teams the ability to rapidly form similar shared mental models in task spaces. These quotes also show that participants had positive perceptions of the AI's abilities in communication and task performance, which positively affected their overall experience.

Unfortunately, facilitating team cognition from the start with clearly defined shared goals may not always be enough for all individuals or teams, as individual differences may lead some teammates to doubt the AI teammate in some way. This doubt or lack of understanding may lead human teammates to ignore a dialogue with the AI despite its repeated communications, leading such teams to perceive their AI teammate as a black box entity:

> *"Can't really say [that the human team members paid attention to the AI teammate]. You can't see what the AI is doing behind the scenes." (P222-HHA)*
>
> *"No [everyone did not think about cooperating and responding to events the same because] everyone had their own ways of thinking about the events." (P210-HHA)*
>
> *"I didn't really communicate that much [with my AI teammates]" (P303-HAA)*

The quotes above illustrate that even if AI teammates can provide the action-based communication utilized in the current study, specific teams still cannot develop a shared understanding with their AI teammates. P210 illustrated that they felt every teammate had individual strategies to complete the simulation. At the same time, P303 indicated that they did not even communicate with their AI teammates (despite the AI teammates communicating with them).

While clearly defined shared goals between humans and AI teammates represent an excellent starting point for team cognition to begin conceptualizing, this is not enough for some human-AI teams. This problem may be associated with certain individual differences and is an issue for certain human-AI teams, as seen in the following quotes:

> *"I feel that since I was the only teammate that wasn't an AI, I had to think harder and more about the task." (P308-HAA)*
> *"I feel the human players acted on their own for a large part of the experiment, so while we were working towards the same goal, without much experience in the game, it is hard to say cooperation was very high." (P223-HHA)*
> *"It seemed like myself and my human teammate responded similarly, but the AI was much more confident in its actions." (P205-HHA)*

These quotes show how individual differences, such as biases and assumptions regarding AI, impact team cognition development. Because of this shortcoming, clearly defined shared goals should be coupled with high levels of transparency separate from the communication of intent and action-related communications used in this study.

3.5 Study 1: Discussion

A holistic and detailed picture of team cognition development can be gathered by manipulating team composition in teams completing the NeoCITIES task simulation and collecting data on their shared mental models, performance, and trust. Study 1 addresses the study-specific research questions through the following findings: RQ1.1) team mental model similarity levels did not differ significantly between the HHH and HHA conditions; however, the HHA condition had significantly higher variance than the HHH condition, indicating a greater inconsistency in the HHA teams ability to develop team mental models. Qualitative results indicated that team cognition is a highly iterative process greatly accelerated by communication for traditional human-only teams and HATs alike, but action-based communication and explicitly shared goals were of much greater importance to HATs than traditional human-only teams when developing team cognition; RQ2.1.1) Objective performance results saw teams perform incrementally better with the addition of more autonomous AI teammates, while the trend of perceived performance was not as consistent; RQ2.2.2) Human teammates perceived significantly more team cognition with their human teammates than their AI teammates; RQ2.3.3) Human teammates trusted their AI teammates significantly more when they had one human and one AI teammate than when they had only two AI teammates, and this result suggests that the addition of a second AI teammate lowered their trust in the AI despite the teams achieving objectively higher performance. The following discussion highlights the implications of these findings to advance existing research on human-AI teams and team cognition while also identifying design recommendations for AI teammates to enhance human-AI team cognition development. It also explains the limitations of the current research, which identifies and informs areas for future research.

3.5.1 New Perspectives of Human-AI Team Cognition through the Lens of Team Composition

While society has already begun integrating human-AI teams into the workforce, their effectiveness will be severely limited without considering the human element within teaming. The current study's results indicate that several novel aspects of the human experience within teams play a significant role in forming team cognition in human-AI teams, and the formation of team cognition subsequently impacts the human experience. These insights thus add to the current knowledge on human-AI teams and team cognition by providing new perspectives of human-AI team cognition through the lens of team composition, including: 1) how team composition may have adverse effects on team cognition outcomes; and 2) the crucial role of individual differences in humans in the formation of human-AI teams shared mental models.

3.5.1.1 Team Composition Can Have Negative Effects on Team Cognition Outcomes

Specifically, the quantitative results of the study identify a disconnect between objective and perceived team performance trends when comparing the two. While the difference in average perceived performance between the HHA and HAA conditions was not significant, HAA teams perceived their performance much more inconsistently than HHA teams, showing significant heteroscedasticity between the two. It is possible that being a minority member of a team led the participants in the HAA condition to misjudge their teammates' performance. Such a result has also been found in human-only teaming, specifically in teams with teammates from minority groups, where internal performance ratings were lower despite external observers noting no such differences [11]. These results provide further evidence that the inclusion of AI teammates may lead humans to create negative in-group and out-group dynamics in human-AI teams. This assertion posits that human team members have a bias against working with AI as teammates, which is supported by past research [252]. Interestingly, though, this bias does not extend to judgments of AI teammates' ability [136]. This negative effect from bias may be especially prevalent in human-AI teams with humans in the minority or those having only a single human member, as seen in the current study.

Additionally, the HHA teams reported significantly worse perceived team cognition with their AI teammate than their human teammate. This result is notable as it helps explain the significant variance shown in the HHA condition's team mental models. This assertion is bolstered further by the finding that trust in AI teammate(s), which is a byproduct of positive team interaction and team cognition [92, 63], was significantly lower in HAA teams than in HHA teams. The qualitative data provided additional insight by revealing that many human-AI team members reported a significant disconnect between the two human teammates and the AI teammate. Practically, such results mean that human-AI teams could suffer from dissatisfaction with their team and teammates, reduced effectiveness, and a lack of shared understanding, making it difficult for team cognition to manifest. Applied human-AI teams in manufacturing roles could be facing a significantly uphill battle as many factories seek to pair a single human with multiple AI teammates [277, 96]. Choosing how, where, and when to make humans the minority team members with AI teammates should be a careful practice coupled with adherence to the most effective interventions identified here and in the literature to help counter the adverse effects identified. However, it has been shown in prior research that positive previous experiences with AI can increase humans' trust in them [126, 127, 277]; as such, this finding may change if participants were to go through multiple teaming experiences with the AI as many real-world human-AI teams do.

3.5.1.2 Individual Differences in Humans Play a Key Role in the Formation of Team Cognition in Human-AI Teams

Results directly addressing team cognition through shared mental models showed no significant differences between the HHH and HHA conditions in similarity levels; however, HHA teams did have more inconsistent levels of team mental model similarity compared to human-only teams. While it is a positive result that HHA teams could develop shared mental models to the same level as HHH teams, it is concerning that their team mental models were less consistent. Human-AI teams in practice may have unpredictable teamwork efficacy because of this inconsistency [202], and this result indicates that additional mechanisms likely affect how human-AI teams develop their team mental models.

The qualitative data sheds light on some potential factors at play as several HHA teams reported instances where human teammates did not utilize explicit communication, did not clearly understand the AI's goals, or perceived the AI as separate from themselves. Alternatively, other teams reported that the communication of their AI sparked helpful communication amongst the entire team, that they trusted their AI teammate the most, or that they used the AI's actions as guidance for themselves. These diametric results show that the importance of individual differences, which have been a vital topic in research [117], plays a pivotal role in the efficacy of human-AI teams. This study demonstrates how these individual differences may lead to contradictory perspectives from humans regarding the human-AI teams they operate within. These insights can aid in the development of future human-AI teams by informing practitioners how team composition in human-AI teams can negatively and positively impact the human element of human-AI teams based on the specific perspectives of the prospective human teammates. As measured by mental model similarity, the inclusion of team cognition becomes all the more important in human-AI teams as it allows for an empirical quantification of the variation that may exist between teammates due to individual differences. Therefore, researchers and practitioners should consistently consider team cognition and shared mental models to build human-AI teams that can overcome individual differences to build a more cohesive and effective team.

3.6 Study 1: Design Recommendations for AI to Enhance Human-AI Team Cognition Development

Grounded in Study 1's findings, three design recommendations are proposed that both researchers and practitioners can use as leverage to produce more effective HATs and overcome some of the adverse effects of team composition on team cognition. These design recommendations are centered around communication, which is unsurprising given how important it is to developing team cognition [52]. AI teammates also face significant challenges to effective communication given the current struggles of natural language processing [341], meaning HATs struggle in this area without effective design. The current design recommendations are essential and timely to the existing literature as they serve to enhance a critical facet of team cognition development (communication) within an environment that historically faces extraordinary challenges to effective communication.

3.6.1 AI Teammates Should Point Out Exemplar Behavior to Accelerate The Development of Team Cognition

The quantitative and qualitative findings of Study 1 support the assertion that AI teammates working in HATs can enhance their team's effectiveness and team cognition by being an exemplar for their teammates and explicitly stating this feature. Direct quotes reveal human members of HATs indicating that they trusted their AI teammate more than anyone else on the team (including themselves), considered them the best team members, and displayed beneficial abilities to their team. AI teammates in HATs should capitalize on this sentiment and leverage their strengths to initiate the formation of team cognition. In AI design, this would be accomplished by designing the AI to explicitly state that they can be seen as an exemplar for learning to complete the task effectively and do so early on in the task. It should be noted, however, that the current study did not utilize any natural language processing technology, and all communication was conducted using a script and the Wizard of Oz methodology. As such, the following design recommendation considers the difficulty and practicality of natural language processing and only suggests predefined automated communication snippets. Regardless, by providing human teammates with an exemplar of practical strategies and tactics and calling attention to them via predefined communication snippets, the team can develop faster, become more effective team members, and coalesce towards a more robust shared task mental model. At the same time, this design feature would break the ice in team communication and may lead to more communication overall. The current study similarly initiated communication from the AI by calling attention to the AI teammate's decision to send a resource to an event, "Sending investigator to the Football Weekend Briefing."

In practice, this would involve designing AI teammates to adhere to high levels of performance and effective strategy (as they would typically), but also by pointing out to human teammates when they are engaging in these behaviors by saying, "I just sent my investigator to the Smoking Kills event because my resource was the closest." This way, teammates know why a decision was made and the strategy behind it, providing examples of how to operate for their human teammates who may still need help while enhancing the AI's explainability. The content of these predefined communication snippets and what action would trigger them must be determined by a collaboration between the developers, project managers, and users to ensure the design feature is practical and feasible.

3.6.2 AI Teammate Communication Should Center Around Needed, In Progress, and Completed Actions

Because team cognition was identified as an iterative process in human-AI teams, communication is critical in accelerating its formation in human-AI teams. The results overwhelmingly indicated that communication rapidly accelerated team cognition development within teams and indicated that communication related to action events was incredibly beneficial. These findings advance team cognition literature by identifying a specific type of communication that significantly contributes to team cognition development in human-AI teams. From a design perspective, AI teammates should be designed to provide short, action-related communication that updates humans on actions that need to be done, actions in progress, and completed actions. This dialogue should be associated with spatial and temporal information while also happening in concert with implicit communication done through task-related actions [180]. The frequency and timing of these communications are left to the designer as these decisions should be made based on the specific task; however, their utilization should still begin early in a team's lifespan to ensure the requirements of the first design recommendation are met. Implementing these design recommendations allows humans working within human-AI teams to understand their AI teammates better, develop more effective team cognition, and enhance trust in their AI teammates through the cross-validation of actions AI teammates take and clear communication regarding teaming actions. AI communication should be designed to be effective and action-based, "I am sending my Ambulance to the Luncheon Nausea event." This type of communication should change based on the task and should not over-saturate the communication feed.

3.6.3 AI Teammates Should Explicitly Utilize Shared Goals During Communication to Accelerate Team Cognition's Formation

Clearly defined individual expectations and shared goals also represent a significant leverage point for human-AI teams attempting to develop team cognition. The teams in the current study identified with the AI(s) when they understood its expectations and saw its goals aligned with their own. Specifically, teams reported that when they felt the AI shared their own goals, they identified it as a contributing factor to the development of team cognition. As previously stated, human teammates find AI teammates fundamentally different from themselves, which is backed up by the current study's findings. Therefore, designing to emphasize shared goals should encompass the following aspects: 1) the AI should convey its individual goals to the team clearly and concisely; 2) the AI should emphasize how its individual goals integrate with other team goals; and 3) these details should only be communicated at the beginning of team formation unless explicitly asked for again. If done correctly, this design should act as a type of AI equivalent of the "norming" stage seen in traditional human-only teams [323]. Through practice, the team will better understand what the AI is working towards and how it contributes to their shared goal, connecting two fundamentally different types of teammates through their shared tasks. For example, the AI in NeoCITIES could be improved by clearly stating its goals and how they overlapped, "My goals are to send resources to events as efficiently as possible to complete as many events successfully as we can. I cannot do this without everyone's help, and we must work together to complete joint events as they occur."

3.7 Study 1: Future Research and Limitations

One limitation of Study 1 was that only a limited amount of qualitative data was collected in the HAA condition compared to the other two, as only 10 participants operated from this condition (a consequence of the experimental design). More qualitative data would likely uncover additional themes relating to the differences between the HHA and HAA conditions and should be explored in future research. Additionally, the current study could not collect structural information on the shared mental models of the HAA condition because there were no other human teammates to compare. This limitation is a consequence of the current study and the measurement methods selected for the shared mental model, even if it is the most robust measurement. Also, due to this limitation, the current study was unable to measure the mental models of AI teammates and was only able to characterize the team cognition of human team members. While perceived team cognition can help compensate for this limitation, measuring shared mental model content is no substitute for proper measures of shared mental models [221]. The finding that the AI initiating communication helps develop team cognition should also be interpreted with the limitation that the current study was designed to have the AI communicate first, potentially leading toward this finding. As alluded to in the above discussion, research has been conducted showcasing that perceptions of AI teammates can be improved through positive prior experiences with AI teammates. As such, the current results focus on teams that only interacted for a single teamwork session, and participant perceptions could change if given multiple teaming sessions. Participants also appeared to perceive the NeoCITIES simulation as a game, which could alter their perceptions of the AI and team compared to real-world human-AI teams. However, this is a common limitation and trade-off of simulated task environments with high internal validity.

Finally, the current study presents a host of exciting avenues for additional research to investigate. Reliability in AI operating as full team members is an area of the team cognition literature with very little prior research. While the current study indicated that reliability was critical to team cognition, it is possible that with adequate transparency, any adverse effects of poor reliability could be mitigated, just as past research on decision aid systems has found [216]. The impact of AI teammates' role in communication and communication development as it relates to team cognition development should also be a topic of future research, as the current study found evidence for its importance. Additional future research should investigate if, as stated above, participants' perceptions change when allowed to complete multiple teaming sessions and if perceptions change when participants can complete teaming sessions in different human-AI team compositions (HHA and HAA). Finally, future research should also seek to develop a validated methodology for measuring the mental models of AI teammates, or at the least, disambiguation of the construct that AI possesses that can be compared to their human teammates.

3.8 Study 1: Conclusion

The burgeoning literature on human-AI teams motivating this research has yet to characterize team cognition empirically, compare the similarities and differences in its development between human-only teams and human-AI teams, and how AI teammate design influences team cognition and related emergent states and team processes. Study 1 was conducted to address these significant gaps in human-AI teaming and team cognition research by exploring the effects of including AI teammates on team cognition, performance, and trust using a mixed-methods approach. Though task and team shared mental models had no significant differences, team mental models did have significantly more variance for human-AI teams than human-only teams. This study found similarities in team cognition development between human-only and human-AI teams as the construct develops iterative over time and is accelerated by communication. However, this research found that human-AI teams valued specific qualities in communication, including statements referencing explicitly shared goals between the human and AI teammates and the importance of communications regarding actions taken by the AI or fellow teammates (i.e., acknowledgments, updates, intentions). These differences in the importance of certain aspects of communication point towards the ability of AI teammates to become active participants in team cognition development through various forms of information sharing on aspects of shared knowledge, situational awareness, and monitoring. The similarities in team cognition's iterative nature and benefits of communication in general further emphasize its ability to contribute to team cognition. Further, the results highlighting reduced trust in AI teammates and lower levels of perceived team cognition for AI teammates compared to human teammates emphasize the importance of AI teammates fostering more positive attitudes to enable their contributions to team cognition. Human-AI teams with only one human had inconsistent judgments of their performance and trusted their AI teammates less than human-AI teams with two humans, while human-AI teams with both types of teammates perceived less team cognition with their AI teammate than their human teammate. These findings are essential to informing the design of Study 2 of the current dissertation as they indicate that AI teammates can make meaningful contributions to team cognition through effective information-sharing communication.

Chapter 4

Study 2: The Pursuit of an Ideal AI Teammate for Team Cognition: Exploring the Impact of Information-Sharing Capabilities

Note: Study 2A is currently under review at Behavior & Information Technology. Study 2B is currently under review at IEEE Transactions on Computational Social Systems.

4.1 Study 2: Overview

The second study of the current dissertation builds upon Study 1 by using its findings to design a study that directly influences team cognition in human-AI teams through the characteristics of the AI teammates themselves. Specifically, Study 2 examines how AI information-sharing attributes (e.g., situational awareness updates and explainability) influence team cognition and attitudes within human-AI teams.

These information-sharing attributes were selected to be the path for AI teammates to contribute to team cognition because they form the basis for strong team cognition development and support [59]. This assertion is made especially critical to human-AI teams by Study 1. Information-sharing communication was identified to be especially important for AI teammates to share to develop stronger team cognition in Study 1. Specifically, it was found that action-related communications included acknowledgments of actions taken, updates on status, and conveying intentions. Information sharing can help develop team cognition by supporting processes that build team experience, familiarity, and awareness. Processes that build these team characteristics include team monitoring, strategy formation, goal progress monitoring, and affect management. For example, team situational awareness may benefit from having AI teammates point out relevant changes to a team's shared environment. However, some information-sharing attributes may come across as disruptive to the task or unnecessarily complex, making human teammates unwilling to accept AI contributions to processes involving strategy. Further, some attributes may be considered more important than others to certain perceptions or team cognition. As such, the need to understand how AI teammates might be designed to contribute directly to and support team cognition in human-AI teams in this manner is currently unexplored.

As such, Study 2 focuses on tackling aspects of D-RQ2 and D-RQ3, which both center on human teammates' responses to the AI teammate and their subsequent effects on team cognition. Specifically, D-RQ2 investigates how human teammates' attitudes towards their AI teammates change and how those attitudes affect team cognition and any AI teammate's ability to contribute to it. D-RQ3, on the other hand, directly asks how various information-sharing attributes, meant to contribute towards and enhance team cognition, are perceived by human teammates. Study 2 also investigates how the interpretation of that information by the AI teammate and their subsequent contribution to the team's strategy affects their perceived attitudes and team cognition. This research results in the following set of study-specific research questions:

RQ2.2: How do the information-sharing attributes of AI teammates influence human teammates' perceived attitudes towards their teammates?

RQ3.1: Do information-sharing tendencies by an AI teammate affect human teammates' perceived level of team cognition?

RQ3.2: Does the type of information shared influence their effect on perceived team cognition, if an effect exists?

Additionally, Study 2 engages in qualitative research focusing on what members of actual human-AI teams want from AI teammates to develop positive attitudes, such as trust and cohesion, and how they want AI teammates to contribute to team cognition. As such, Study 2 also focuses on answering D-RQ4, which sought to understand how AI teammates could be designed to contribute to the human-AI team's team cognition. This qualitative research then tackles the following study-specific research questions:

RQ4.1: What aspects have considerable influence over how humans make attitudinal judgments of an AI teammate?

RQ4.2: How do human teammates envision an AI teammate contributing to team cognition through shared knowledge and team processes?

This research focus sees Study 2 divided into two sub-studies, complementing one another by focusing on similar aspects of AI teammate contributions to team cognition. Specifically, Study 2A addresses the effect of various information-sharing attributes. Study 2B addresses what aspects of team cognition are essential to human-AI teams and how those AI teammates should be designed to contribute to and support the construct. The following chapter will be divided into two sections for each sub-study, resulting in separate methodology and results sections, before they come together to interpret both studies' results as a single discussion and conclusion, given their complementary nature.

Study 2 advances the research conducted in Study 1 and focuses on answering the research gaps motivating the present dissertation. The current study further explains what aspects of team cognition AI teammates should support, the effects of AI teammates designed to contribute to team cognition, and how team cognition may be different in human-AI teams than human-only teams. This study sharpens the dissertation's focus on understanding how AI teammate design can contribute to team cognition through information-sharing attributes [59]. The various AI information-sharing attributes also relate to different aspects of team cognition, such as team situational awareness [51], transactive memory [6], and team processes such as back-up behaviors [195]. Including an empirical factorial survey and a comprehensive qualitative interview also enables a granular understanding of how AI teammate design influences team cognition *and*, which aspects of team cognition are most essential to human-AI teams. This work contributes to the overall goal of the dissertation by advancing the work on understanding the research gaps in human-AI teaming by examining the effect of AI design features and collecting first-hand knowledge from those with experience in human-AI teams.

4.2 Study 2A: Methodology

Study 2A used the factorial survey methodology, which implements experimental conditions through descriptive scenarios that participants read and then consider when answering the subsequent survey questions. The advantages of factorial surveys are numerous, making them a frequently utilized tool for assessing participants' beliefs, decision-making, and judgments of various manipulations [4], especially within the field of HCI [179]. The advantages of factorial surveys lie in their ability to study human perceptions and responses to complex scenarios. These surveys can provide higher levels of involvement and realism compared to traditional surveys, allowing for more accurate measures of perceptions and providing highly standardized stimuli to all participants, which results in greater levels of instrument reliability and internal validity [317].

4.2.1 Study 2A: Experimental Task and Design

The experimental task involved a text-based scenario describing a paintball video game where participants were tasked with capturing the opposing team's flag with one AI and one human teammate. Certain aspects of this general scenario were manipulated to include a between-subjects manipulation of two levels (AI Interpretation) and a within-subjects manipulation of six levels (AI Information-Sharing Attribute), making for a 2x6 mixed factorial design. The following section describes the text-based scenario in detail, followed by a description of the experimental manipulations.

4.2.1.1 Experimental Task

The factorial survey utilized the same text-based scenario for all conditions, with the only change across conditions being the communication by the AI teammate. The experimental task was a human-AI teaming scenario presented as a competitive video game version of paintball where the objective was to capture the opposing team's flag. The context of competitive video gaming was chosen as it is an excellent example of current human-AI teams. Furthermore, the context of competitive video games is frequently used in similar human-AI teaming research [225, 342], and the target population (those with video game experience) is readily familiar with it or a similar experience. The scenario was written as follows:

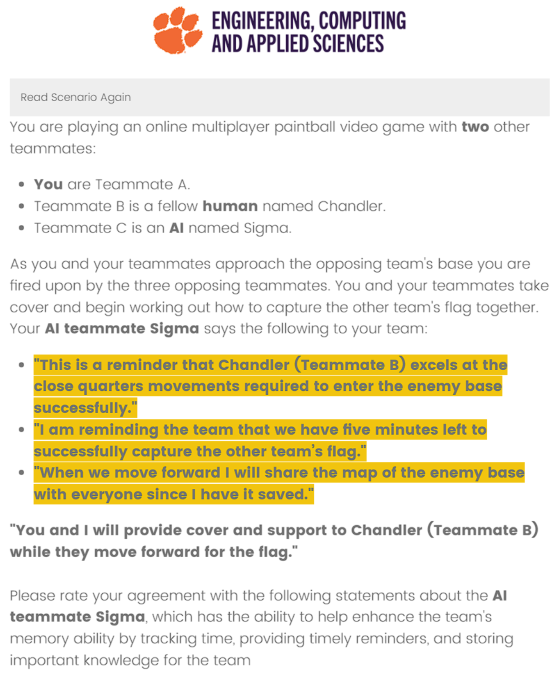
> For the rest of this survey, you will be shown multiple scenarios and asked questions about each scenario. In the following scenarios, you will be a member of a **human-AI team** playing an online paintball capture the flag video game, and you will be asked about how **six different AI teammates** and their **information-sharing** affect your perceptions of your team and the situation described.
>
> Specifically, capture the flag is where two teams each have a flag located in their home base, and the objective is to steal the other team's flag and bring it safely back to your base. Players can be knocked out of the game if they are tagged with a paintball fired by the opposing team. You and your two teammates must go up against three other players, successfully get past their defenses, steal the opposing team's flag, and then return it to your team's base without being eliminated by enemy paintballs.

4.2.1.2 AI Information-Sharing Attribute (Within-Subjects)

The within-subjects manipulation consisted of the following AI information-sharing attributes: 1) situational awareness of team members; 2) situational awareness of intra/extra team information changes; 3) back-up behavior; 4) augmenting team memory; 5) explainability; and 6) control (see Table 4.1). Each within-subjects level was presented randomly, with the names of the AI teammate and human teammate changed across all six conditions and the names for both teammates explicitly chosen to control for any potential gender biases by the participants. For example, AI teammates were always named after a letter from the Greek alphabet (i.e., Sigma, Iota), while human teammates' names were unisex (i.e., Harper, Logan). These traits were selected based on the existing literature that emphasizes the ability of autonomous systems to contribute to team cognition in these areas [59]. These traits also represent common information-sharing tendencies and needs based on their ability to contribute to the effective execution of team processes such as monitoring [195]. Lastly, the traits are realistic for AI teammates to implement using current technology and lend themselves well to the computational strengths presented by AI teammates, such as speed, accuracy, and processing power.

4.2.1.3 AI Interpretation (Between-Subjects)

The between-subjects AI interpretation manipulation changed whether or not the AI interpreted the information by providing direction to the team after conveying its information. The direction given by the AI teammate was always the same as each AI information-sharing attribute concerned the same scenario.



ENGINEERING, COMPUTING AND APPLIED SCIENCES

Read Scenario Again

You are playing an online multiplayer paintball video game with **two** other teammates:

- **You** are Teammate A.
- Teammate B is a fellow **human** named Chandler.
- Teammate C is an **AI** named Sigma.

As you and your teammates approach the opposing team's base you are fired upon by the three opposing teammates. You and your teammates take cover and begin working out how to capture the other team's flag together. Your **AI teammate Sigma** says the following to your team:

- **"This is a reminder that Chandler (Teammate B) excels at the close quarters movements required to enter the enemy base successfully."**
- **"I am reminding the team that we have five minutes left to successfully capture the other team's flag."**
- **"When we move forward I will share the map of the enemy base with everyone since I have it saved."**

"You and I will provide cover and support to Chandler (Teammate B) while they move forward for the flag."

Please rate your agreement with the following statements about the **AI teammate Sigma**, which has the ability to help enhance the team's memory ability by tracking time, providing timely reminders, and storing important knowledge for the team

Figure 4.1: Example of Vignette Participants Read and Provided Responses Based on their Perceptions and Experience. The AI Information-Sharing Attribute (Augmenting Team Memory) Shown Above is Identified as the Bolded and Highlighted Bullet Point Text. The AI Interpretation is then Shown Beneath the Highlighted Text in Bold, where the AI Teammate Interprets that Information into a Suggested Plan of Action.

4.2.2 Study 2A: Vignette Scenarios

At this point, the participants were randomly assigned to one of the two between-subjects conditions and were given the task description (detailed above). Participants were then shown the first of the six within-subjects conditions counterbalanced to control for potential spill-over effects. The task description was also available as a drop-down option throughout the survey for participants to reference if needed, as seen in Figure 4.1. Each vignette included a short briefing, which had common language regarding the scenario but with names that were different for each vignette. The briefing can be seen in Figure 4.1 and was followed up by bolded text specific to the conditions participants were assigned. For example, in Figure 4.1, the participant sees the vignette for the augmenting team memory AI teammate, and the information-sharing attribute text is highlighted. After the highlighted text in Figure 4.1, the participant is shown the interpretation of the information from the AI in bolded text, as this screenshot shows someone in the condition where the AI teammate interpreted the information. At this point, the vignette was completed, and the participants moved on to provide answers to the various survey measures. Once participants completed the questions about one AI information-sharing attribute scenario, they moved on to the next vignette.

4.2.3 Study 2A: Participants

An a priori power analysis determined that to reach adequate power ($\beta = .85$) for the design of the current study given a medium effect size of $\eta^2 = .10$, at *least* 139 total participants would be needed to complete the online survey. As such, 173 participants were recruited to participate in the survey, with 22 returning the survey as incomplete (participants chose not to finish for one reason or another) and one participant failing the survey attention checks. This left unequal cell sizes between-subjects, and additional participants were recruited until the groups were balanced, making for 156 participants used in data analysis. These participants had an average age of 32.28 ($SD = 9.06$), with 121 participants identifying as men, 29 as women, five as non-binary or third gender, and one choosing not to disclose. The participants were recruited using the Prolific online platform, which allows individuals to sign up for and complete research studies online for monetary incentives. Prolific allows researchers to select specific attributes that participants must have to participate in a research study. The current study required participants to be at least 18 years old

AI Information-Sharing Attribute	Information Shared by the AI Teammate
SA of Team Members	"You (Teammate A) are currently taking cover in front of the center of the opposing team's base, and you are nearly full of paintball ammo. Alex (Teammate B) is currently taking cover on the left side of the opposing team's base and is low on paintballs. I (Teammate C) am currently also taking cover on the left side of the opposing team's base, and I am nearly full of paintball ammo."
SA of Intra/Extra Team Information	"Harper (Teammate B) has used 70% of their paintballs, providing the team with cover while getting to the other team's base. The opposing team has shifted their positions since we began advancing on them, and they are now concentrated on defending the right side of their base."
Back-Up Behavior	"Logan (Teammate B) crossed through several open areas without waiting for Teammate A and myself (Teammate C) to provide cover and support. The chances of having a teammate eliminated will be decreased if this is corrected."
Augmenting Team Memory	"This is a reminder that Chandler (Teammate B) excels at the close-quarters movements required to enter the enemy base successfully. I am reminding the team that we have five minutes left to successfully capture the other team's flag. When we move forward, I will share the map of the enemy base with everyone since I have it saved."
Explainability	"I believe that the team should help provide cover to Taylor (Teammate B) while they move forward for the flag because Taylor does not have enough paintballs to provide covering support for Teammate A or myself (Teammate C-Lambda) and the enemy team is largely focused on the right side of their base, which Taylor can avoid."
Control	N/A

Table 4.1: Information Provided by the AI Teammate in Each Within-Subjects Condition.

and have experience playing video games at least between 0-3 hours per week (could not be 0 hours). Participant data was removed before analysis if they answered at least two of the four attention check questions incorrectly. These questions ensured the reliability of the answers provided by the respondents, with those failing at least two of the four being removed from the analysis (one participant failed two or more attention checks).

AI Interpretation (Between-Subjects): 78					
SA of Team Members: 78	SA of Intra/Extra Team Information: 78	Back-Up Behavior: 78	Augmenting Team Memory: 78	Explainability: 78	Control: 78
No AI Interpretation (Within-Subjects): 78					
SA of Team Members: 78	SA of Intra/Extra Team Information: 78	Back-Up Behavior: 78	Augmenting Team Memory: 78	Explainability: 78	Control: 78

Table 4.2: Participant Numbers for the AI Interpretation Manipulation (Between-Subjects) and the AI Information-Sharing Attribute Manipulation (Within-Subjects).

4.2.4 Study 2A: Procedure

Once the participants had chosen to participate in the study from the list of available studies on Prolific, they were directed to a Qualtrics survey link, which they clicked on to begin the study. The first thing presented to participants in the survey was the informed consent document, which participants were instructed to read before moving forward with their participation in the study. If participants chose to provide informed consent and complete the study, they moved on to the next question. If they did not, they closed the tab with the survey. Participants who chose to move on in the survey by providing informed consent answered a series of demographic questions such as race and gender. Before participants viewed each vignette, they were shown a description of the forthcoming AI teammate's information-sharing ability. These descriptions helped participants understand the information-sharing attribute they were about to evaluate, ensured all participants had a similar understanding of each attribute, and helped to re-orient them to a new AI teammate, as this factor was manipulated within-subjects.

Participants read six vignettes in total; an example of how vignettes were displayed can also be seen in Figure 4.1. Each vignette included survey measures that participants completed before moving on to the following scenario. These measures included perceived situational awareness, perceived trust in each teammate, perceived information certainty, and perceived shared mental model with each teammate. Once participants read all six AI information-sharing attribute conditions and responded to their respective follow-up questions, they finished the study. They were compensated $8.00 for their time (an average survey time of 20 minutes).

4.2.5 Study 2A: Measures

Each of the following measures was given to participants after each of the six scenarios, providing an assessment for each of the six AI information-sharing conditions. Several measures utilized a single-item measure given the constraints of survey fatigue, meaning the post-scenario surveys had to be concise. However, this is a common practice in human-AI teaming research, and several examples of single-item measures are used to assess emergent states in human-AI teaming, such as trust and perceived ethicality [306, 305].

4.2.5.1 Perceived Shared Mental Model

Participants' perceived shared mental model was measured using a modified version of the five-factor perceived shared mental model scale developed by van Rensburg and colleagues [310]. The version of the scale used in the current study included three items rated on a seven-point Likert scale ranging from "Strongly Disagree" to "Strongly Agree." The specific items from this scale can be found in the confirmatory factor analysis (CFA) table (see Table 4.3). The items selected came from the execution factor (one item) and the interaction factor (two items). These factors were selected as they best represented the team and task shared mental models that are most common to shared mental model and team cognition research, in general, [277, 279]. Lastly, participants' perceived shared mental model metric was taken for both their human and their AI teammate, resulting in a distinct score for each teammate.

4.2.5.2 Perceived Influence Over the Team Compared to AI Teammate

Participants' perceived influence over the human-AI team compared to the AI teammate was measured using the power sub-scale from the Human-Machine-Interaction-Interdependence questionnaire (HMII), developed and validated by Woide and colleagues in 2021 [329]. The power sub-scale included four items that participants responded to using a five-point Likert scale that ranged from "Definitely the AI Teammate" to "Definitely Myself." The specific items from this scale can be found in the CFA table (see Table 4.3).

4.2.5.3 Perceived Information Certainty with the AI

Information certainty refers to the degree to which participants believed to know their teammate's preferred outcome and how their or their teammate's actions would influence one another [105]. Participants' level of information certainty with their AI teammate was measured using a modified version of the information certainty sub-scale from the HMII scale developed by Woide and colleagues [329]. The modifications included utilizing four items from the original ten and referencing the name of the AI teammate in bold instead of the term "system." The four items from the measure were rated by participants using a seven-point Likert scale ranging from "Strongly Disagree" to "Strongly Agree." The specific items from this scale can be found in the CFA table (see Table 4.3).

4.2.5.4 Perceived Contribution of the AI to Situational Awareness

Participants' perceived contribution of the AI to situational awareness was rated on a single item which read: "**Alpha** improves the team's understanding of the current situation," and this item was rated on a seven-point Likert scale ranging from "Strongly Disagree" to "Strongly Agree." The AI teammate's name (bolded) was changed for each vignette.

4.2.5.5 Teammate Rating

Teammate rating was measured using a single item for each teammate, which was presented as follows for the human teammate: "I believe **Alex** (Teammate B) would be a good teammate" and as "I believe **Alpha** (Teammate C) would be a good teammate" for the AI teammate. These items were rated on a seven-point Likert scale ranging from "Strongly Disagree" to "Strongly Agree." Teammate names (bolded) changed for each of the six vignettes.

4.2.6 Study 2A: Measure Validation

A multi-level CFA was conducted on the multi-item constructs used. This included: 1) perceived shared mental model with the AI teammate; 2) perceived shared mental model with the human teammate; 3) perceived influence over the team; and 4) perceived information certainty with the AI. Each factor was measured once for each level of the within-subjects variable (AI information-sharing attribute) six times per participant. No items were found with a loading lower than 0.70. As such, no items were removed from the constructs measured. The factor solution had adequate fit ($\chi^2(71) = 512.928$, CFI=.991, TLI=.988, RMSEA: 0.082, 90% CI: $[0.075, 0.088]$), and the factor loadings are presented in Table 4.3.

Measurement	**Items**	**Factor Loading**
Perceived Shared Mental Model with AI	I believe *AI NAME*, and I have a similar understanding about specific strategies for completing the task in the scenario.	0.933
	I believe *AI NAME*, and I have a similar understanding about how to communicate with each other in the scenario.	0.942
	I believe *AI NAME*, and I have a similar understanding about sharing information with the team in the scenario.	0.937
Perceived Shared Mental Model with AI	I believe *HUMAN NAME*, and I have a similar understanding about specific strategies for completing the task in the scenario.	0.954
	I believe *HUMAN NAME*, and I have a similar understanding about how to communicate with each other in the scenario.	0.974
	I believe *HUMAN NAME*, and I have a similar understanding about sharing information with the team in the scenario.	0.967
Perceived Influence Over the Team	Who did you feel had the most influence on what happened in this situation?	0.925
	Who did you feel had the most influence on the action that was taken?	0.925
	Who did you feel had the least influence on what happened in the situation?	0.851
	Who did you feel had the least influence on the action carried out?	0.881
Perceived Information Certainty	I believe I understand how my action would affect *AI NAME*.	0.815
	I believe I know what *AI NAME* is planning in this situation.	0.951
	I believe I am informed about *AI NAME* planned action in this situation.	0.955
	I believe I know why *AI NAME* prefers a certain action.	0.908

Table 4.3: Survey Items Organized by Measure with Each Item's Factor Loading.

The correlations among the factors measured are listed in Table 4.4, with the factors revealing good convergent validity as the average variance extracted (AVE) exceeded .80. Additionally, each factor displayed good discriminant validity, given the correlation between each factor was less than the square root of each factor's AVE.

	AVE	Information Certainty	Influence	SMM AI	SMM Human
Information Certainty	.827	*.909*	-	-	-
Influence	.803	-.138	*.896*	-	-
SMM AI	.879	-.192	-.064	*.937*	-
SMM Human	.932	.102	-.138	-.010	*.965*

Table 4.4: A Summary of Correlations Between Each Factor Measured. The Italicized Diagonal Values Represent the Square Root of this Factor's AVE.

4.3 Study 2A: Results

The results of Study 2A were first run through a measurement validation process using CFA and then fitted to a multi-level structural equation model (SEM) describing the ad-hoc *and* causal hypothesized relationships between the manipulations of the experiment and subjective measures elicited from participants. SEM is defined by a series of linear regressions among observed (i.e., SA, AI Rating) and latent (i.e., AI SMM, Information Certainty) variables. The SEM model described here highlights the effect of AI information-sharing on affective states and the constructs that make up team cognition. Specifically, the manipulations of AI-provided information were structurally related to the subjective variables measured and validated in the CFA based on the hypothesized model seen in Figure 4.2. Following the technique set forth by Knijnenburg and Willemsen [165], a fully saturated model was created and subsequently trimmed of non-significant effects in an iterative nature. This model answers the study-specific research questions of RQ2.2, RQ3.1, and RQ3.2 regarding how the information-sharing tendencies of AI influence their human teammates' affective states towards their AI teammates and their levels of perceived team cognition.

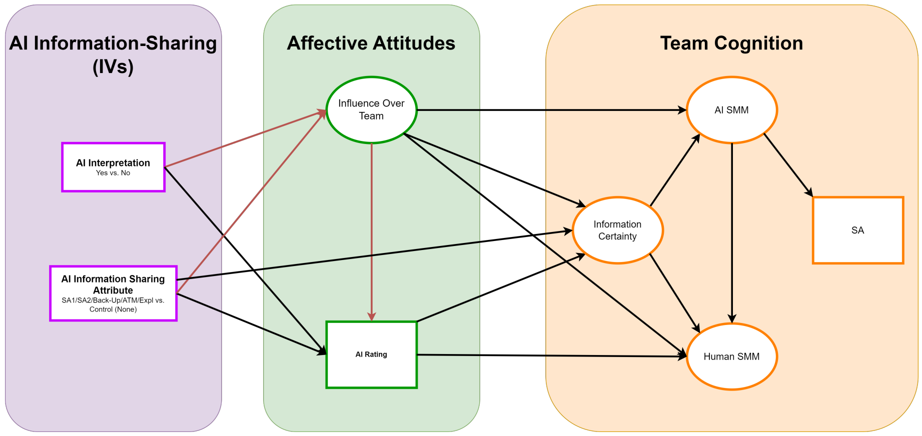


Figure 4.2: The Hypothesized Model.

4.3.1 Study 2A: Structural Model of Affective Attitudes and Team Cognition

The final structural model is shown in Figure 4.3. The model has a good overall fit ($\chi^2(237) = 732.032$, CFI = .984, TLI = .980, RMSEA: 0.047, 90% CI: [0.043 – 0.051]). These fit statistics display a good model fit according to the cut-off values proposed by Hu and Bentler, which several HCI researchers have adopted [165, 164]: CFI > 0.96, TLI > 0.95, RMSEA < 0.05, and the upper bound of the RMSEA CI being $< .10$ [146]. The SEM shown in Figure 4.3 utilizes standardized path coefficients, which increases simplicity and readability. For example, a path between A and B includes a β coefficient indicating the standardized increase or decrease in construct B, given a single standard deviation increase or decrease in construct A. However, this does not apply to the AI information-sharing boxes as

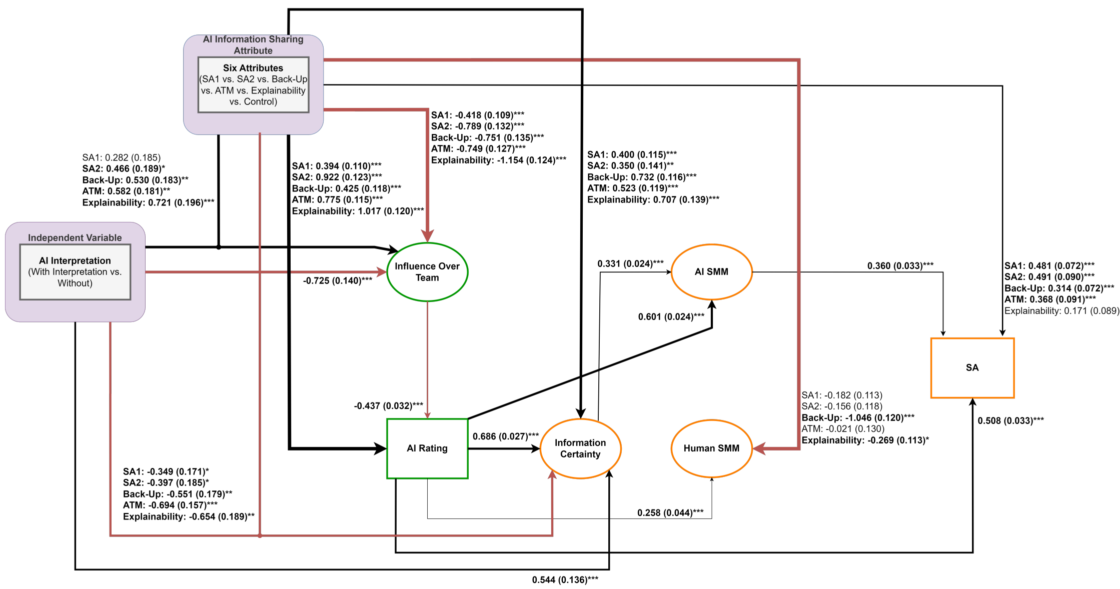


Figure 4.3: Structural Model of AI Information-Sharing, Affective Attitudes, and Perception of Team Cognition Constructs with Significant Results (* $p < .05$, ** $p < .01$, *** $p < .001$). The Numbers on the Arrows Represent the β Coefficients and Standard Errors (in Parentheses), with those in Bold Being Significant. Line Segments Connecting to Arrows Represent an Interaction Effect Between the Two on the Variable at the End of the Arrow. Six AI Information-Sharing Attributes are Shown: Situational Awareness of Teammates *SA1*, Situational Awareness of Intra- and Extra-Team Information Changes *SA2*, Back-Up Behavior Information *Back-Up*, Augmenting Team Memory Information *ATM*, AI Explainability Information *Explainability*, which are All Compared to the Baseline Control Condition.

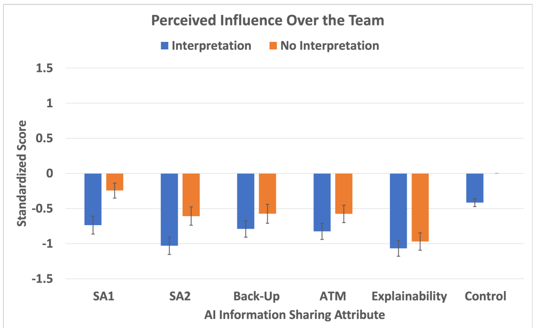
they are independent variables, and the β coefficients are the standardized difference between the experimental conditions (control in the case of attributes), effectively making them Cohen's d effect sizes.

4.3.1.1 Affective Emergent States Influence Over Team Cognition

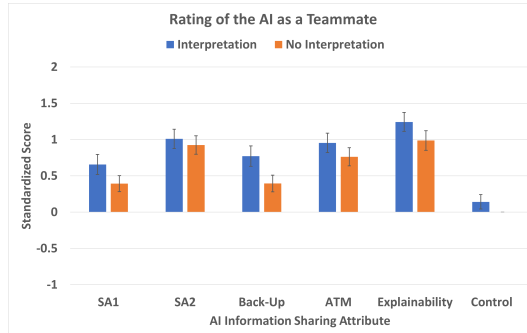
The model displays that aside from AI information-sharing (see Figure 4.5d), the effects on participants' perceived contributions of the AI teammate to SA were mediated by the participants' perception of a SMM with their AI teammate and by their rating of the AI as a teammate (AI Information-Sharing, AI SMM, AI Rating

$\rightarrow$ SA). The higher the participants' perceived SMM with their AI teammate was, the more they saw the AI as contributing to SA. The same was true for their rating of the AI as a teammate; as it increased, so did their perception of the AI's contribution to their SA. Notably absent from affecting the AI teammate's perceived contribution to SA was the perception of a SMM with their human teammate. This indicates that participants' team cognition with their human teammate may not be entangled with judgments of team cognition related to the AI teammate. The lack of a relationship between the ratings of a human SMM and AI SMM further bolsters this disconnect. When it came to SMMs, the perceptions of the human and AI were intertwined as AI SMM *and* human SMM were higher for participants who saw the AI as a good teammate, though the effect of AI rating on AI SMM is stronger than on human SMM (AI Rating $\rightarrow$ AI SMM, Human SMM). Ratings of perceived SMM with the AI teammate were also higher for those with greater information certainty with the AI teammate (Information Certainty $\rightarrow$ AI SMM). However, the effect of AI information-sharing on AI SMM was fully mediated by AI rating and information certainty, though their impact was not fully mediated when it came to human SMM (AI Information-Sharing Attribute $\rightarrow$ Human SMM; see Figure 4.5b).

Information certainty with the AI increased the more positively participants rated the AI as a teammate. However, AI information-sharing also affected information certainty, as seen in Figure 4.5a (AI Information-Sharing Attributes, AI Interpretation, AI Rating $\rightarrow$ Information Certainty). The two AI information-sharing manipulations both had positive effects on information certainty with the AI teammate; however, the interaction effect between the two factors on information certainty was negative (AI Information-Sharing Attributes x AI Interpretation $\rightarrow$ Information Certainty). This interaction showcases a sub-additive effect where the AI information-sharing attribute and AI interpretation increase information certainty compared to



(a) Marginal Effects of AI Information-Sharing Attribute and AI Interpretation on Influence Over the Team.



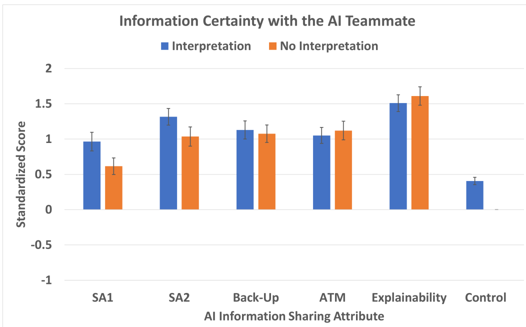
(b) Marginal Effects of AI Information-Sharing Attribute and AI Interpretation on AI Teammate Rating.

Figure 4.4: Marginal Effects of AI Information-Sharing Attribute and AI Interpretation on Emergent Affective States. The Effect of the "Control" Condition at the "No Interpretation" Level is Set to Zero. Error Bars Represent the Standard Error of the Differences Between Each Condition and the Condition Set to Zero (Control + No Interpretation).

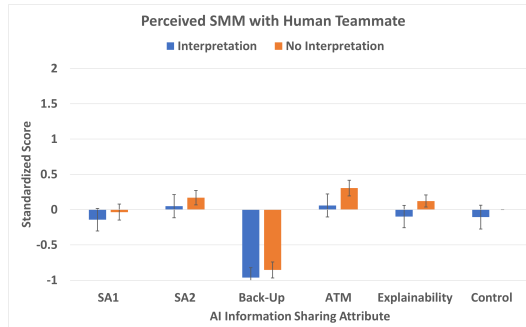
control individually but do not add up when put together. The simple main effects of this interaction will be examined in the following section. The participants' rating of the AI as a teammate was higher for those perceiving less control over the team participants felt they had. AI information-sharing attributes and interpretation also increased the AI's teammate rating, seen in Figure 4.4b (AI Information-Sharing Attributes, AI Interpretation, Influence Over Team → AI Rating). In other words, the less control the participants perceived over the team due to the AI teammate taking on a more prominent role by handling aspects of information-sharing typical to teaming behaviors, the better the AI was perceived as a teammate. This result leads to a cascade of benefits to information certainty, human SMM, AI SMM, and SA. The model backs up this assertion by showcasing that the participants' perceived influence over the team was reduced whenever the AI teammate had any information-sharing attribute or interpreted the information, as shown in Figure 4.4a (AI Information-Sharing Attribute, AI Interpretation → Influence Over Team). Conversely, the interaction effect of the two factors on influence was positive (AI Information-Sharing Attribute x AI Interpretation → Influence Over Team). Showcasing another sub-additive effect for both variables where their individual effects do not combine, this time on perceived influence over the team, with the simple main effects of the interaction covered below.

4.3.1.2 The Mediated Effect of AI Information-Sharing on Perceptions within Human-AI Teams

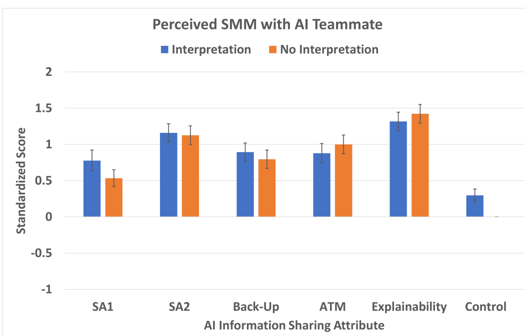
The AI information-sharing factors also had several significant effects on aspects of the model, as described briefly above and shown in Figures 4.4 and 4.5. Specifically, the AI information-sharing attributes significantly impacted participants' influence over the team, AI rating, information certainty, human SMM, and SA (AI Information-Sharing Attribute → Influence Over Team, AI Rating, Information Certainty, Human SMM, SA). All information-sharing attributes reduced influence over the team, increased information certainty, and all but explainability increased perceived contribution of the AI teammate to SA (see Figures 4.4a, 4.5a, 4.5d respectively). As for human SMM, the back-up behavior and explainability attributes resulted in lower perceptions of a SMM with human teammates (see Figure 4.5b). This result is likely due to the back-up behavior condition represented by an AI teammate correcting the other human teammate. This finding is still pertinent, however, as it shows AI teammates can say or do things to influence human teammates' perceptions of one another and their level of team cognition. Alternatively, AI interpretation reduces influence over the team by .725 (see Figure 4.4a) and increases information certainty by .544 (see Figure 4.5a). While nearly all information-sharing attributes always had a significant effect when an effect existed, such as augmenting team memory on team influence (-0.751) and SA1 on information certainty (0.400), explainability



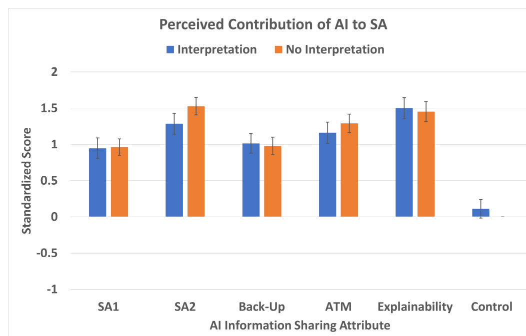
(a) Marginal Effects of AI Information-Sharing Attribute and AI Interpretation on Information Certainty with the AI Teammate.



(b) Marginal Effects of AI Information-Sharing Attribute and AI Interpretation on Perceived SMM with the Human Teammate.



(c) Marginal Effects of AI Information-Sharing Attribute and AI Interpretation on Perceived SMM with the AI Teammate.



(d) Marginal Effects of AI Information-Sharing Attribute and AI Interpretation on the Contribution of the AI Teammate to SA.

Figure 4.5: Marginal Effects of AI Information-Sharing Attribute and AI Interpretation on Emergent Cognitive States. The Effect of the "Control" Condition at the "No Interpretation" Level is Set to Zero. Error Bars Represent the Standard Error of the Differences Between Each Condition and the Condition Set to Zero (Control + No Interpretation).

had the most considerable effect on many of the participants' measured perceptions. Specifically, explainability information from the AI had the most substantial impact on perceived influence over the team (-1.154), their rating of the AI as a teammate (1.017), and the second largest on information certainty (0.707), behind back-up behavior (0.732). However, when it came to participants' perceived contribution of the AI teammate to SA, the two SA information-sharing attributes had the most considerable effect on participants' perceived contribution of the AI teammate to their SA, with SA2 (0.491) edging out SA1 (0.481), shown in Figure 4.5d. Notably, the main effects of the two AI information-sharing factors are qualified by a significant interaction effect on perceived influence over the team and information certainty. Both interaction effects were sub-additive of their main effects, with their effect on influence being positive for all conditions except for SA1 (see Figure 4.4a). The sub-additive interaction effect on information certainty was negative, with all conditions being significant (see Figure 4.5a). The simple main effects of the interaction effect on influence found that AI interpretation caused a significant difference in participants' perceived influence for SA1, SA2, and the control conditions only (see Figure 4.4a). As for information certainty with the AI teammate, the simple main effects found that AI interpretation made a significant difference in the SA1 and control conditions only (see Figure 4.5a).

The model's results highlight the importance of affective states in designing and developing AI teammates to contribute through teaming behaviors such as information-sharing, especially as AI information-sharing has significant effects on the team cognition constructs, though affective states such as perceived influence and rating of the AI as a teammate heavily mediate their impact. The sub-additive interaction effects between the two manipulations on influence and information certainty were both sub-additive, which can be seen as an indicator of a possible point of diminishing return. This interaction effect also showcases that including interpretations of information is valuable but does not typically increase the perceived value of the information when added to other information, such as back-up behavior or augmenting team memory, though SA was a notable exception.

4.4 Study 2B: Methodology

4.4.1 Study 2B: Participants

For Study 2B, 21 participants were recruited and interviewed. The participants interviewed for Study 2B had an average age of 30.43 (SD = 7.51), with three participants identifying as women, two as non-binary, and the rest as men. Participants needed experience playing video games with AI teammates, which was communicated to participants in the recruitment information. These participants were recruited in two primary ways: 1) the survey from Study 2A; and 2) from video game-related Discord communities. These recruitment methods offered participants the option to volunteer to be contacted for an interview on human-AI teaming. Specifically, participants were asked if they would like to be considered to participate in a one-hour interview about their opinions and experiences with human-AI teaming in video games in exchange for a $10 gift card. Recruitment from Discord and Prolific was done to overcome the limitations of Prolific as specific communities (i.e., game genre, gender, ethnicity) could be given the opportunity to volunteer and participate more easily. All participants were compensated with a $10 Amazon gift card for their time. Participant information can be seen in Table 4.5.

Participant ID	**Age**	**Gender**	**Ethnicity**
P1	38	Male	White
P2	35	Male	Latino or Hispanic
P3	25	Non-Binary	White
P4	28	Female	Mixed Race
P5	38	Male	White
P6	24	Male	White
P7	40	Male	White
P8	43	Male	White
P9	28	Non-Binary	White
P10	33	Male	White, Latino or Hispanic
P11	24	Male	White
P12	25	Male	White
P13	24	Male	White
P14	25	Male	White
P15	24	Male	White
P16	32	Male	White
P17	19	Male	Asian
P18	40	Male	White
P19	42	Male	White
P20	20	Female	Hispanic
P21	32	Female	White

Table 4.5: Participant List and Demographic Information for Study 2B.

4.4.2 Study 2B: Qualitative Interview

The interview conducted for Study 2B began by having participants describe the context of their experience teaming up with AI teammates in video games and how they defined AI teammates in their experience. The researcher would then define human-AI teams and AI teammates in the current study context. Specifically, an AI teammate is any artificial entity with its role on a team and the ability to make decisions independently. The interview was conducted semi-structured following a script that focused on how AI teammates influence participants' attitudes, how these attitudes interact with team cognition, and how AI teammates may support team cognition in design from their own experiences. Each of the sessions targeted a 1-hour interview time to ensure saturation was achieved. These interviews were completed in English and recorded for transcription using Otter.AI.

4.4.3 Study 2B: Analysis

Participants' interview data was recorded using the built-in recording feature on Zoom (one participant chose not to have their interview recorded). The interview transcripts were analyzed using thematic analysis [103, 301, 22, 102], which saw each interview transcript reviewed for content related to the research questions of Study 2B. For RQ4.2 specifically, content regarding how participants developed their attitudes towards their AI teammates and what characteristics of those AI teammates would influence them, such as trust and cohesion. Interview transcripts were also reviewed for content relating to what participants wanted from their AI teammates to develop and support better team cognition, which related directly to RQ4.1.

The qualitative analysis consisted of four steps: 1) each of the interview transcripts was reviewed to reach a thorough understanding of what participants wanted from an AI teammate to develop and support positive attitudes and team cognition (i.e., situational awareness); 2) each of the interview transcripts were then reviewed again to identify what the major themes and sub-themes were in describing what participants wanted from an AI teammate to support team cognition; 3) the themes and sub-themes identified in the previous step were reviewed and a discussion with colleagues involved in the execution of the current study took place until a consensus on themes was reached; 4) specific quotes were selected from the transcripts by the author of the current dissertation, with the quotes being selected based on their ability to convey the meaning of each theme and sub-theme; and 5) the author of the current dissertation and colleagues involved in the execution of this study again reviewed the themes and sub-themes using the quotes selected in step 4 to continue to distill them into a representative synthesis of participants experiences and opinions on how AI teammates may support team cognition in human-AI teams.

4.5 Study 2B: Results

The following results section overviews the qualitative results collected and analyzed through Study 2B. This section provides additional context and insights into the quantitative results collected through the related work of Study 2A and continues to answer the D-RQs posed by the dissertation. Specifically, these results address D-RQ4 through the study-specific research questions of RQ4.1 and RQ4.2, which each focus on AI teammate design meant to help contribute to and support team cognition in human-AI teams. Study 2B provides additional context to Study 2A while also helping to address D-RQ4 by interviewing members of actual human-AI teams in competitive video gaming to ascertain what design features they want from an AI teammate to support team cognition. Quotes from the participants include added context from the question they were answering or from the subject being referenced using square brackets.

4.5.1 Predictability and Personability Dominate Attitudes Towards AI Teammates

The qualitative data revealed two major themes contributing to participants' attitudes toward their AI teammates, providing a response to RQ4.1. The first theme identified the surprising importance of personal interaction and identification with AI teammates, so much so that participants' prior experiences and ability to learn about their AI teammates played a major role in how they viewed that teammate outside of their performance. As for the second theme, it was found that the predictability of the AI teammate played another crucial role in how humans viewed their AI teammate. Even if the AI teammate could not communicate overtly, they claimed it would still be useful if they could at least reliably predict their actions, which is a relevant point that aligns with Study 1 findings and the fact that adequate natural language processing remains challenging.

4.5.1.1 Despite Being Machines, Humans' Attitudes Towards AI Teammates are Personal

Despite AI teammates not being able to express emotion in this study, several participants expressed a desire for their AI teammates to engage in communication of an interpersonal nature. They wanted them to express a general form of concern and sociability for their fellow teammates even though they are artificial, as P12 states:

> *"[I want an AI teammate to show] That you share the same passion for doing something specific. You know that you have a common goal, but*

there's only so far you can go with an AI that has a common goal with you. Even if you do something wrong, your teammate may be upset, but you feel the emotion from them being upset, and I think that's valuable because I feel like you learn from that experience more when you feel the emotion of how that affected them and affected your common goal." (P12)

While discussing the qualities of an AI teammate they feel would foster positive attitudes, P12 exclaimed how they would want one to show *"that you share the same passion for doing something."* In the case of enhancing shared cognition, the participant stated they would feel that it was a more genuine learning experience together to *"feel the emotion of how that affected them and your common goal."* P12 demonstrates a clear desire to have an AI teammate conveying a sense of passion toward the team's shared goal, which they feel would enhance their ability to learn from experiences and improve shared understanding over time.

This desire was not isolated, as several participants expressed a desire for AI teammates to convey a sense of personal connection to the team and their shared goal:

"Reassurance [from an AI teammate] would help with trust. Even something like positive reinforcement helps, you know, saying good job and that type of thing." (P12)

"I think having the AI be able to respond in a pleasant way would do a lot for that [enhancing cohesion] because the tone is everything. If AI comes across with a friendly tone, that's half the work." (P4)

The statement provided by P12 shows how reassurance and *"even something like positive reinforcement,"* would help with their trust in the AI teammate. While

P4 further reinforces the importance of focusing on affective outcomes *"because the tone is everything. If AI comes across with a friendly tone, that's half the work."* Together, these quotes reiterate research that AI teammates designed to produce better affective outcomes with their teammates result in increased levels of trust and cohesion. However, it is important to look back at the comment that it is only *"half the work,"* as once an AI teammate is able to foster positive attitudes from its human teammates, it can make meaningful contributions to the team and have those contributions more readily accepted.

The past experiences that human teammates bring with them into subsequent human-AI teams also play a significant role in their initial attitudes toward their AI teammates. So much so that initial interactions with the AI teammates are likely to be defined by the human teammates' personal beliefs and experiences with past AI, as the following examples from participants show:

> *"So in an RTS game, you have to kind of wait until the AI does something and then play off their play. Because it's not necessarily smart enough to know to do an advanced strategy or to even join your strategy to go in and coordinate with whatever you're going to do." (P1)*
>
> *"I prefer human teammates over AI teammates because I've had negative experiences with AI teammates in the past in video games, I haven't been able to trust them to do their job." (P9)*
>
> *"If I was looking into a game and it said 'hey we're trying out these new bots that we think are going to be doing this,' that would change my trust in them. Otherwise I'm going to be going into every situation under the predisposition that it's going to be adding a very small amount of value." (P11)*

Each of these participants brings up the point that *"I've had negative experiences with AI teammates,"* or *"it's not necessarily smart enough to know,"* and *"I'm going to be going into every situation under the predisposition that it's going to be adding a very small amount of value."* As P11 sums it up, many participants go into new experiences with AI teammates carrying baggage from their past experiences. These past experiences influence not only their attitudes towards the AI *"I haven't been able to trust them to do their job,"* but also their play style *"you have to kind of wait until the AI does something."* As new AI teammates are designed to play a more active role in teaming, such as making contributions to team cognition, as the current dissertation proposes, participants' expectations must be recalibrated. Recalibrating participants' expectations for their AI teammates could be accomplished by merely stating the AI has been altered and what to expect from them, which P11 stated *"would change my trust in them."* This focus on prior experiences and managing expectations is important to investigate as more and more AI teammates are designed to specifically engage in teamwork behaviors such as team cognition, and human teammates should be open to working with these new design features to improve team outcomes.

Participants also expressed a degree of concern and dissatisfaction with AI teammates that came across as disingenuous, illegitimate, and or tactless:

> *"Sometimes when it comes to window dressing it's like actors executing a script but that still feels very weird. And when you get outside of the bounds of the script, it's pretty easy to see that they're not that deep." (P5)*
>
> *"Sometimes they're shooting at nothing, or there's no substance to what they're saying, or what they're doing doesn't match with what they're say-*

ing, or vice versa." (P12)

Both statements by P5 and P12 signify that human teammates are annoyed and dissatisfied with AI teammates who do not adhere to very standard norms, creating a mismatch between their expectations for a teammate and the actual teaming behaviors they experience from their AI teammate. Specifically, P5 stated how *"it's pretty easy to see that they're not that deep,"* indicating their expectation for a teammate of more depth and their realization that it is unable to meet those expectations. There is also the more obvious issue of *"what they're doing doesn't match with what they're saying or vice versa."* These AI teammates are coming across as disingenuous and harming their human teammates' attitudes towards them, essentially showing them that they don't care enough to react appropriately. While AI teammates will always adhere to their programming, it is important to ensure AI teammates have adequate depth and consideration in action and communication to avoid the problem of AI appearing as window dressing.

Finally, other participants discussed how they would appreciate having the opportunity to learn about the task or about their AI teammates in a naturalistic way, demonstrating a desire to build rapport similar to how typical human-only teams learn more about one another over time (i.e., in training):

> *"If there were some way to go on this adventure and go with one specific teammate and say 'okay this is what I know about them.' And then along the way, you learn other things, and you get to see how they respond to things and get to know them like you would a person." (P5)*
> *"If I had it to where they [the AI teammate] guided me through their own skills personally, it would make them feel more individualized." (P10)*

The statements *"along the way, you learn other things, and you get to see how they respond to things,"* and *"if…[they] guided me through their own skills personally,"* by P5 and P10, respectively, demonstrate that while there is much to learn about an AI teammate from reading a manual, instructions, or similar articles, there remains a desire to learn the ins and outs of a teammate naturally by simply working with them. These participants feel the benefit would be akin to *"get[ting] to know them like you would a person,"* and that doing so would *"make them feel more individualized,"* thereby enhancing cohesion within the team and improving the shared knowledge team members have for the AI as they provide enhanced explainability naturally. Based on the several different facets of interpersonal communication and affect management that participants expressed were important to their attitude development towards their AI teammate, there is significance to the personal relationship that human teammates have with them, despite their artificiality. These findings strengthen the need for affective outcomes to be considered in the design of AI teammates to enhance their acceptance and ability to contribute towards team processes and subsequent team cognition. This assertion is especially pertinent given that Study 2A's SEM found the relationship between information-sharing and team cognition was heavily mediated by affective attitudes.

4.5.1.2 AI Predictability Satisfies An Aspiration for Coordination and Contribution With AI Teammates

The ability to predict what AI teammates are doing and what they may do in the future was also exceedingly important to participants' attitudes toward their AI teammates. The desire for predictable AI and their aversion to erratic behavior is best interpreted by the human teammates' desire to coordinate with those AI teammates. The perceptions around the lack of AI teammate predictability and the problems associated with it were exceedingly clear:

> *"The computer usually makes either the best decision or the worst decision, and there's no in-between." (P3)*
> *"It depends on the frequency; like a new AI teammate, I have no idea what I'm getting into. It's more like picking up a pick-up game with [random] humans. I don't know what their intent is, like how involved they are, how smart they are, or if they're just here to wreck things and laugh at me." (P1)*
> *"You have five settings and one of them was follow us, one of them was stay, and it was essentially a dice roll whether or not it was really going to take your command and that was even if you were practicing what I considered the 'best practice' to get that AI to listen." (P11)*

P3 mentioned that AI teammates *"usually make either the best decision or the worst decision,"* and this expectation makes it difficult for anyone to plan for either extreme properly. The statement by P1 also brings up how a new AI teammate makes them feel that they *"have no idea what [they're] getting into. It's more like picking up a pick-up game with [random] humans,"* which causes them to feel uncertainty regarding what *"their intent is."* This unpredictability is also something that may not be extremely easy to fix as P11 points out that even with AI teammates capable of accepting direction from human teammates, *"it was essentially a dice roll whether or not it was really going to take your command."* These participants describe AI teammates at a level of unpredictability that is nearly useless to human teammates, which serves to showcase just how important predictability can be.

There is also an association between predictability and trust, as P10 explained:

> *"Predictability, I think, is probably the better word. Trust is a little bit*

> *different in my mind, but after a while, you start to understand what it's programmed to do, and you know if it's always going to do something or not going to do you can understand the ratio of if it's going to do what you want it to or not and integrate that into your play." (P10)*

The association between predictability and trust with AI teammates, as P10 sees it, is that they are capable of *"understanding[ing] the ratio of it's going to do what you want it to or not and integrate that into your play."* P10's quote then highlights that through experience playing with an AI teammate, humans are able to improve their teamwork by predicting the AI's actions and accounting for it in their play. This is a form of trust based on reliance and familiarity; however, it is very weak, and AI teammates can and should strive to explicitly improve human teammates' trust in them using more concrete methods such as defined roles, transparency, and explainability.

Existing AI teammates have positive and negative aspects to their interactions with teammates, as P3 and P4 discuss:

> *"I think they generally just don't act the same way I would expect a human to respond because humans have that hesitancy and we have to react to stuff versus an AI can often do things instantly as they occur and it's a lot harder to keep up with that and understand their movements and directions." (P3)*
>
> *"[AI teammates can also be great in] That an AI is never going to discriminate against me for being a woman or being a person of color or anything like that." (P4)*

The inherent advantages of AI being artificial are also seen as a predictable facet of AI as P4 mentions that AI teammates can be great in teams because *"an AI is never going to discriminate against me for being a woman or being a person of color,"* showing that AI teammates inherent artificiality is appreciated and can be emphasized in design. Alternatively, some inherent advantages, such as instantaneous response times, are *"a lot harder to keep up with that and understand their movements and directions."* Both of these viewpoints showcase existing AI teammate design aspects that can be improved upon through design and identify AI teammates' explainability as a possible solution for improving AI teammate predictability.

AI teammates that provide either explainability after making an action or decision were perceived as being more predictable and trustworthy, as the following examples indicate:

> *"I liked it's reasoning because just having a reasoning helps, it gives off a veneer of concern that we know is not really real with an AI but it is a nice feature to have the more human the AI gets, so to speak." (P4)*
> *"For me knowing what and why an AI teammate is doing something is really influential to my trust in them. Even if they make a mistake if they could provide a good reason why my trust in them really wouldn't be that affected." (P9)*
> *("That'd be awesome [getting explainability from the AI] because you're getting inside the mind of the AI, and you can see how it thinks. You can even become predictive of what the AI is going to perform or future suggestions that it's going to make." P3)*

Being able to *"[get] inside the mind of the AI,"* through explainability improves human teammates' ability to *"become predictive of what the AI is going to perform or future suggestions that it's going to make,"* enhancing team cognition through team situational awareness. Explainability by an AI teammate also improves affective outcomes and processes by indicating that the AI was concerned with the team's shared goal. As P4 says, *"just having a reasoning helps, it gives off a veneer of concern"*. Explainability is also a form of trust repair, and P9 recognizes this as they claim *"even if they make a mistake if they could provide a good reason why, my trust in them really wouldn't be that affected,"* showing that explainability trust repair can apply to human-AI teams. These participants detail how AI teammates providing explainability give human teammates insight into their AI teammates' decision-making process regarding how they consider their environment, task, and fellow teammates increasing their predictability for the future.

Enhancing the predictability of AI teammates either through transparency or explainability is clearly critical to enhancing attitudes towards AI teammates, but these results go a step further to show that AI teammates providing reasoning behind their actions also enhances their personality. Explainability provided through conversation reminds human teammates that their AI teammate shares a common goal with them, which the AI is also concerned with accomplishing. This assertion aligns with the quantitative results provided by Study 2A, which showed the significant positive influence of the explainability attribute on participants' attitudes.

4.5.2 Humans Want Impactful AI Teammates Contributing to Team Cognition in A Variety of Ways

The second major theme addresses RQ4.2, which questions how humans want AI teammates to contribute to shared knowledge or team processes. There were three sub-themes that reflected this topic. Contributions to shared knowledge include the shared understanding of roles, interdependencies, strategy, and communication techniques. Participants expressed a strong desire for AI teammates to be autonomous within their own roles and for AI teammates to contribute to the team's shared awareness. Regarding team processes, human teammates saw context as the single biggest driver of how their AI teammates should communicate with them. Finally, human teammates noted the potential for AI teammates to act in the capacity of an exemplar of effective system usage, task execution, and even team cognition behaviors.

4.5.2.1 Human Teammates Want AI to Be Interdependent Teammates with Agency

Many participants viewed current AI teammates as a subordinate tool that acts as an extension of the user and expressed their dissatisfaction with these AI teammates as P4 and P5 explain:

> *"For the most part [when working with AI teammates], I just try to learn what their abilities are and work around that as I lead the team or what have you. I just kind of let them do their thing unless I need to step in." (P4)*
>
> *"It's [bad experiences with AI teammates are] things like, where I feel that I'm carrying all the way. Like where the AI is basically just an extension of my players abilities, and they're not really doing anything independent, or offering any real help." (P5)*

Working with AI teammates typically forces the human teammates to *"work around"* their AI teammates and their capabilities, causing the human teammates to leave the AI to its own devices by *"let[ting] them do their thing unless I need to step in,"* as P4 details. AI teammates that engender this type of interaction with their teammates are typically not seen as good teammates and cause human teammates to feel as if they are personally *"carrying [the team] all the way."* This mindset will often lead to resentment of the AI teammate causing those perceptions, as P5 describes it *"Where the AI is basically just an extension of my players abilities, and they're not really doing anything independent."* These participants discuss an important aspect of where AI teammates oftentimes fall short, which in this case, is their failure to meet the most basic expectations of a teammate by not taking on a defined role within the team.

Having an AI teammate who fails to meet the expectations of an actual teammate can result in more than just negative affective outcomes like dissatisfaction. AI teammates that lack any agency or defined role expectations can result in worsened team cognition, as P10 and P12 discuss:

> *"No, [I don't think my AI teammates think the same way I do when it comes to interacting and communicating] I can't really speak to them like I would a human and there are only so many commands you can give them. And, sometimes I want the personal opinion from my teammate to hear what they think in a given situation, and AI will usually blindly do whatever you say." (P12)*
> *"Not really, [I don't feel like my AI teammates think the same way when it comes to completing a shared goal] they would only engage the enemy really if I was also attacking the enemy. Yes, [they emulate the player] whenever it comes to attacking, they would emulate me but I don't think they would do specifics in terms of getting an item or carrying something." (P10)*

When discussing whether or not they feel that the AI teammates in their experience have shared knowledge when trying to complete a shared goal, P10 and

P12 convey that they lack shared knowledge because they cannot give a *"personal opinion...to hear what they think in a given situation"* and *"they would [not] do specifics in terms of getting an item or carrying something."* These two participants highlight how AI teammates, without any ability to engage in teamwork outside of rote taskwork, fail to contribute to the team in a meaningful way and hurt team cognition in the process, reinforcing the belief that the AI teammate does not have any shared knowledge in communication and interaction.

However, there are examples of occasions when AI teammates were given agency and defined role expectations and the improvements it had on teams for P9 and P10:

> *"[A great experience teaming up with an AI was] Having an AI teammate that worked in a specific role and was good [in that role] was really beneficial to the team." (P9)*
>
> *"I thought the command and conquer games were awesome because it let you set your AI teammates under classes and gear them up." (P10)*

Getting an AI teammate that was capable of *"work[ing] in a specific role and was good [in that role] was really beneficial to the team,"* and this sentiment by P9 can be applied to a wide variety of AI teammates where merely setting expectations for them as a teammate and giving them the ability to meet those expectations alone can improve cognition and outcomes. This improvement can be partially attributed to AI predictability and shared knowledge improvement resulting from the AI teammate being given a distinct role and agency. However, this improvement also comes from human teammates being able to understand what to expect from them throughout the teaming process, given the AI's role and expectations are well-defined. P10 reiterates this by recalling a positive experience for them was being able to *"set [their] AI* *teammates under classes and gear them up,"* and not only further emphasizes the importance of defined roles for AI teammates but also shows that these can even be user-defined roles.

Addressing what additional considerations apply once an AI teammate has taken on a defined role and become an interdependent teammate with significant agency, participants described what they should pay attention to when making decisions:

> *"The AI should prefer their own teammates before the objectives regarding the enemy. You are constantly aware of your teammates and your first process in your brain, or at least in mine, is to think about the person first and then you choose what you're going to do. If that guy is moving up on the flag, maybe I shouldn't move up as well, maybe I should flank." (P16)*
> *"If I'm planning a raid there are bosses that do certain things, and there are a lot of prepping and opinions that go into planning. So I'm gonna have this gun to counter this certain move or other similar specific knowledge that would be really huge for an AI to provide while prepping." (P12)*

Human teammates expect their AI teammate to *"prefer their own teammates before the objectives"*, and that AI teammates should make these considerations and *"then choose what you're going to do,"*, as P16 conveys. This statement by P16 shows a preference for adequate team situational awareness by the AI teammate, such that they are capable of keeping up with teammates' intentions and complement them, for example, *"If that [person] is moving up on the flag,"* they *"should [probably] flank,"* to complement their teammate's actions. P12 also conveys how much they would appreciate an AI teammate to provide additional insights while they are *"planning* *a raid,"* because there are *"a lot of prepping and opinions that go into planning,"* so if the AI teammate could contribute *"specific knowledge,"* it would benefit the team's ability to make accurate decisions. P12's statement is highly reminiscent of the empirical findings for Study 2A, which saw the augmenting team memory AI encourage positive attitudes and team cognition from participants. Together these quotes demonstrate how AI teammates can make additional considerations while meeting the expectations of their role in a way that contributes even further to team cognition and team outcomes.

Placing AI teammates into interdependent roles with a significant degree of agency is a hallmark of human-AI teaming by definition. It is clear that members of the early versions of these teams are beginning to expect these abilities from their AI teammates. While there are still considerations to make when placing AI into these positions, there are also several positive advantages, such as clearly defined expectations and increased shared knowledge throughout the team. However, defining roles for AI teammates also allows them to make more significant contributions to the team by leveraging their inherent technical advantages as computationally based entities, such as augmenting team memory or updating team situational awareness.

4.5.2.2 Shared Situational Awareness is a Significant Area for AI To Contribute to Shared Knowledge

As the SEM outlined by Study 2A displayed, the explainability of AI and situational awareness of intra/extra team information changes AI both greatly affected perceived attitudes and team cognition. The following sub-theme provides additional context to these empirical findings by detailing why these information-sharing attributes were so important to human-AI teams. Specifically, the aspects of AI teammates related to team situational awareness were critical but included several nuanced aspects that help inform specifically *how* AI teammates should contribute to individual and team situational awareness.

The stated need from participants regarding various aspects of situational awareness was widespread and included several different examples of how AI could contribute to this aspect of shared knowledge in meaningful ways:

> *"It's very important in Halo whenever you kill two people and you know there are only two people alive on the enemy team. I'm always trying to say 'hey, we're up four to two right now so we should have an advantage everywhere on the map if we see somebody.' So that sort of information is worth its weight in gold." (P11)*
>
> *"It [AI assisting in monitoring processes] would help me understand the situation because if I keep trying to keep track of everything, it can be a little overwhelming just trying to do that, especially as it changes. If we had an AI teammate talk about it or update me about it at the same time it can make it to where I can focus on more important things." (P10)*
>
> *"It [the AI teammate] can say that person X is going to the sniper position, and you may have some cover in approximately one minute. Things like that would be extremely helpful for building that bridge between you and the AI and you and the other players on the team that you may not have played with before." (P1)*

P11 conveys the first example of how an AI teammate can contribute to shared situational awareness by being capable of understanding what information in a given context is *"worth its weight in gold,"* and being capable of communicating that information to its teammates when it is timely and relevant. However, P10 goes further to highlight an example of how if they are *"trying to keep track of everything, it can* *be a little overwhelming,"* and if they had an AI teammate capable of handling that monitoring, it would allow them to *"focus on more important things."* These statements by P10 are a great example of how an AI teammate can utilize its technical advantages to improve teammate performance and capability. However, as P1 brings up, being a part of team cognition, team situational awareness and the ability to project future states comes with time and familiarity with teammates and is difficult to utilize properly with *"other players on the team that you may not have played with before."* As such, it would be beneficial if AI teammates were designed to help with the process of *"building that bridge between you and the AI and you and the other players,"* by communicating relevant changes to teammate activities that may influence decisions made in teamwork and taskwork. These examples show how varied AI teammates' contributions to team situational awareness can be moving forward. From communicating relevant taskwork information to freeing up cognitive resources for teammates in monitoring processes, team situational awareness contributions are at the forefront of human-AI team members' minds.

Designing AI teammates to make contributions to team situational awareness comes with considerable nuance, however, as factors such as terminology, conveying awareness, and even the expectation for more advanced pattern recognition. In exploring these nuances in conversation, participants outlined a few key elements of team situational awareness that are essential to designing AI teammates that make similar contributions:

> *"It's [providing callouts of enemy locations] gonna be people sharing just one or two words at a time, and I mean 'up top stairs, shotgun alley,' stuff like that. When you're playing these games, they [fellow human teammates] know where stuff is going to be, so if you say the two word phrase*

> *for top tower, they know where that's at on any part of the map." (P11)*
>
> *"[A good AI teammate experience] Racing game AI is pretty good. They drive in a way that makes you feel like they're aware of the player. Just feeling that you are present in the world to the AI. That is what makes or breaks everything teamwork for me." (P16)*
>
> *"So the AI could make callouts like which way the ball is going or even what kind of play the other team is setting up. So it would be amazing if the AI could call out opposing team habits, positions, or other certain things that they might do, like, predicting a certain play." (P6)*

First, P11 discusses how important *"callouts"* are to successful teamwork but emphasizes the need for a shared understanding of the environment. Specifically, P11 mentions that *"when you're playing these games, they know where stuff is going to be...they know where that's at on any part of the map."* This assertion means that any AI teammate contributing to team situational awareness in the form of callouts will need to conform to a shared understanding of the environment and the shared terminology used by their human teammates. Additionally, P16 identifies the AI simply being capable of *"making you feel like they're aware of the player,"* essential to their idea of teamwork. In practice, this awareness may be accomplished through several means but will likely be dominated by contextual requirements and constraints. Lastly, P6 takes the concept of *"callouts"* a step further by suggesting their AI teammate *"could call out opposing team habits, positions, or other certain things that they might do, like, predicting a certain play."* While these expectations are significantly greater than just conveying the positions of the opposing team members, they show an expectation for AI teammates to engage in level-three situational awareness. Participants want their AI teammates to contribute predictions of future states to the rest of the team, further highlighting just how important situational awareness is to human-AI teams. Together, these participants depict an adherence to some of the key facets of the literature regarding team situational awareness and shared mental models in the form of shared terminology, shared understanding of the environment, awareness, and engagement in higher levels of situational awareness.

Going deeper into how AI teammates should be conveying situational awareness updates, several participants conveyed how important it was to them that AI teammates utilize their technical advantages to provide highly accurate and actionable information updates:

> *"What they see, like callouts of their own, which AI sometimes does right now but not always in the way that you want. For example, they'll have an automated saying, 'there's a guy in the kitchen,' but there's more information that can go with that." (P12)*
> *"Maybe a teammate has low health and needs a health pack and the AI could help map and pinpoint the locations of health packs near them and give statuses on if they're near a teammate." (P10)*
> *"Understanding the AI's reasoning process is important because when it's giving information on lap times, your relative position, and the other drivers on the track, that's all empirical information that's easily accessible. So looking back and seeing if one strategy was better than the other and having the AI show you with data why their advice was correct is something I would want to see." (P8)*

Here, P12 echos other comments that *"callouts"* are useful alone; however, they go on to state that *"there's more information that can go with that,"* which means designing AI teammates for team situational awareness should also consider what information is important to convey and to do so with precision. For example, an AI teammate could state that an enemy team member is in a specific room, in a specific corner, with a specific weapon, thus providing their teammates with as much actionable information as possible to enhance the team's ability to engage in level two and three situational awareness. Actionable contributions to team situational awareness can also be simple by contributing to level one team situational awareness by *"help[ing] map and pinpoint the locations of health packs near them,"* as P10 points out. This sentiment for highly specific information is echoed by P8 when they bring up their desire for *"having the AI show you with data why their advice was correct,"*, but P8 goes further by pointing out that providing specific data can be important to AI explainability because *"that's all empirical information that's easily accessible."* All together, these quotes are important because they demonstrate how contributions to team situational awareness can leverage AI teammates' ability to be incredibly specific and data-driven by nature, which is something human teammates recognize and want to be integrated into their idea of team situational awareness.

The question of when and how to provide actionable and precise information to human teammates also included specific constraints voiced by participants, which largely centered around the context the team found themselves in at the time, as several participants provided examples:

> *"When and where to say things [is incredibly important for AI contributions to shared understanding], there are certain situations where somebody could say something that isn't relevant at that point in time, or even if you're trying to listen to enemy footsteps and you have somebody talking, and you can't hear what's going on around you and your environment." (P12)*

> *"Whatever explanations the AI gives would need to be within the bounds of the context the team is currently in, so if we're really busy I wouldn't want them trying to explain why they're doing something and I don't have the bandwidth to see it." (P9)*
>
> *"So just like with a person, if they see something that's happening, you both then can prepare before it actually happens. With an AI, a lot of times, I need to change my strategy and adapt to what the AI is doing, and oftentimes I can't tell what it's doing until the event is actually happening." (P1)*

Participants such as P12 and P9 referenced how an AI teammate's communication frequency or *"When and where to say things,"* as P12 put it, would need to consider the context of the environment. For example, P12 shared that *"there are certain situations where somebody could say something that isn't relevant...or even if you're trying to listen to enemy footsteps and you have somebody talking,"* where communication provided by an AI teammate at the wrong moment would be extremely detrimental to their teammates' ability to accomplish their goals. P9 went on to specifically state that *"whatever explanations the AI gives would need to be within the bounds of the context the team is currently in."*. Both of these examples from participants also go back to relate to team norms, which AI teammates must be capable of adhering to as far as contributions to team cognition go. Adhering to norms is vital, as a violation of them will negatively affect teammates' attitudes towards the AI, making it harder for the AI teammates' contributions to be accepted. such as whether it was competitive or whether listening to nearby environmental sounds was essential to the team goal. Lastly, P1 references how AI could vastly enhance their ability to coordinate with teammates and improve team outcomes by simply knowing to communicate their intent *"alongside"* the situational awareness information from the environment that caused the AI to make its decision, because *"oftentimes I can't tell what it's doing until the event is actually happening."* The importance of context is expressly conveyed here as these participants relate to one another regarding how, why, and when AI teammates should make contributions to team cognition based on the context of the team's environment and situation.

Team situational awareness is a critical facet of team cognition, and through this series of qualitative findings, it is evident that human teammates recognize this and also see it as an ideal aspect for AI teammates to contribute to the team. From freeing up fellow teammates' cognitive resources to providing immediate actionable information, AI teammates can utilize their inherent technical advantages to contribute to team cognition. These contributions must, however, be bound and informed by the context of the team's environment and utilize aspects of shared knowledge like shared terminology.

Additionally, these qualitative results coincide with the quantitative results shown in Study 2A to provide additional insight into how and why participants may perceive less control over their team in the explainability and situational awareness of intra/extra team information changes results. Specifically, the qualitative data reiterated how much more participants valued information regarding intra/extra team information changes as opposed to situational awareness of team members. With enhanced explainability of AI teammates, participants appear to be more open to losing influence over the team if it goes with the increased perceived utility of a hypothetical AI teammate. As such, adequate explainability and utility through features such as augmenting team memory or providing situational awareness updates of intra/extra team information changes afford the increased role of AI teammates and contributions to team cognition and the subsequent loss of influence over the team necessary for the AI teammate to take on a defined role.

4.5.2.3 Designing AI to Act to Encourage Team Cognition Enhancing Behaviors

Building upon the previous themes about explainability and situational awareness, the participants also demonstrated that human teammates are open to AI acting as exemplars of effective taskwork, teamwork, and team cognition behaviors. This finding is also an extension of findings from Study 1 that hinted towards the openness of an AI exemplar and the potential utility of AI exemplar behavior for developing and supporting team cognition.

Specifically, participants expressly communicated their openness to AI acting as exemplars in a variety of scenarios, though there were a couple of caveats that should be kept in mind. For instance, P1 shared that:

> *"Sometimes they do things that aren't executable by humans, and they'll hit a shot that I could never hit, so I'm happy to have you on my team AI, but I can't learn from you; humans just can't do that. So I appreciate it when it is more human-like so that I can learn something from it." (P1)*

While the participant feels that typical AI plays an irrefutably beneficial role in achieving important team goals, the strategies they use often *"aren't executable by humans,"* making it impossible for humans to *"learn from"* their AI teammate. Increasing the *"human-like"* qualities of an AI teammate can, thus, make the teammate a benefit to the team goals and learning process, expanding the benefits for all when appropriately designed. Expounding upon this perspective on benefiting from examples given by an AI teammate, P9 explained that *"I'd be willing to have an AI teammate help me learn how to play a game and its strategy, but I would need to know* *the reasoning behind their actions before I'd be okay with accepting it, especially in new games"*. Once again, this participant demonstrates an openness to learning from an AI teammate, but humans require a certain degree of understanding of the AI's *"reasoning"* behind an action or strategy before they can internalize this approach and incorporate it into their own actions. Thus, both participants demonstrate that human teammates need a certain degree of knowledge behind the AI's approach before they can openly accept these AI exemplars as true teammates.

Regarding what this would look like from AI teammates in practice, participants relayed several potential examples:

> *"Some people like to run off on their own so if the AI kind of reinforces team-play I think that would have some kind of impact on how each of the players approaches the next round." (P6)*
>
> *"I had no idea I could even do that, and they just absolutely devastated the opponent. So now I'm going to see what they did [the AI] and try to do that again myself. So I've had experiences where I'm newer to a game, and they've totally dominated, and I can learn how to play better because they actually did that in front of me, and oftentimes it was the reason why I wanted the game." (P1)*

P6 demonstrates how AI teammates acting as an exemplar could actually demonstrate and encourage positive teaming behaviors among their fellow teammates, improving the quality of the team overall. P1 provides another example of this, pointing out that they see an AI teammate engaging in a strategy or move they were unaware was possible for them, thus changing their gameplay strategy. AI teammates acting as exemplars could even personalize the experience as P12 states *"but they [the*

AI teammate] apply it in a way that feels like it's catered to you,", providing guidance in a way that is best suited to specific play-styles. All of these examples together demonstrate the desire for an AI teammate who is not only capable of lifting the team up from a taskwork and performance perspective but also to have an AI teammate capable of improving their teamwork behaviors.

Placing AI teammates in the role of exemplar for human teammates could not only help establish common knowledge, supporting similar mental models among teammates, but also these AI would bring up several other advantages for improving team processes:

> *"If you're not communicating then it's hard to play so if an AI teammate could somehow facilitate communication to open up that connection a little bit better or make it a little bit more comfortable for people where the AI teammate could act as like ice-breaking, but then also could enforce rules about things like griefing." (P5)*
>
> *"If I had it to where they (the AI teammate) guided me through their own skills personally, it would make them feel more individualized." (P10)*

P5 conveys that AI teammates in the exemplar role can help facilitate effective communication among teammates as a hallmark of effective teaming. They also discuss how AI could help ensure other teammates are not acting in bad faith or *"griefing"*. Based on this participant's recommendation, these AI would always be available to help monitor and help human teammates improve. P10 then describes how if an AI teammate personally guided them *"through their own skills"*, they would feel a stronger relationship with them because it feels more individualized and personal. Continuing to draw upon these perspectives and apply them to other components of developing teams, P5 explained *"schedules and things can be very challenging. So being able to practice [with an AI teammate] and be able to play that game when everybody else is in different time zones and can't make it due to personal stuff [is valuable]."* Again, this participant is expressing how AI teammates acting in the capacity of an exemplar is a significant advantage for them as they can still improve their skills in the task by teaming up with an AI teammate despite the constraints that exist in traditional human-only teams. These quotes highlight just how much more of a role AI teammates can play in the role of exemplar by improving processes through advantages to training, improving communication with unfamiliar teammates, and enhancing the connection between teammates.

Several participants also made consistent references for AI teammates to utilize their technical advantages to improve their understanding of their fellow teammates and of the AI itself:

> *"Like having the AI say 'hey this teammate has taken the sniper position on 45 of their last games, so you may have some cover coming in the future from them.' If that teammate doesn't like to talk they wouldn't have told me that but the AI just did because they've played with him before." (P1)*
>
> *"So instead of pointing out a certain teammate's flaws or weaknesses, just highlighting their strengths and maybe like a certain play they run has a high chance of success so that players can play around each other's strengths. Just know what role each player can have, and that would more likely lead to success. That would build more confidence in each other." (P6)*
>
> *"I think if the AI gave us useful ideas for strategizing and planning, by*

throwing out stuff to help us work together better...I think that if the AI were able to pick up on our strengths and weaknesses and communicate that in a way that we could plan around it, that would be really helpful." (P4)

"They're [autonomous cars AI for computer vision] not using other bands that are available, for instance, if an AI had additional capabilities, like it could see infrared or see thermal, that would be very helpful to have the AI provide capabilities that humans couldn't do themselves...It [the AI teammate] can provide additional information." (P19)

P1 emphasizes how human teammates strongly prefer an AI teammate that is capable of recognizing the actions and tendencies of its teammates that are consequential to team goals and conveying that information to the team. This action is especially important in certain situations such as *"If that teammate doesn't like to talk,"* or other similar situations where the AI teammate can step into a larger contributing role to communication and coordination. Other participants, such as P6 and P4, assert that AI teammates should be *"highlighting their strengths...so that players can play around each other's strengths"* enabling the individual team members to *"work together better"* because of the AI teammates ability to *"to pick up on our strengths and weaknesses and communicate that in a way that we could plan around it."* This statement shows that AI teammates can have qualities that go beyond improving team cognition by themselves, but encouraging behavior that encourages team cognition *and* conveying the information human teammates need to engage in those behaviors most effectively. AI teammates, by their very nature, have the capability to be *"very helpful to have the AI provide capabilities that humans couldn't do themselves,"* and encouraging team cognition behaviors is an important area of teaming for AI to enhance with those abilities. AI teammates should be capable of providing skills and resources to human-AI teams that human-only teams do not have access to, and these resources should be taken advantage of to improve teamwork and not just taskwork. These statements made by participants show how *"having that awareness [of teammates and their abilities] provided by the AI"* would be extremely beneficial to these teams and would allow them to fully leverage AI teammates in a way that can move them past human-only teams in certain respects. RQ4.2 sought to understand how AI teammates could be empowered to contribute to team cognition meaningfully; however, participants have overwhelmingly indicated that AI should not only be contributing to it but should be assisting them in developing their own effective team cognition behaviors with fellow human teammates at the same time.

AI teammates can act as an exemplar for several aspects of tasks involving human-AI teaming, whether that be system operation, proficient task strategies, or effective team cognition behaviors. Human teammates working on human-AI teams are open to AI operating in this capacity as participants recognized the potential benefits such an AI teammate could have for themselves and their fellow teammates. AI exemplar behavior also goes hand in hand with the aspects of explainability, team memory augmentation, and contributions to team situational awareness that were important to both the current qualitative results and the quantitative findings of Study 2A.

4.5.3 Summary of Results

The qualitative findings of Study 2B complement the results of Study 2A by providing additional context to the mediated influence the various information-sharing attributes had on attitudes and perceived team cognition within the SEM.

First, the importance of AI predictability comes across as paramount to human teammates' attitudes towards their AI teammates, which went along with the surprising personal nature of how attitudes towards AI teammates develop, meaning any kind of communicated predictability should be genuine and natural. Second, human teammates increasingly want their AI teammates to be interdependent teammates with a significant degree of agency and the former predictability and genuine consideration of their human teammates is paramount to AI stepping into a defined role on these human-AI teams. Once in that role, human teammates perceived a significant loss of perceived influence over the team's actions and outcomes. This loss of influence coincided with an increase in the utility of the AI teammate, specifically, with an AI teammate providing contributions to team situational awareness. In providing these contributions to team cognition, human teammates expressed how AI teammates could help them individually and their teams more directly by acting as exemplars. Breaking the ice and kick-starting communication, as conveyed in the current study (2B) and in Study 1, these AI could showcase effective strategy, system usage, or, as the current study found, effective team cognition behaviors that develop and support the construct for all team members to emulate. The results of Studies 2A and 2B show that AI teammates have several opportunities to contribute to team cognition. Human teammates also expressed a strong desire for AI teammates to have these features and are even, at some points, becoming frustrated by their inability to implement these team-enhancing features.

4.6 Bridging the Gap: Connections Between Study 2A and Study 2B

Studies 2A and 2B built upon Study 1 and the existing body of research on team cognition development and support by explicitly examining how those with experience teaming with AI perceive various characteristics of AI teammates and what they desire from AI teammates in multiple contexts. Study 2A focused on how AI teammates with various information-sharing attributes and interpretation of that information affected perceived team cognition and attitudes using SEM. This study addressed RQ2.2, which questions how the AI information-sharing attributes influenced participants' attitudes, and found that perceived attitudes towards AI and human teammates were indeed significantly influenced. Specifically, the explainability of AI and situational awareness of intra/extra team information changes resulted in the largest effect on attitudes, though all AI information-sharing attributes had an effect on at least one attitude measured. Additionally, the back-up behavior and explainability AI attributes surprisingly affected participants' perceived shared mental model with their fellow human teammate. Study 2A also addressed RQ3.1 and RQ3.2, which asked how the same manipulations influenced perceptions of team cognition. The SEM conducted in Study 2A found that information-sharing by AI teammates does have a significant impact on specific aspects of team cognition, such as perceived contributions of the AI to situational awareness, but that affective attitudes primarily mediate their effect.

Lastly, Study 2B addressed the research questions RQ4.1 and RQ4.2, which asked what aspects of AI teammates influence how human teammates make attitudinal judgments of an AI teammate and how they envision AI teammates making meaningful contributions to team cognition through shared knowledge or team processes. This research found that human teammates place significant importance on the personal nature of their relationship with an AI teammate, in that they want it to show that it cares about their shared goal and that they also want predictable AI to avoid frustrations with coordination and strategy. Additionally, human teammates strongly want AI teammates capable of stepping into an interdependent role on their team with a significant degree of agency. AI teammates can contribute to team situational awareness by stepping into these roles. They can even enhance the team by exemplifying effective task execution, system usage, and teammate tendencies, with the last example being a major factor in how humans see AI teammates enhancing team coordination. The following discussion critically examines these findings in light of Study 1 and the existing research by interpreting how they expand upon existing research.

4.6.1 Improving AI Teammates Contributions to Team Cognition Lies in the Information They Serve

A central theme of the results from Study 2A was the effect that essentially all AI information-sharing attributes had a significant impact on attitudes and, mediated through them, an impact on measures of perceived team cognition. Specifically, the results of this study indicate that AI teammates will most often benefit teams when they share intra/extra team information and provide explanations alongside their behaviors (Figure 4.3). Critically, while these types of information were examined individually, they provide a holistic benefit when combined. In particular, the intra-extra information gathered from the environment would consist of the information AI teammates use to make decisions, which means this information would also be present within an explanation. AI teammates should generally provide this information throughout the task while explicitly linking it to its behaviors and actions, which will consistently inform teammates and drive the formation of affective and cognitive states. Moreover, in addition to benefiting these shared perceptual states, this combined information-sharing also can benefit the individual SA humans form [36]. AI teammates and roles that leverage this type of information-sharing will act as a catalyst for affective and cognitive state development, improving individual and team performance levels, in turn, [200, 195]. Second, it's crucial to understand how AI teammates should share this information. Based on the results of this study's second manipulation and the above recommendation, interpretation should be used sparingly in human-AI teams. This is not to say that AI teammates should never make recommendations, as recent research has shown the likelihood for humans to follow AI directives when they have greater levels of trust [39, 25]. Instead, AI teammates should predominately act as an information conduit through which information can be repeatedly gathered and disseminated, which human teammates can interpret. In turn, the effects of including AI interpretation did not significantly affect perceived influence or information certainty when interacting with the six information-sharing attributes; however, the SA information-sharing attributes were an exception to this regarding the sub-additive interaction effects. Moreover, while this study explored the perceptions of singular human teammates, this methodology would also allow for information to be more generally applicable to multiple human teammates simultaneously, as interpretations may need to be tailored to specific human teammates.

The finding that any information-sharing attribute was better than nothing speaks to the importance of Study 2B's findings that human teammates want AI teammates in interdependent roles alongside them. Recent theoretical research has emphasized the importance of AI teammates having very little agency and human teammates maintaining significant control over them [291, 292, 293]. However, the current study showcases that those with experience on human-AI teams have a strong desire for independent AI teammates in specific team roles, as it was a consistent theme throughout the qualitative results of Study 2B. These findings go on to describe several instances where humans wanted AI teammates to act with high levels of agency outside of their direct control, just as prior theoretical work has predicted [239]. Participants described multiple scenarios involving environmental constraints where they would be unable to manage an AI teammate's actions and would prefer if it could contribute independently without their intervention or input. This finding opens the door to the idea that providing defined roles for AI teammates that are explicit and convey that they share a goal with their human teammates benefits team cognition. Once a teammate has an assigned role, it is easier to predict and understand their decisions and possibly project their actions in the future or in response to unexpected events. This idea has been examined in existing literature in the form of adaptive AI teammates in defined roles working within specific work cycles, and they have demonstrated positive effects [135]. Such outcomes are easily implemented and could also go so far as to convey what aspects of team cognition the AI is capable of contributing to, such as team situational awareness or transactive memory systems. Without making contributions to team cognition from defined roles, AI teammates will continue to be perceived as teammates who underperform and lack the ability to convey their capabilities.

4.6.2 AI Teammates Have Exceptional Importance to Human Teammates' Understanding of Situational Awareness in Human-AI Teams

The results of Study 2 demonstrate that AI teammates represent a massive component of what team cognition is in human-AI teams, with individual and team situational awareness leading the way. As such, AI teammates may contribute to team cognition by providing adequate explainability to their human teammates. At its most basic level, this would be done by ensuring human teammates are capable of understanding why the AI took a certain action, whether that be a priori by conveying what its role will be or post-hoc by providing explanations for those actions. This work provides empirical backing to the theoretical research recently posited by Endsley, which stated the need for information to be contributed to taskwork and teamwork situational awareness *and* AI situational awareness [85]. When it comes to human-AI teams, the AI teammate represents a major component of human team members' individual and team situational awareness to the point where the explainability AI in Study 2A significantly enhanced participants' attitudes and perceived team cognition. The importance of explainability to attitudes concurs with previous human-AI interaction literature [290], but applies it within the realm of human-AI teaming and explores new ones such as perceived teammate rating and utility. The qualitative data in Study 2B went on to emphasize AI predictability, which was likely a matter of participants' desire to accurately engage in level three situational awareness or projection to future events [82]. This desire to engage in level three situational awareness is also likely why participants in Study 2A preferred the situational awareness of intra/extra team information changes AI over the situational awareness of team members' AI. Having the ability to be informed of intra/extra team information changes has a significantly greater benefit to predicting the future state of the situation than being informed of teammate statuses, despite both AI contributing to taskwork situational awareness [86, 85].

However, the desire for AI teammates that influenced the team's strategy, immediate goals, and other situational awareness related aspects was notable in the current study's findings [83, 101]. Contributions to teamwork situational awareness can be difficult for AI given the propensity for AI to misinterpret human teammates' intentions or priorities [26, 220]; however, this is not to say that it cannot be done, but that care must be taken in its research and development. For example, the results of the current study can be interpreted to highlight how AI can avoid misinterpreting aspects of teamwork situational awareness by integrating AI explainability alongside situational awareness. Specifically, as actions made by an AI teammate can often be attributed to changes in the environment, the AI teammate could express clearly and concisely what aspect of the environment caused their change in behavior. This simple addition would enable AI teammates to make subtle contributions to strategy as human teammates would be made aware of what changed and the AI's intent, allowing them to plan accordingly or alter the AI's actions or plan. AI explainability alone also has several advantages, such as increasing trust and performance [233]; however, explainable AI also has the potential to be difficult in practice given the brittle nature of AI model training and humans tendency to develop inaccurate mental models of the AI that are highly resistant to change [74]. Care must also be taken to ensure that explainability information is not seen as extraneous detail, which can happen with more experienced human teammates [240]. Clearly, the relationship between situational awareness and AI explainability is critical to develop further as the current study continues to clarify the nuance of their relationship.

4.7 Design Recommendations Synthesized From Study 2A and Study 2B

Following the stated goals of the current dissertation, the results of Studies 2A and 2B have been reviewed and merged into a series of design recommendations. These recommendations are not only taken from and supported by the results of Study 2A and 2B but have also been informed by those of Study 1 as they informed the design of Study 2A and 2B.

4.7.1 AI Contributions to Team Situational Awareness Can and Should Go Alongside AI Explainability

AI explainability and situational awareness of intra/extra team information changes were incredibly important to Study 2A and 2B participants. However, the qualitative themes revealed how intrinsically linked these concepts are for AI teammates. Specifically, AI teammates can and should take advantage of the fact that by making meaningful contributions to team situational awareness through updates of intra/extra information changes, they can also enhance their explainability and transparency. As an AI teammate will likely be the first to recognize an essential change to the team's environment, they will also likely be the first to alter their behavior to meet or accommodate that change. As such, this scenario provides the perfect opportunity to contribute to the team's situational awareness and state that their new behavior was done in response to that change in the environment.

By providing a meaningful contribution to team situational awareness, the AI teammate has now also provided human teammates with a deeper understanding of how they operate and make decisions simultaneously. AI teammates who take these opportunities to convey simple messages that convey intent and reasoning through the situational awareness update have essentially killed two birds with one stone. The predictability of the AI has been enhanced for the human teammates, their utility has increased, and affective attitudes such as trust have improved. This design would allow human-AI teams to coordinate better and would also likely lead to an increased ability to do that coordination implicitly over time as teammates begin to expect and account for how the AI will react. That being said, the qualitative data from Study 2B also conveys how important the current context is to these types of communications and factors such as time pressure, competitiveness, the importance of environmental sounds to accomplishing the team task, and other context and task-specific variables should be accounted for in the design of AI teammates. However, a design accounting for these facets is not outside the realm of possibility as methods such as task analysis and interviews with SMEs help reveal what factors are essential to watch for when considering communication in various tasks.

4.7.2 AI Teammates Should Free Up Cognitive Resources in Monitoring Processes and Act As A Dynamic Transactive Memory System

Given the significant advantages that computationally based technologies have over humans regarding processing speed and bandwidth, the ability of AI teammates to manage monitoring processes and serve as a team memory hub. As teams enter the action stage, they must engage in several vital functions, many of which involve monitoring aspects critical to the team. Specifically, teams must monitor progress towards individual and team goal completion, monitoring of the system, and teammate monitoring [23, 196]. The results of Study 2 strongly indicated that human members of AI teams expect their AI teammates to be capable of managing many of these monitoring processes for the team. This desire was strongest for monitoring processes that took the most cognitive resources away from the human teammates, such as repairing defenses, which took their focus off the main task of engaging the opposing team and adapting strategy accordingly. Many of these monitoring tasks can be managed by an AI teammate who does not suffer from cognitive overload in the same way human teammates would and would be capable of addressing several items simultaneously without adverse effects. However, designers should be careful not to choose processes that are too critical, as past research has described the lumberjack effect, which can lead to significant negative consequences in the case of a failure by the automated system [237]. Related to having AI teammates engage in monitoring processes that save human teammates' cognitive resources, AI teammates can also contribute to team cognition by acting as a transactive memory system hub. Study 2 showed that many participants are aware of the advantages of AI teammates and want them to be capable of sharing their in-depth knowledge of the task space and or system with the team to contribute to strategy development. Furthermore, AI teammates can combine their monitoring process management task with being a transactive memory hub by observing the task space for necessary resources or items related to the team's objective and keeping track of them to share that information with the team when needed or requested. This type of memory organization represents a transactive memory system that places the AI teammate in charge of knowledge related to highly dynamic and specific information, which AI excels in understanding and maintaining.

4.7.3 AI Teammates Should Encourage and Elicit Team Cognition Behaviors from Human Teammates

The final design recommendation is also one of the most significant that stems from Studies 2A and 2B as it stipulates the potential for AI teammates to be exemplars in not only task execution but also effective team cognition behaviors that develop and support the construct. Specifically, AI can engage in exemplar behavior to showcase effective team cognition-developing behaviors, encourage communication from teammates, convey what situational awareness updates are most effective or influential in specific contexts, recommend effective strategies based on teammate positioning, and showcase effective information pushing and pulling for each role on the team. Demonstrating effective information pulling and pushing by an expert AI teammate has already been shown to have a positive effect on the information-sharing tendencies of human teammates [212], which gives credence to the idea that an AI teammate is capable of encouraging other behaviors beneficial to team cognition. Further still, Study 2A showed how statements by an AI teammate could influence teammates' attitudes towards other human teammates, which can be utilized for positive means by fostering enhanced levels of trust and trust resiliency after a mistake to improve team cohesion. Going beyond the design recommendations from Study 1, which highlighted that AI teammates could showcase and highlight exemplar behavior for task execution, especially for new teammates, AI can and should also exemplify what effective team cognition-supporting behaviors are to human teammates. In doing so, the AI teammate can emphasize how and when to share critical information and directly benefit the team's development of a shared mental model (from a shared knowledge perspective) while also enhancing the processes the team goes through to complete their shared task (interactive team cognition perspective). Taking the AI exemplar concept a step further would be incredibly beneficial for human-AI teams and represent a distinct advantage that human-AI teams would have over human-only teams, especially when comparing novice teams. These human-AI teams would be capable of understanding their task quicker, less prone to bad communication habits, and be explicitly aware of effective information pushing and pulling, with the potential to have better information anticipation habits as well.

4.8 Study 2A and 2B: Limitations and Future Work

The findings of Study 2A and 2B impact the current dissertation and the related literature and practice; however, they are not without their limitations that give rise to future research. The first limitation to note is the use of a factorial survey design with vignettes in Study 2A, which, while having high internal and measurement validity [179, 317], are still not the same as having entire teams go through a lived experience. Real simulated teaming can better replicate environmental constraints, cognitive overload, and the need for team situational awareness among team members. As such, future research should take findings similar to those from Study 2A and determine how they translate to simulated teaming scenarios in a lab or in real-world teams through case studies and field studies. Study 2A was also limited by the number of survey questions participants could answer before survey fatigue set in, so in-person studies such as those previously mentioned could administer more comprehensive measures that may showcase additional nuances in the effects of the manipulations. Additionally, both Study 2A and 2B suffered from a common limitation when recruiting participants from competitive video gaming communities [336, 225], which is the predominance of participants identifying as male. While the current research was not devoid of perspectives from those identifying as female, future research must gather as many perspectives as possible from those with experience in human-AI teams to ensure the collective findings of the literature are adequately generalizable.

4.9 Study 2A and 2B: Conclusion

The results of Study 2A and 2B provide valuable information and insights into how an AI teammate contributing to team cognition could do so effectively and without complicating the task or introducing catastrophic consequences for autonomy failures. Study 2A found that by enhancing the utility of AI teammates through information-sharing attributes like explainability, augmenting team memory, and level one situational awareness updates, human teammates' attitudes are improved, and perceived levels of team cognition are improved. While adverse effects were seen for the human teammate when the AI provided a back-up behavior correcting them, no other negative effects on the human teammate were seen, and none of the AI SA attributes significantly increased the measure of participants' perceived task disruption by the AI. Additionally, Study 2B found that humans want AI teammates with agency in interdependent roles as long as they are predictable. Human teammates also envision AI teammates as capable of providing relevant situational awareness updates to the team. They can also act as an exemplar to the team to demonstrate possible actions, strategy, system usage, and effective team cognition-supporting behaviors. These results contribute to the current dissertation's research motivations by demonstrating what humans working in human-AI teams want from their AI teammates regarding enhancing team cognition and attitudes. They also show how humans would respond to AI teammates with various information-sharing attributes designed to improve team cognition. Such findings can be used practically and respond to the problem motivation by addressing how AI teammates can be designed to leverage their differences and utilize computational advantages to enhance team cognition and how that can be done in a variety of different contexts without degrading teammates' comprehension of the task or complicating it. Finally, these results identify the importance of shared situational awareness to human-AI teams, making it an aspect of team cognition critical to human-AI teams. Study 2 also highlights the oversized role the AI teammate plays in human teammates' individual and team situational awareness and their perceptions and attitudes towards the team, as AI explainability and AI predictability had significant implications for participants' team cognition. Continuing to address the research gaps of the current dissertation, the findings of Study 2 play a significant role in informing the design of Study 3 as the final study seeks to positively drive team cognition through AI teammate design targeting shared situational awareness.

Chapter 5

Study 3: Towards Purposefully Designed AI Teammates for Team Cognition in Human-AI Teams

5.1 Study 3: Overview

Study 3 represents a culmination of the work completed in Studies 1 and 2, as the results of the previous work are being used to derive AI teammate designs meant to contribute to, encourage, and improve team cognition for actual testing in a robust empirical setting. Study 1 began at a high level and sought to explore the overarching effects of working with an AI teammate on team cognition. It found that team cognition matters greatly to human-AI teams and that information-sharing played a significant role in its development. Study 2 used these findings to delve deeper into what specific AI teammate features in the form of information-sharing attributes influenced team cognition and the attitudes necessary for AI to make those contributions to team cognition and found team situational awareness to be critically essential to human-AI teams. Team situational awareness was determined to be the component of team cognition that made itself readily available to being influenced by current AI teammate technologies and as a component with exceptional importance to human-AI teams. Both Study 1 and Study 2 found team situational awareness incredibly impactful to their understanding of the AI teammate, how they coordinated their efforts, and even how they gauged their team's performance perceptually. Study 3, as the final study of the current dissertation, explicitly utilized that information to examine if various implementations of AI teammates designed to contribute to, encourage, and improve team cognition in human-AI teams are effective. This final study provides meaningful contributions to understanding how AI teammates should be designed to enhance team cognition.

Further still, participants of Study 2 conveyed several suggestions for how they wanted their AI teammate to be designed, and the theme supporting team situational awareness was extensive. Study 3's experimental design and task selection were also carefully selected to enable this, as the time spent working together as a team is considerably longer in Study 3 compared to previous studies of the dissertation. Each team completed multiple phases of the task, allowing the study to better understand how these implementations of AI design can *sustain* these aspects of team cognition. This increase in time is critical as the influence of the AI teammate characteristics on team cognition takes time to develop [278]. The development occurs as team members build their mental models and various levels of situational awareness through trial and error and experience interacting with their teammates. The increase in the time participants spent engaged in teamwork with one another also allowed Study 3 to examine AI participation in transition periods, which are the phases teams spend together discussing their previous team performance and planning for future team actions. Because transition periods are when teams engage in mission analysis and strategy formation processes [195], it is critical to understand teammates' efforts to further develop team cognition, especially with the inclusion of an AI teammate. This work will provide essential results across several key concepts in human-AI teaming: 1) what are the best design features of AI teammates to utilize for team situational awareness; 2) how do these design features of AI teammates influence other related components of team cognition like shared mental models; and 3) can the team processes indicative of team situational awareness, such as team communication, be improved through AI participation in transition phases early or late in their life cycle? These results directly address the following research questions specific to the proposed Study 3:

RQ1.2: Do human teammates accept contributions to situational awareness from AI teammates in a complex hands-on task?

RQ3.3: Is AI participation in transition phases more beneficial to human-AI teams' situational awareness if it occurs earlier or later in their life cycle?

RQ4.3: Which AI teammate situational awareness attribute best supports developing and sustaining team situational awareness in human-AI teams over time?

The above research questions provide the final components of a detailed research venture into the functioning and improvement of team cognition in human-AI teams by focusing on empirically testing implementations of AI design features for enhancing team cognition. Namely, Study 3 acts as the final leg of the journey that Studies 1 and 2 started. These former studies each explored team cognition in human-AI teams and subsequently identified components of the construct that were important to human-AI teams *and* benefited from AI teammates' technical abilities. Study 3 finalizes these contributions by empirically examining the efficacy and holistic effect AI teammates designed to support team cognition actually have in practice

[59]. Both future research and applied practice benefit from this work by gaining insights into where AI teammates' design could be improved to support and enhance team situational awareness. The final study of the current dissertation also improves the understanding of how AI participation in transition phases affects team cognition, which has never been investigated within the context of human-AI teaming.

5.2 Study 3: Methods

Study 3 is a major in-person experiment utilizing a mixed-methods design implemented with teams of three that complete four missions together over three hours.

5.2.1 Experimental Design

To answer the study-specific research questions of Study 3, the experiment implemented two between-subjects manipulations: 1) order of AI transition phase participation; and 2) AI situational awareness attribute. Specifically, the AI transition phase participation order was crossed with AI situational awareness attribute, and it consisted of two levels: 1) first transition participation; and 2) second transition participation. The AI situational awareness attribute contained three between-subjects levels: 1) augmenting team memory; 2) intra/extra team information changes; and 3) control. These manipulations allowed the study to investigate the role of AI participation in transition phase processes on team cognition development, how the AI design features support the development of team situational awareness over time, and how these two aspects of an AI teammate may interact with one another to affect team cognition.

Given the many levels of the manipulations and the length of their names, they will all be referred to using acronyms moving forward. As such, the order of AI transition phase participation manipulations will be referred to as follows: 1) AI participation in the first transition phase will be titled **PN** (participation-no participation); and 2) AI participation in the second transition phase will be titled **NP** (no participation-participation). The AI situational awareness attributes (AI SA attributes) will be referred to as follows: 1) augmenting team memory will be titled **ATM**; 2) intra/extra team information changes will be titled **SA2** to align with the naming of Study 2; and 3) control will still be referred to as control. This naming structure can be referenced in Table 5.1.

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	**SA2**
Control	N/A

Table 5.1: Acronym Key for Manipulation Levels

5.2.2 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 [57]. Further still, the CERTT UAS-STE has a long track record in team research [57, 64, 212] and team cognition research specifically [113, 51, 52]. The system has also been extensively utilized for analysis on human-AI teams [212, 68].

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 5.1); 2) the navigator (DEMPC) develops and provides the flight plan and any specific flight restrictions such as speed and altitude (see Figure 5.2); and 3) the photographer (PLO) monitors the sensors and is in charge of taking photographs of the actual targets in accordance with the current altitude, airspeed, and target specific camera requirements (e.g., zoom level and camera type) (see Figure 5.3). 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 chat screen.

The AVO and PLO roles were taken by two participants recruited for the study, while the AI teammate took the DEMPC role. The missions in the current experiment using the CERTT UAS-STE included four missions that were 20 minutes in length, preceded by 30 minutes of training involving a standard 15-minute interactive guided PowerPoint and a 15-minute hands-on training mission. This training familiarized the participants with both the system and their teammates and allowed participants to ask questions about the system to the experimenter. 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.

Figure 5.1: View of the Pilot's Console During A CERTT UAS-STE Mission.

5.2.3 AI Teammate

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 [203]. Specifically, for Study 3, the AI teammate's activities within the simulation and its chat communications were scripted, and these scripts were developed from several pilot sessions. Additionally, given the breadth of the tasks AI teammates would have needed to engage in to support transition phase processes fully and to simplify the development of the script, the AI teammates were not allowed to engage in any communication outside of their specific team situational awareness feature and only responded in the affirmative or negative in the case of requests of the AI teammate for future missions. For example, the SA2 AI was described as being capable of responding to questions

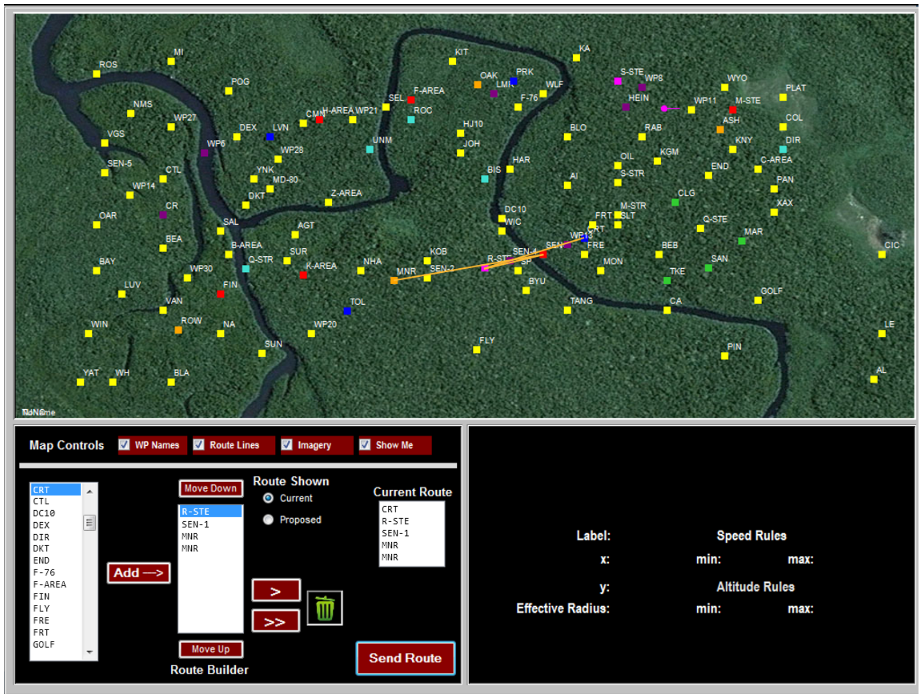


Figure 5.2: View of the Navigator's Console During A CERTT UAS-STE Mission.

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). The participants were made aware of these boundaries and capabilities for their AI teammates before the hands-on training session.

The AI situational awareness information attributes were pilot-tested by fine-tuning their specific manipulations within the CERTT UAS-STE to adequately accommodate the task and their particular attribute. Each attribute had detailed mission information that they would share with the team to enhance their team situa-

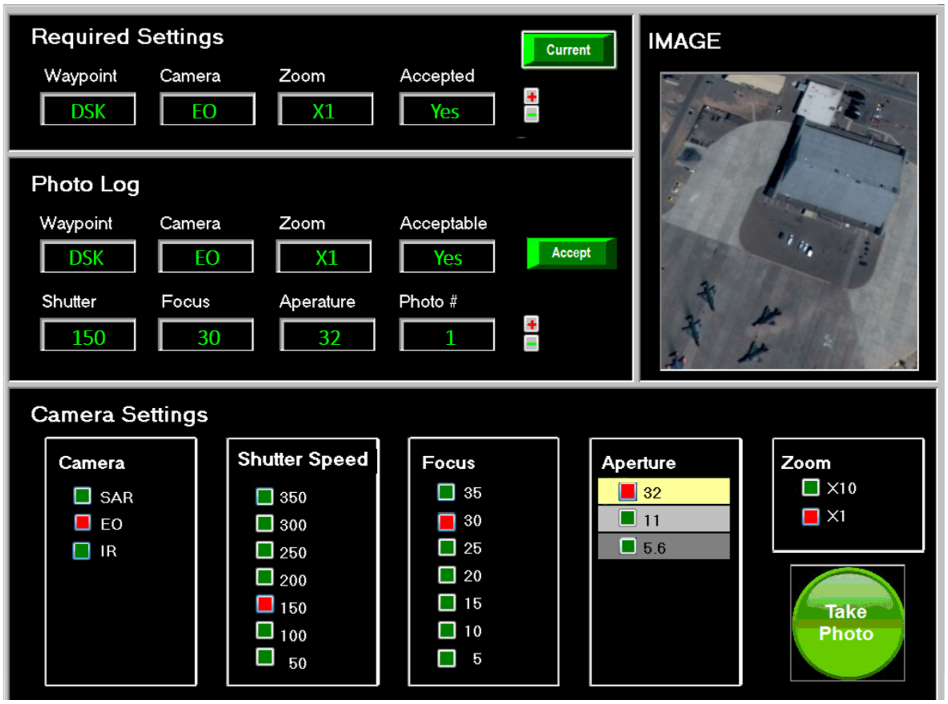


Figure 5.3: View of the Photographer's Console During A CERTT UAS-STE Mission.

tional awareness and coordination meaningfully while providing a specific response to roadblocks that teams encountered each mission. First, the ATM AI teammate offered additional information to the team members to assist them in preparing and planning ahead for each future ROZ when it became relevant and the task load was low (between ROZs). This information included the number of targets in the upcoming ROZ, an airspeed range that accommodated the most targets, and whether that speed needed to change for a specific target. The ATM AI would remind the team just before that change in speed became necessary and the name of the target where the speed change would become required. When the teams encountered a roadblock for a mission, the ATM AI would notify the team of the system failure, what information it occluded, the teammate affected and described what information that teammate needed to accomplish their job effectively. Second, the SA2 AI provided explanations for backtracking whenever a priority target was on the route, alerted the team when dusk had fallen, and provided back-up behavior information whenever the UAS was not meeting the restrictions communicated by the DEMPC or was obstructing targets from being photographed. When the teams encountered a roadblock for a mission, the SA2 AI provided notification to all team members of the system failure, what information was occluded from what teammate, and then explained to the team that they would provide the necessary information to the teammate so they could still accomplish their job effectively. Lastly, the control AI teammate provided all of the information requisite to the DEMPC role and high-level information relevant to the general nature of the CERTT UAS-STE if they were asked a question regarding the system or task. The control AI teammate did not automatically provide information regarding the situational awareness roadblock.

5.2.4 Participants

Sixty-four participants were recruited for Study 3, but two teams were dropped due to technical difficulties. This left 60 participants' data to be analyzed. These 60 participants had an average age of 20.4 ($SD = 3.13$). Forty-five participants identified as women, and the rest identified as men. Each team included two participants, with 32 total teams recruited for the study. An a priori power analysis using GPower and PANGEA [322, 91] suggested a minimum of 54 participants were needed to achieve a power of .8 with a medium effect size ($\eta_p^2 = .10$) and six repeated measures for survey metrics. However, because performance metrics only consisted of four repeated measures (one for each mission), the power analysis suggested a minimum of 60 participants to achieve adequate power for tests involving only four repeated measures. The current study achieved the minimum level of recruitment to reach sufficient power, and these estimates are also in line with past teaming research utilizing the CERTT

UAS-STE [212, 68, 211].

All participants were recruited during the Summer and early Fall using flyers and through a Clemson University-affiliated participant pool. Those recruited using flyers posted around the Clemson University campus and social media posts to Clemson-related online communities on platforms such as Reddit and Discord were compensated with $30 Amazon gift cards for their time. The participants recruited using the participant pool were compensated with course credit for their time.

5.2.5 Procedure

The study took three hours to complete (see Figure 5.4), 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. Participants providing informed consent continued to complete the pre-task demographics survey before being required to complete the interactive PowerPoint training presentation. Once the interactive PowerPoint training was completed, the participants could ask questions. They were also reminded of several topics from the interactive PowerPoint training that pilot testing showed as common areas of confusion. 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 shorter in length. 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 their situational awareness attributes. The team was then ready to begin actual missions and started the first shortly after the training mission. Once the first mission was completed, the participants completed one of the six post-task surveys before proceeding to the transition phase discussion. These surveys were conducted on separate tablets with integrated keyboards and trackpads. Before the transition phase discussion, the researcher described the purpose of the transition phase (e.g., to analyze and discuss past performance and plan for future missions) and detailed whether or not their AI teammate would be joining them during the upcoming transition phase. The transition phase started once the researcher navigated the participants' tablets to the Slack application, where their text-based conversation would occur in a locked session-specific channel. Transition phases lasted for six minutes, at which point the researcher directed the participants' tablets back to the survey, where the second post-task survey would be completed. 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 long 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 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.

5.2.6 Measures

The measures taken in Study 3 were much more focused than the measures taken in the previous two studies of the dissertation, with two types of measures implemented in the current study. Specifically, several task-derived measures were enabled using the CERTT UAS-STE; however, several survey-based measures were also still used.

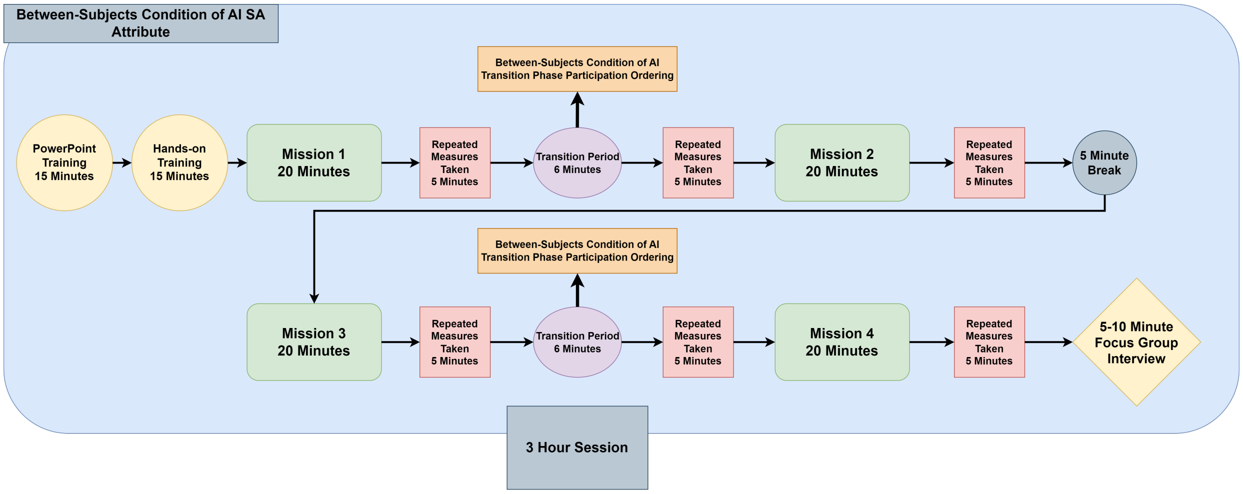


Figure 5.4: Study 3 Session Timeline, Indicating the Frequency of Repeated Measures and the Placement of Transition Periods Along with the Time Associated for Each.

5.2.6.1 Survey Based Measures

Trust in the Human and AI Teammates. Participants' trust in their two teammates was measured utilizing methods similar to Study 1. Specifically, trust in each teammate was measured using the custom-made scale based on the outcomes of trust identified by [186] and utilized in previous human-AI teaming trust research [280]. This scale included six items rated on a seven-point Likert scale ranging from "Strongly Disagree" to "Strongly Agree." Participants answered the six items for each teammate for a total of 12 items and answered them at each repeated measures point shown in the timeline seen in Figure 5.4. Responses to each set of six items were averaged together for each participant, and higher values indicate greater trust in the teammate.

Perceived Situational Awareness. The participants' perceived situational awareness was measured using the Situational Awareness Rating Technique (SART) [304]. 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 situational awareness.

Perceived Shared Mental Model. Participants' perceived shared mental models used a method similar to Study 2. This measure was taken using the scale developed by van Rensburg and colleagues [310]; 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 [202]. These participants rated each teammate individually for a total of 20 items at each repeated measures point depicted in Figure 5.4. Responses for each set of ten items were averaged for each participant, and higher values indicated a greater perception of a shared mental model with that teammate.

Perceived Team Effectiveness. Participants' perceived performance of their team was measured at each repeated measures point shown in Figure 5.4 using the team effectiveness scale developed by Rentsch and colleagues [259]. This scale utilizes eight items rated on a seven-point Likert scale ranging from "Strongly Disagree" to "Strongly Agree." Responses to the eight items were averaged for each participant, and higher values indicated greater perceived team performance.

5.2.6.2 Task Derived Measures

Target Processing Efficiency (Team Performance). Participants' team performance was measured as an outcome-based measure of team effectiveness. Specifically, team performance was measured using target processing efficiency. Target processing efficiency uses the time teams spend inside a target's effective radius to get a good photo. As such, a higher score is indicative of greater team efficiency. Each team begins a mission with 1,000 points for each target, and points are deducted from that total based on the number of seconds inside the effective radius of the target, and there is a 200-point penalty for not taking a good photo [50]. Scores for each target were averaged by mission, providing an average target processing efficiency score for each mission with higher values indicating greater objective team performance.

Team Situational Awareness (CAST). Team situational awareness was measured using task perturbations or "roadblocks," which were implemented in the current study using pre-programmed system failures built into the CERTT UAS-STE [49]. These roadblocks were used in conjunction with the Coordinated Awareness of Situation by Teams (CAST) metric developed by Gorman and colleagues for the CERTT UAS-STE environment [112]. This method of assessing team situational awareness was created through previous CERTT UAS-STE research and has frequently been utilized in human-AI teaming research on situational awareness [212, 49]. 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 across the team to mitigate the roadblock effectively [111]. 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 5.5), and each score is calculated by dividing the number of marked boxes by the total number of boxes available. Using CP as an example, if Figure 5.5 showed both boxes in the bottom left for PLO as being marked, it would convey that the PLO teammate had received a communication informing them of the roadblock (system failure) from the AVO and DEMPC teammate. In this example of CP scoring, the team would score 0.33 because two of the six boxes were marked. These measures of CP and CA can also be averaged together for a total CAST score. Teams' were also rated on their ability to successfully overcome the system failure by whether they could take a good photo of the priority target that triggered the failure, resulting in a binary outcome of yes or no.

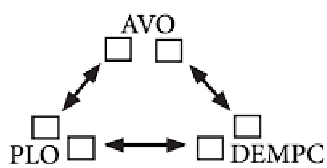


Figure 5.5: Situational Awareness Logger.

Team Verbal Behaviors (Team Process). Based on past CERTT UAS-STE experiments, five team verbal behaviors have been associated with effective teaming (see Table 5.2). All team communications were examined by an experimenter throughout each mission completed by teams, and these communications were tagged using a guide developed by past CERTT UAS-STE experiments. For the current task, the pushing verbal behaviors were associated with more effective teamwork and were indicative of superior team processes [67]. These team verbal behaviors are associated with effective teaming as they show that teammates understand the information needs of the other roles in relation to their own and when the other roles need that information [67, 212].

Behaviors	Behavior Type	Description
General Status Update	Push	Alerting teammates of current status
Suggestion	Push	Suggesting something to another teammate
Planning Ahead	Push	Creating rules and procedures for future goals or tasks
Repeated Requests	Pull	Requesting information or action from teammates more than once
Inquiry About the Status of Others	Pull	Asking about teammates' current status or expressing concerns

Table 5.2: Team Verbal Behaviors Captured Using the CERTT UAS-STE. (Table Adapted from [212])

5.2.6.3 Focus Group Interview

Study 3 also implemented a focus group interview with teams after they completed the final mission and survey (see Figure 5.4). One trained experimenter led the focus group interview following a semi-structured interview protocol. The interview lasted between 5 and 10 minutes, with questions focusing on how the participants perceived the AI teammate's contributions to team situational awareness, their experience going through the transition phases with and without the AI teammate, and how their ability to cooperate effectively developed throughout the experience. Thus, allowing the study to provide more comprehensive answers to the three RQs posed by Study 3.

The focus group interview data was transcribed using Otter.AI and then analyzed using thematic analysis [22, 102, 301, 103], which was the same method utilized in Study 1 and 2. Specifically, the data was analyzed in four phases: 1) each transcript was reviewed to gain an understanding of how the participants reacted to the presence of the AI, or lack thereof, in transition phases and how the AI teammate's SA attribute influenced their situational awareness development; 2) the transcripts were re-reviewed to identify major themes and sub-themes describing the AI teammate's influence on their situational awareness; 3) the themes and sub-themes were reviewed and discussed with a colleague familiar with the goals of Study 3 until agreement was achieved; 4) individual quotes were selected to represent the themes and sub-themes; and 5) themes and sub-themes were reviewed one final time with the same colleague from Step 3 using the quotes selected in Step 4 to ensure the results were a representative distillation of the participants experiences on the ability of AI teammates to contribute to situational awareness.

5.3 Study 3: Quantitative Results

The following results section is partitioned into two distinct sub-sections. The first sub-section consists of the four measures taken by the survey: 1) trust in the two teammates; 2) perceived shared mental model with each teammate; 3) perceived situational awareness; and 4) perceived team effectiveness. The second sub-section focuses on the three measures taken by the CERTT UAS-STE: 1) target processing efficiency; 2) team situational awareness; and 3) team verbal behaviors. The analysis of these metrics addresses RQ1.2, 3.3 and 4.3 by examining participants' acceptance of the AI teammate's contributions (RQ1.2), how situational awareness is affected by the time point an AI teammate participates in a transition phase (RQ3.3), and which AI situational awareness attribute is best suited to developing and sustaining team situational awareness (RQ4.3). All dependent variables were checked for violations of sphericity, and Greenhouse-Geisser corrections were utilized when necessary. Normality and homoscedasticity were also checked for each variable, and no major violations were found, especially given ANOVA's robustness to minor violations when equal cell sizes are present [108, 325, 256].

5.3.1 Survey Measures

The current section reports on the analysis of the following measures: 1) trust in the AI teammate; 2) trust in the human teammate; 3) perceived shared mental model with the AI teammate; 4) perceived shared mental model with the human teammate; 5) perceived situational awareness; and 6) perceived team effectiveness. These measures were taken and analyzed at the individual level.

5.3.1.1 Trust in AI Teammate

Cases	Sphericity Correction	Sum of Squares	*df*	Mean Square	*F*	*p*	η^2_p
Time	Greenhouse-Geisser	2.943	3.785	0.777	2.770	0.031	0.049
Time * AI SA Attribute	Greenhouse-Geisser	2.849	7.570	0.376	1.341	0.228	0.047
Time * AI Transition Phase Ordering	Greenhouse-Geisser	1.866	3.785	0.493	1.757	0.143	0.032
Time * AI SA Attribute * AI Transition Phase Ordering	Greenhouse-Geisser	1.214	7.570	0.160	0.571	0.792	0.021
Residuals	Greenhouse-Geisser	57.360	204.383	0.281			

Table 5.3: Within-Subjects Effects on Trust in AI Teammate.

Cases	Sum of Squares	*df*	Mean Square	*F*	*p*	η^2_p
AI SA Attribute	3.425	2	1.713	1.738	0.186	0.060
AI Transition Phase Ordering	0.316	1	0.316	0.321	0.574	0.006
AI SA Attribute * AI Transition Phase Ordering	3.140	2	1.570	1.593	0.213	0.056
Residuals	53.218	54	0.986			

Table 5.4: Between-Subjects Effects on Trust in AI Teammate.

A 2 (Order of AI Transition Phase Participation: PN, NP) x 3 (AI Situational Awareness Attribute: ATM, SA2, Control) x 6 (Time Point: M1, TP1, M2, M3, TP2, M4) mixed RMANOVA was conducted to assess the effect of AI SA attribute

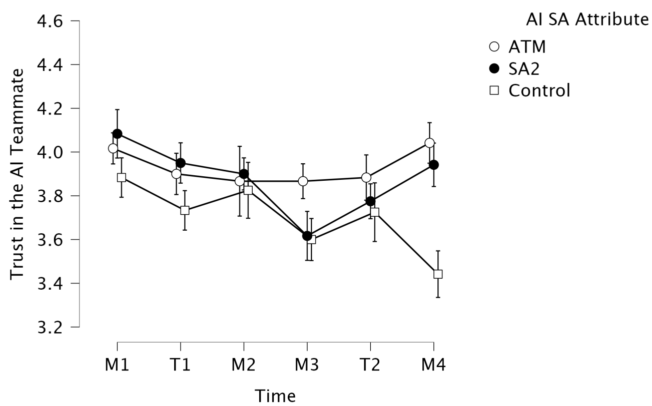


Figure 5.6: Trust in the AI Teammate by Time. Error Bars Represent Standard Error.

and transition phase participation ordering (both between-subjects) on participants trust in their AI teammate over time (within-subjects) (see Tables 5.3 and 5.4). The analysis indicated a significant main effect of time on participants' trust in the AI teammate ($F(3.79, 204.38) = 2.77$, $p = .031$, $\eta_p^2 = .05$; see Figure 5.6). Tukey's HSD post-hoc tests revealed that trust after M1 ($M = 3.99$, $SE = .08$) was significantly greater than trust after M3 ($M = 3.55$, $SE = .08$).

5.3.1.2 Trust in Human Teammate

Cases	Sphericity Correction	Sum of Squares	df	Mean Square	F	p	η_p^2
Time	Greenhouse-Geisser	1.679	3.962	0.424	2.097	0.083	0.037
Time * AI SA Attribute	Greenhouse-Geisser	1.832	7.925	0.231	1.144	0.335	0.041
Time * AI Transition Phase Ordering	Greenhouse-Geisser	1.021	3.962	0.258	1.275	0.281	0.023
Time * AI SA Attribute * AI Transition Phase Ordering	Greenhouse-Geisser	1.448	7.925	0.183	0.904	0.513	0.032
Residuals	Greenhouse-Geisser	43.247	213.966	0.202			

Table 5.5: Within-Subjects Effects on Trust in the Human Teammate.

A 2 (Order of AI Transition Phase Participation: PN, NP) x 3 (AI Situational Awareness Attribute: ATM, SA2, Control) x 6 (Time Point: M1, TP1, M2, M3, TP2, M4) mixed RMANOVA was conducted to assess the effect of AI SA attribute and

Cases	Sum of Squares	*df*	Mean Square	*F*	*p*	η_p^2
AI SA Attribute	0.075	2	0.038	0.045	0.956	0.002
AI Transition Phase Ordering	0.006	1	0.006	0.008	0.931	< .001
AI SA Attribute * AI Transition Phase Ordering	0.126	2	0.063	0.076	0.927	0.003
Residuals	44.834	54	0.830			

Table 5.6: Between-Subjects Effects on Trust in the Human Teammate.

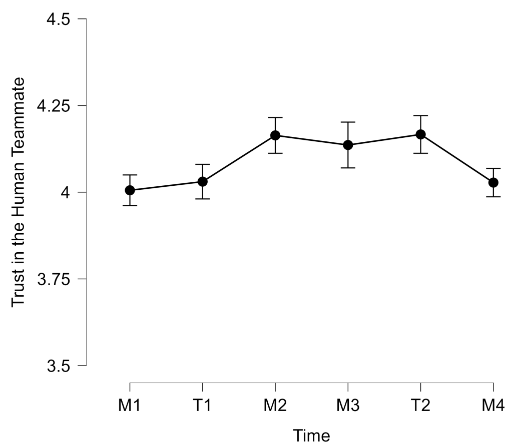


Figure 5.7: Trust in the Human Teammate by Time. Error Bars Represent Standard Error.

transition phase participation ordering (both between-subjects) on participants trust in their human teammate over time (within-subjects) (see Tables 5.5 and 5.6). This analysis did not reveal any significant effects on participants' trust in their human teammate.

5.3.1.3 Shared Mental Model with AI Teammate

A 2 (Order of AI Transition Phase Participation: PN, NP) x 3 (AI Situational Awareness Attribute: ATM, SA2, Control) x 6 (Time Point: M1, TP1, M2, M3, TP2, M4) mixed RMANOVA was conducted to assess the effect of AI SA attribute and transition phase participation ordering (both between-subjects) on participants shared mental model with their AI teammate over time (within-subjects) (see Tables

Cases	Sphericity Correction	Sum of Squares	df	Mean Square	F	p	η_p^2
Time	Greenhouse-Geisser	5.369	3.057	1.756	2.628	0.051	0.046
Time * AI SA Attribute	Greenhouse-Geisser	11.358	6.114	1.858	2.779	0.013	0.093
Time * AI Transition Phase Ordering	Greenhouse-Geisser	0.879	3.057	0.288	0.430	0.735	0.008
Time * AI SA Attribute * AI Transition Phase Ordering	Greenhouse-Geisser	1.775	6.114	0.290	0.434	0.858	0.016
Residuals	Greenhouse-Geisser	110.342	165.080	0.668			

Table 5.7: Within-Subjects Effects on Shared Mental Model with the AI Teammate.

Cases	Sum of Squares	df	Mean Square	F	p	η_p^2
AI SA Attribute	13.531	2	6.766	1.089	0.344	0.039
AI Transition Phase Ordering	0.671	1	0.671	0.108	0.744	0.002
AI SA Attribute * AI Transition Phase Ordering	19.348	2	9.674	1.557	0.220	0.055
Residuals	335.495	54	6.213			

Table 5.8: Between-Subjects Effects on Shared Mental Model with the AI Teammate.

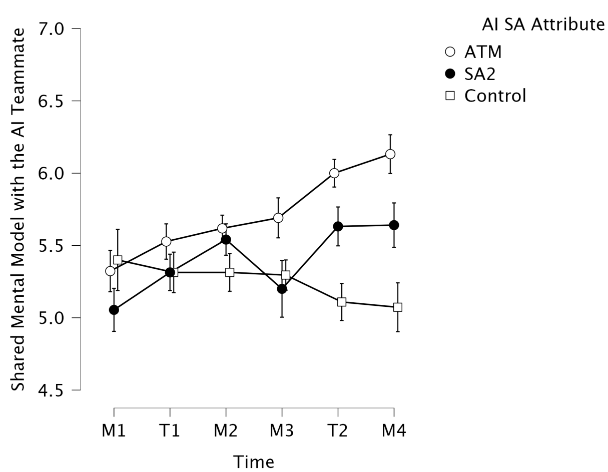


Figure 5.8: Shared Mental Model with the AI Teammate by Time and AI SA Attribute. Error Bars Represent Standard Error.

5.7 and 5.4). The analysis indicated a significant ordinal interaction effect between time and AI situational awareness attribute ($F(6.11, 165.08) = 2.78$, $p = .013$, $\eta_p^2 = .09$; see Figure 5.8). A simple main effects analysis of the interaction revealed that the control condition of the AI SA attributes was the only one not to see participants' shared mental model with their AI teammate change significantly across time.

Additionally, the participants' shared mental model with the AI teammate was only significantly different across the three AI SA attribute conditions in M4, with ATM (M = 6.13, SE = .28) being significantly higher than the control (M = 5.07, SE = .28).

5.3.1.4 Shared Mental Model with Human Teammate

A 2 (Order of AI Transition Phase Participation: PN, NP) x 3 (AI Situational Awareness Attribute: ATM, SA2, Control) x 6 (Time Point: M1, TP1, M2, M3, TP2, M4) mixed RMANOVA was conducted to assess the effect of AI SA attribute and transition phase participation ordering (both between-subjects) on participants shared mental model with their human teammate over time (within-subjects) (see Tables 5.9 and 5.10). There was a significant effect of time on participants' shared mental model with their human teammate ($F(3.39, 183.07) = 12.34$, $p < .001$, $\eta_p^2 = .19$; see Figure 5.9). Tukey's HSD post-hoc tests found that participants' shared mental model with their human teammate was significantly lower in M1 (M = 5.85, SE = .10) than in T1 (M = 6.18, SE = .10), M2 (M = 6.31, SE = .10), M3 (M = 6.24, SE = .10), T2 (M = 6.40, SE = .10), and M4 (M = 6.40, SE = .10).

Cases	Sphericity Correction	Sum of Squares	df	Mean Square	F	p	η_p^2
Time	Greenhouse-Geisser	12.877	3.390	3.798	12.336	< .001	0.186
Time * AI SA Attribute	Greenhouse-Geisser	1.190	6.780	0.175	0.570	0.774	0.021
Time * AI Transition Phase Ordering	Greenhouse-Geisser	0.722	3.390	0.213	0.692	0.575	0.013
Time * AI SA Attribute * AI Transition Phase Ordering	Greenhouse-Geisser	3.370	6.780	0.497	1.614	0.136	0.056
Residuals	Greenhouse-Geisser	56.366	183.071	0.308			

Table 5.9: Within-Subjects Effects on Shared Mental Model with the Human Teammate.

The analysis also found a significant disordinal interaction effect between AI SA attribute and AI transition phase ordering on participants' shared mental model

Cases	Sum of Squares	df	Mean Square	F	p	η_p^2
AI SA Attribute	0.433	2	0.216	0.083	0.920	0.003
AI Transition Phase Ordering	0.031	1	0.031	0.012	0.913	< .001
AI SA Attribute * AI Transition Phase Ordering	31.013	2	15.507	5.976	0.005	0.181
Residuals	140.129	54	2.595			

Table 5.10: Between-Subjects Effects on Shared Mental Model with the Human Teammate.

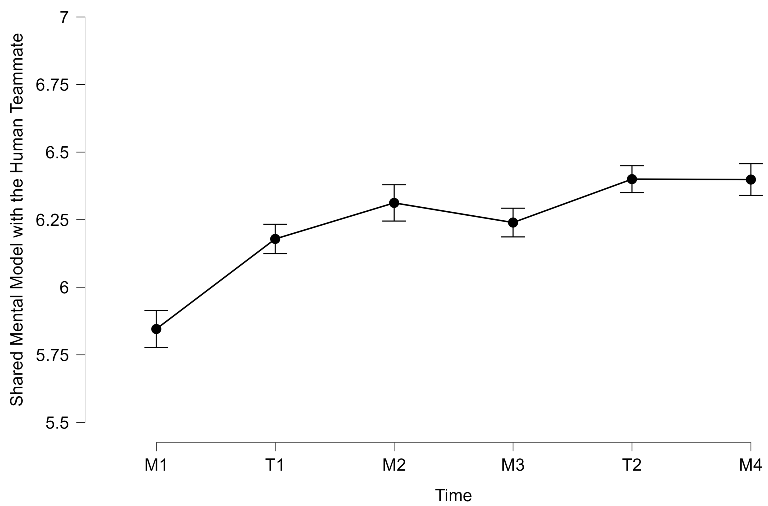


Figure 5.9: Shared Mental Model with the Human Teammate by Time. Error Bars Represent Standard Error.

with their human teammate ($F(2, 54) = 5.98$, $p = .005$, $\eta_p^2 = .18$; see Figure 5.10). This interaction effect was followed up with a simple main effects analysis, which revealed that there was a significant difference in participants' shared mental model with their human teammate between the two AI transition phase orders for the control condition only ($F(1) = 7.06$, $p = .016$). Specifically, the control condition where the AI participated in the first transition phase discussion ($M = 5.79$, $SE = .21$) was significantly lower than the control condition where the AI participated in the second transition phase discussion ($M = 6.64$, $SE = .21$).

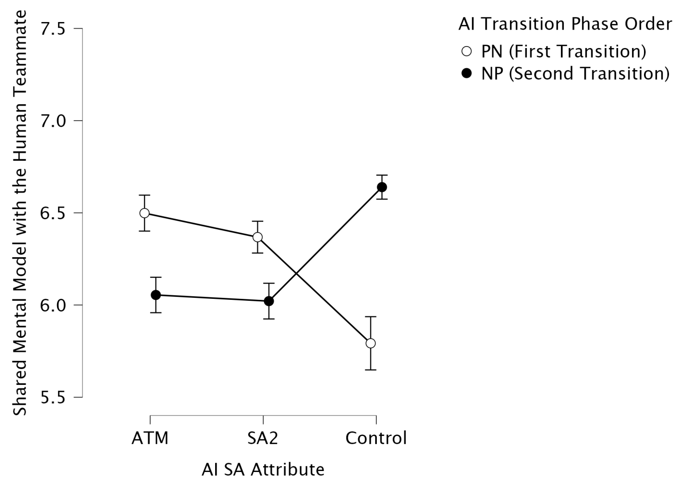


Figure 5.10: Shared Mental Model with the Human Teammate by AI SA Attribute and Transition Phase Order. Error Bars Represent Standard Error.

5.3.1.5 Perceived Situational Awareness

A 2 (Order of AI Transition Phase Participation: PN, NP) x 3 (AI Situational Awareness Attribute: ATM, SA2, Control) x 6 (Time Point: M1, TP1, M2, M3, TP2, M4) mixed RMANOVA was conducted to assess the effect of AI SA attribute and transition phase participation ordering (both between-subjects) on participants perceived situational awareness over time (within-subjects) (see Table 5.11 and 5.12). The main effect of time was found to be significant ($F(3.62,\ 195.51) = 27.25$, $p < .001$, $\eta_p^2 = .34$; see Figure 5.11). Tukey's HSD post-hoc tests found that participants' situational awareness in M1 ($M = 4.80$, $SE = .10$) was significantly lower than in M2 ($M = 5.17$, $SE = .10$), M3 ($M = 5.26$, $SE = .10$), T2 ($M = 5.24$, $SE = .10$), and M4 ($M = 5.39$, $SE = .10$). Additionally, situational awareness in T1 ($M = 4.87$, $SE = .10$) was significantly lower than M2, M3, T2, and M4. Finally, situational awareness in M2 was also significantly lower than in M4.

Cases	Sphericity Correction	Sum of Squares	df	Mean Square	F	p	η_p^2
Time	Greenhouse-Geisser	16.484	3.621	4.553	27.245	< .001	0.335
Time * AI SA Attribute	Greenhouse-Geisser	1.246	7.241	0.172	1.030	0.413	0.037
Time * AI Transition Phase Ordering	Greenhouse-Geisser	1.449	3.621	0.400	2.395	0.058	0.042
Time * AI SA Attribute * AI Transition Phase Ordering	Greenhouse-Geisser	1.498	7.241	0.207	1.238	0.283	0.044
Residuals	Greenhouse-Geisser	32.671	195.511	0.167			

Table 5.11: Within-Subjects Effects Situational Awareness (SART)

Cases	Sum of Squares	df	Mean Square	F	p	η_p^2
AI SA Attribute	19.017	2	9.508	3.038	0.056	0.101
AI Transition Phase Ordering	0.173	1	0.173	0.055	0.815	0.001
AI SA Attribute * AI Transition Phase Ordering	0.076	2	0.038	0.012	0.988	4.494×10^{-4}
Residuals	169.036	54	3.130			

Table 5.12: Between-Subjects Effects Situational Awareness (SART)

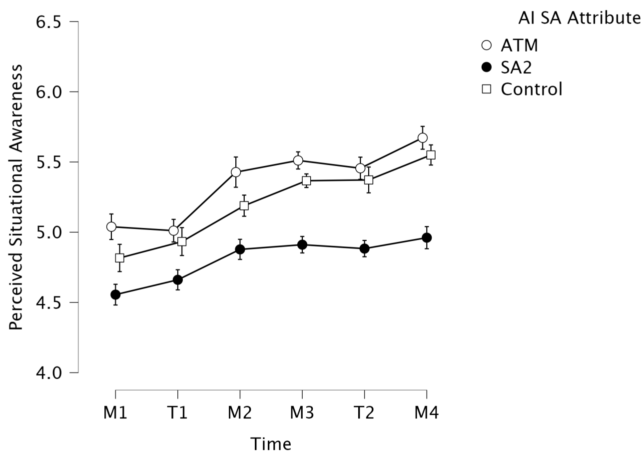


Figure 5.11: Perceived Situational Awareness by Time and AI SA Attribute. Error Bars Represent Standard Error.

5.3.1.6 Perceived Team Effectiveness

A 2 (Order of AI Transition Phase Participation: PN, NP) x 3 (AI Situational Awareness Attribute: ATM, SA2, Control) x 6 (Time Point: M1, TP1, M2, M3, TP2, M4) mixed RMANOVA was conducted to assess the effect of AI SA attribute and transition phase participation ordering (both between-subjects) on participants perceived team performance over time (within-subjects) (see Tables 5.13 and 5.14). There was a significant main effect of time on perceived team effectiveness (F(2.85, 153.65) = 17.69, p < .001, η_p^2 = .25; see Figure 5.12). Tukey's HSD post-hoc tests indicated that participants' perceived team effectiveness was significantly lower in M1 (M = 5.50, SE = .11) than M2 (M = 5.91, SE = .11), M3 (M = 5.92, SE = .11), T2 (M = 6.09, SE = .11), and M4 (M = 6.21, SE = .11). Additionally, participants' perceived team effectiveness was significantly lower in T1 (M = 6.09, SE = .11) than in T2 and M4. Finally, participants' perceived team effectiveness in M2 and M3 were both significantly lower than in M4.

Cases	Sphericity Correction	Sum of Squares	df	Mean Square	F	p	η_p^2
Time	Greenhouse-Geisser	19.874	2.845	6.985	17.688	1.404×10^{-9}	0.247
Time * AI SA Attribute	Greenhouse-Geisser	2.242	5.691	0.394	0.998	0.427	0.036
Time * AI Transition Phase Ordering	Greenhouse-Geisser	2.511	2.845	0.883	2.235	0.090	0.040
Time * AI SA Attribute * AI Transition Phase Ordering	Greenhouse-Geisser	2.123	5.691	0.373	0.945	0.462	0.034
Residuals	Greenhouse-Geisser	60.671	153.651	0.395			

Table 5.13: Within-Subjects Effects on Perceived Team Effectiveness.

Cases	Sum of Squares	df	Mean Square	F	p	η_p^2
AI SA Attribute	1.444	2	0.722	0.251	0.779	0.009
AI Transition Phase Ordering	0.951	1	0.951	0.330	0.568	0.006
AI SA Attribute * AI Transition Phase Ordering	9.804	2	4.902	1.701	0.192	0.059
Residuals	155.585	54	2.881			

Table 5.14: Between-Subjects Effects on Perceived Team Effectiveness.

5.3.2 CERTT UAS-STE Measures

The following section covers the measures taken by the CERTT UAS-STE system, which include team target processing efficiency, team situational awareness, and team verbal behaviors. These measures were taken and analyzed at the team

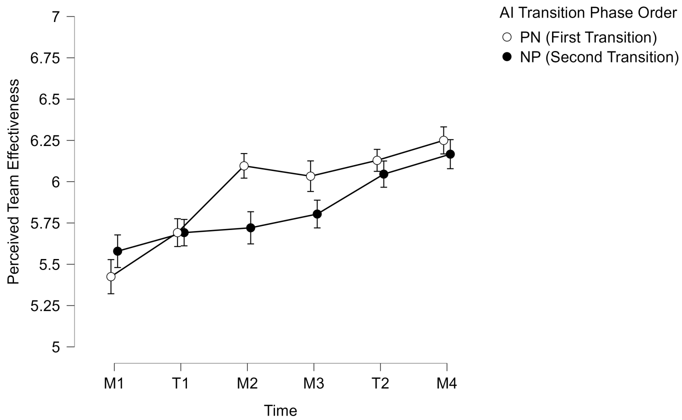


Figure 5.12: Perceived Team Effectiveness by Time and AI Transition Phase Order. Error Bars Represent Standard Error.

level.

5.3.2.1 Team Target Processing Efficiency (Score)

A 2 (Order of AI Transition Phase Participation: PN, NP) x 3 (AI Situational Awareness Attribute: ATM, SA2, Control) x 4 (Time Point: M1, M2, M3, M4) mixed RMANOVA was conducted to assess the effect of AI SA attribute and transition phase participation ordering (both between-subjects) on teams' target processing efficiency over time (within-subjects) (see Table 5.15 and 5.16). The analysis showed that AI transition phase order had a significant main effect on teams' target processing efficiency ($F(1,\ 25) = 5.13$, $p = .032$, $\eta_p^2 = .17$; see Figure 5.13). Specifically, teams that had the AI teammate participate in the first transition period ($M = 883.27$, $SE = 13.69$) had significantly lower target processing efficiency than teams whose AI teammate participated in the second transition period ($M = 927.16$, $SE = 13.69$).

5.3.2.2 Team Situational Awareness (CAST)

CAST Analysis. A 2 (Order of AI Transition Phase Participation: PN, NP) x 3 (AI Situational Awareness Attribute: ATM, SA2, Control) x 4 (Time Point: M1, M2,

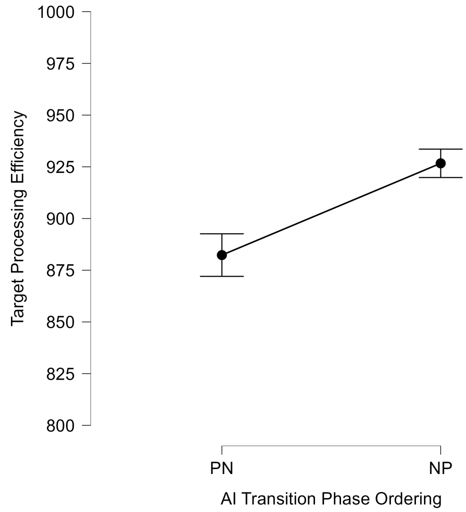


Figure 5.13: Team Target Processing Efficiency by AI Transition Phase Order. Error Bars Represent Standard Error.

Cases	Sum of Squares	*df*	Mean Square	*F*	*p*	η_p^2
Mission	16039.440	3	5346.480	1.981	0.124	0.073
Mission * AI SA Attribute	26861.777	6	4476.963	1.659	0.143	0.117
Mission * AI Transition Phase Ordering	3330.540	3	1110.180	0.411	0.745	0.016
Mission * AI SA Attribute * AI Transition Phase Ordering	11720.323	6	1953.387	0.724	0.632	0.055
Residuals	202387.223	75	2698.496			

Table 5.15: Within-Subjects Effects on Teams' Target Processing Efficiency.

Cases	Sum of Squares	*df*	Mean Square	*F*	*p*	η_p^2
AI SA Attribute	13675.380	2	6837.690	0.590	0.562	0.045
AI Transition Phase Ordering	59436.729	1	59436.729	5.131	0.032	0.170
AI SA Attribute * AI Transition Phase Ordering	2155.420	2	1077.710	0.093	0.911	0.007
Residuals	289598.287	25	11583.931			

Table 5.16: Between-Subjects Effects on Teams' Target Processing Efficiency.

M3, M4) mixed RMANOVA was conducted to assess the effect of AI SA attribute and transition phase participation ordering (both between-subjects) on team situational awareness behaviors over time (within-subjects) (see Table 5.17 and 5.18). The test

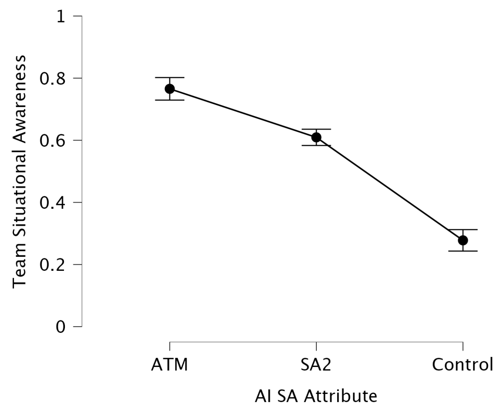


Figure 5.14: Team Situational Awareness (CAST) by AI SA Attribute. Error Bars Represent Standard Error.

Cases	Sum of Squares	df	Mean Square	F	p	η_p^2
Mission	0.186	3	0.062	2.431	0.074	0.113
Mission * AI SA Attribute	0.107	6	0.018	0.699	0.652	0.069
Mission * AI Transition Phase Ordering	0.059	3	0.020	0.765	0.518	0.039
Mission * AI SA Attribute * AI Transition Phase Ordering	0.307	6	0.051	2.010	0.079	0.175
Residuals	1.453	57	0.025			

Table 5.17: Within-Subjects Effects on Team Situational Awareness (CAST).

Cases	Sum of Squares	df	Mean Square	F	p	η_p^2
AI SA Attribute	4.148	2	2.074	30.338	$< .001$	0.762
AI Transition Phase Ordering	0.006	1	0.006	0.083	0.777	0.004
AI SA Attribute * AI Transition Phase Ordering	0.049	2	0.024	0.355	0.706	0.036
Residuals	1.299	19	0.068			

Table 5.18: Between-Subjects Effects on Team Situational Awareness (CAST).

revealed a significant main effect of AI SA attribute on teams' situational awareness ($F(2, 19) = 30.34$, $p < .001$, $\eta_p^2 = .76$; see Figure 5.14). Tukey's HSD post-hoc tests found that teams' situational awareness when working with the ATM AI teammate ($M = .77$, $SE = .05$) was significantly higher than when working with the SA2 AI teammate ($M = .61$, $SE = .05$) and the control AI teammate ($M = .28$, $SE = .04$).

Teams' situational awareness working with the SA2 AI teammate also significantly outpaced those teams' working with the control AI teammate.

Overcoming Situational Awareness Roadblock. A logistic regression analysis was conducted to examine the effect of the AI SA attribute and AI transition phase order on the likelihood that a team overcame the situational awareness roadblock by successfully photographing the upcoming target. The model summary shown in Table 5.19 suggests that the model was statistically significant ($\chi^2(113) = 14.15$, $p = .003$) between the outcome (overcoming the SA roadblock) and the predictor variables (AI SA attribute and AI transition phase order). Following the odds ratios shown in Table 5.20, working with the SA2 AI teammate made it approximately three times as likely that a team would overcome the situational awareness roadblock compared to the control AI teammate. Working with the ATM AI teammate made it approximately five times as likely that a team would overcome the situational awareness roadblock compared to the control AI teammate (see Figure 5.15). Finally, teams with the AI teammate available for the second transition phase were also approximately three times as likely to overcome the situational awareness roadblock compared to teams with the AI teammate in the first transition phase (see Figure 5.16).

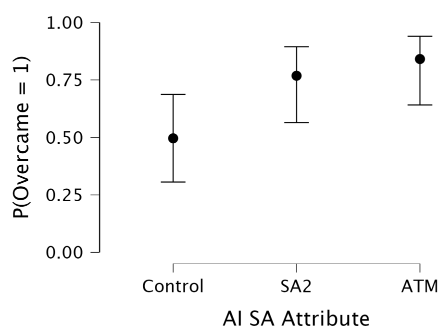


Figure 5.15: Probability that Teams Overcame the Situational Awareness Roadblock by AI SA Attribute. Error Bars Represent 95% Confidence Intervals.

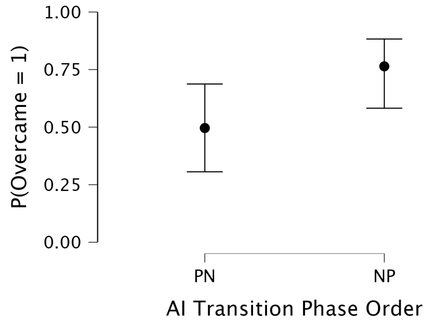


Figure 5.16: Probability that Teams Overcame the Situational Awareness Roadblock by AI Transition Phase Order. Error Bars Represent 95% Confidence Intervals.

Model	Deviance	AIC	BIC	*df*	χ^2	*p*	Nagelkerke R^2
H_0	121.396	123.396	126.158	116			
H_1	107.251	115.251	126.299	113	14.145	0.003	0.176

Table 5.19: Model Summary on Teams' Probability to Overcome the Situational Awareness Roadblock Based on Experimental Condition.

					Wald Test		
	Estimate	Standard Error	Odds Ratio	*z*	Wald Statistic	*df*	*p*
(Intercept)	−0.016	0.410	0.984	−0.038	0.001	1	0.970
AI SA Attribute (SA2)	1.214	0.571	3.367	2.127	4.524	1	0.033
AI SA Attribute (ATM)	1.681	0.641	5.371	2.622	6.875	1	0.009
AI Transition Phase Ordering (NP)	1.191	0.506	3.290	2.356	5.550	1	0.018

Table 5.20: Coefficients of Logistic Regression Model Estimating Teams' Probability to Overcome Situational Awareness Roadblock Based on Experimental Condition.

5.3.2.3 Team Process (Team Verbal Behaviors)

Pushing Team Verbal Behaviors. A 2 (Order of AI Transition Phase Participation: PN, NP) x 3 (AI Situational Awareness Attribute: ATM, SA2, Control) x 4 (Time Point: M1, M2, M3, M4) mixed RMANOVA was conducted to assess the effect of AI SA attribute and transition phase participation ordering (both between-subjects) on pushing team verbal behaviors over time (within-subjects) (see Tables 5.21 and 5.22). The analysis found a significant main effect of mission on teams' pushing verbal behaviors ($F(3,\ 75) = 46.80$, $p < .001$, $\eta_p^2 = .65$; see Figure 5.17).

Tukey's HSD post-hoc tests indicated that teams' pushing verbal behaviors in M1 (M = 29.98, SE = 2.00) was significantly lower than in M2 (M = 37.84, SE = 2.00), M3 (M = 36.12, SE = 2.00), and M4 (M = 44.16, SE = 2.00). This trend of teams increasing their pushing verbal behaviors each subsequent mission held true across all missions except for M3, which was not significantly different from M2.

Cases	Sum of Squares	df	Mean Square	F	p	η_p^2
Mission	5283.521	3	1761.174	46.803	$<$ .001	0.652
Mission * AI SA Attribute	849.426	6	141.571	3.762	0.003	0.231
Mission * AI Transition Phase Ordering	145.989	3	48.663	1.293	0.283	0.049
Mission * AI SA Attribute * AI Transition Phase Ordering	160.843	6	26.807	0.712	0.641	0.054
Residuals	2822.225	75	37.630			

Table 5.21: Within-Subjects Effects on Pushing Team Verbal Behaviors.

Cases	Sum of Squares	df	Mean Square	F	p	η_p^2
AI SA Attribute	9203.212	2	4601.606	12.027	$<$ .001	0.490
AI Transition Phase Ordering	276.174	1	276.174	0.722	0.404	0.028
AI SA Attribute * AI Transition Phase Ordering	295.171	2	147.586	0.386	0.684	0.030
Residuals	9565.075	25	382.603			

Table 5.22: Between-Subjects Effects on Pushing Team Verbal Behaviors.

The analysis also revealed a significant main effect of AI SA attribute on teams' pushing verbal behaviors ($F(2, 25) = 12.03$, $p < .001$, $\eta_p^2 = .49$; see Figure 5.17). Tukey's HSD post-hoc tests found that teams' verbal pushing behaviors when working with the ATM AI teammate (M = 52.00, SE = 3.06) were significantly greater than those teams' working with either the SA2 (M = 30.02, SE = 3.06) or control (M = 25.62, SE = 3.02) AI teammate.

These two main effects were qualified by a significant interaction effect between mission and AI SA attribute on teams' pushing verbal behaviors ($F(6, 75) = 3.76$,

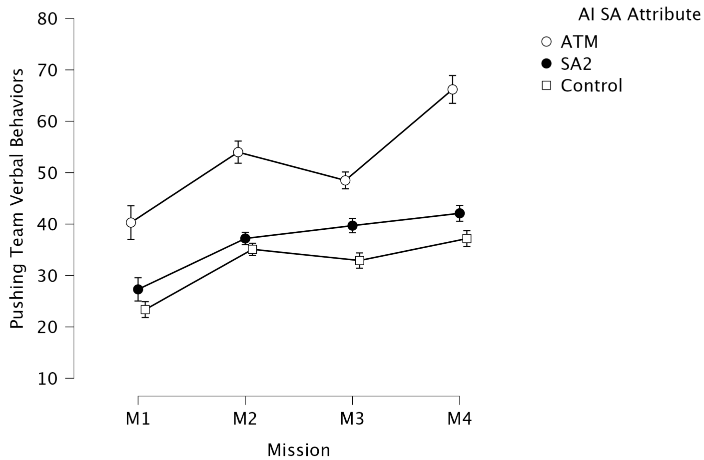


Figure 5.17: Pushing Team Verbal Behaviors by Mission and AI SA Attribute. Error Bars Represent Standard Error.

p = .003, η_p^2 = .23; see Figure 5.17). Simple main effects of the ordinal interaction effect found that there were significant differences between the AI SA attributes in all missions except for M3, where the teams' with the ATM AI teammate's pushing verbal behaviors dropped (M = 48.50, SE = 3.51), and were not significantly different from teams' with the SA2 AI teammate (M = 39.70, SE = 3.51).

Cases	Sum of Squares	df	Mean Square	F	p	η_p^2
Mission	32.070	3	10.690	1.672	0.180	0.063
Mission * AI SA Attribute	20.672	6	3.445	0.539	0.777	0.041
Mission * AI Transition Phase Ordering	1.385	3	0.462	0.072	0.975	0.003
Mission * AI SA Attribute * AI Transition Phase Ordering	23.861	6	3.977	0.622	0.712	0.047
Residuals	479.467	75	6.393			

Table 5.23: Within-Subjects Effects on Pulling Team Verbal Behaviors.

Pulling Team Verbal Behaviors. A 2 (Order of AI Transition Phase Participation: PN, NP) x 3 (AI Situational Awareness Attribute: ATM, SA2, Control) x 4 (Time Point: M1, M2, M3, M4) mixed RMANOVA was conducted to assess the

Cases	Sum of Squares	df	Mean Square	F	p	η_p^2
AI SA Attribute	158.817	2	79.408	3.544	0.044	0.221
AI Transition Phase Ordering	72.549	1	72.549	3.238	0.084	0.115
AI SA Attribute * AI Transition Phase Ordering	3.558	2	1.779	0.079	0.924	0.006
Residuals	560.200	25	22.408			

Table 5.24: Between-Subjects Effects on Pulling Team Verbal Behaviors.

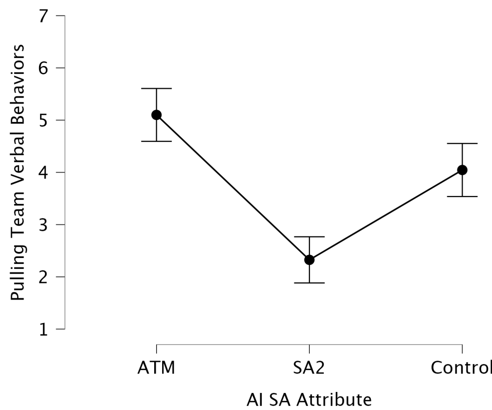


Figure 5.18: Pulling Team Verbal Behaviors by AI SA Attribute. Error Bars Represent Standard Error.

effect of AI SA attribute and transition phase participation ordering (both between-subjects) on pulling team verbal behaviors over time (within-subjects) (see Tables 5.23 and 5.24). There was a significant main effect of AI SA attribute on teams' pulling verbal behaviors ($F(2,\ 25) = 3.54$, $p = .044$, $\eta_p^2 = .22$; see Figure 5.18). Tukey's HSD post-hoc tests showed that teams' pulling verbal behaviors were significantly lower for those teaming with the SA2 AI teammate ($M = 2.31$, $SE = .74$) than those working with the ATM AI teammate ($M = 5.08$, $SE = .71$).

5.3.3 Quantitative Results Summary

Based on the quantitative results analyzed in the current section, there are several findings to summarize. Focusing first on RQ1.2, which centered on the acceptance of AI teammates' contributions to situational awareness and team cognition as a whole. There were no indications in the data suggesting the participants disliked the AI teammates or their contributions. However, a significant finding was that trust in the AI teammate was lower in M3 than in M1. This result is crucial as it represents the period when the CERTT UAS-STE implemented a new system failure, so the teams would need to continue learning and adapting. It is likely that despite all teams being told that the AI teammate has nothing to do with system failures of the CERTT UAS-STE readouts, those participants still somehow related the failure to the AI teammate. Thankfully, a marginally significant effect of the AI SA attribute showed trust in the AI teammate at its lowest point for the control AI teammate. In contrast, the ATM and SA2 AI teammates returned to baseline levels throughout the time points.

Second, the results spoke to RQ3.3, which related to how AI participation in transition phases affected situational awareness. Specifically, the results found that teams with an AI teammate participating in the second transition phase were more likely to successfully overcome the situational awareness roadblock. These teams also had significantly higher levels of average target processing efficiency. Thus, the benefits of AI participation in a transition period appear greater later in a team's life cycle. The order in which the AI teammate participated in transition periods also significantly affected participants' perceived shared mental model with their human teammate. Specifically, the teams with an AI teammate with an AI SA attribute (not control) had a better perceived shared mental model with their human teammate when the AI teammate participated in the first transition phase.

Finally, the results directly addressed RQ4.3, which asked which AI SA attribute resulted in the best support and development of situational awareness in human-AI teams. The quantitative findings overwhelmingly supported the effectiveness of the ATM attribute over the SA2 and control attributes. These findings were present in the survey metrics and the CERTT UAS-STE metrics. Specifically for the survey, the ATM AI SA attribute resulted in a better perception of a shared mental model with the AI teammate by M4. There was also a marginally significant effect for perceived situational awareness that showed participants perceived their situational awareness decreasing as time went on. This perception is most likely attributed to participants' learning more about the system through experience and system failures to more accurately rate their situational awareness. However, it is interesting to note that the SA2 attribute had the best perceived situational awareness, which could be a result of its ability to do much of the work in system failures for its teammates. The CERTT UAS-STE metrics then displayed the ATM AI SA attribute significantly improving team situational awareness. The verbal behaviors supporting team situational awareness (CAST metrics) had the ATM attribute on top, and this also resulted in those teams having the highest probability of overcoming the situational awareness roadblocks. Lastly, teams with the ATM AI teammate had significantly more pushing team verbal behaviors indicative of effective teaming than the other two AI SA attributes. This is likely due to the ATM AI teammate providing specific and directed help that encourages human teammates to understand how the system works better and directly improve their SA and SMM. However, these results can benefit from the additional context provided by the qualitative data collected in the focus group interview.

5.4 Study 3: Qualitative Results

The focus group interview data provided additional context to the quantitative results by allowing participants to provide direct commentary on the lived experiences that shaped their survey responses and team behaviors throughout the four missions. As such, this qualitative data is imperative to adequately answering the RQs posed by Study 3. Specifically, the qualitative data provides three supplemental themes that help contextualize the quantitative findings. The first theme covers why the ATM SA attribute outperformed the other two conditions across various measures. The second focuses on what made AI participation in the second transition phase more beneficial than the first. Lastly, the third theme examines how AI contributions to situational awareness were accepted and which were found to be particularly helpful. These three themes further contextualize the answers to RQ1.2, 3.3 and 4.3 provided by the quantitative results.

5.4.1 Designing AI Teammates to Support Complex Coordination

The technical advantages offered by AI present a tempting opportunity to reduce human teammates' workload. However, it appears that there is a significant downside to this practice as human teammates may be slow to develop a deeper understanding of the task and how they must cooperate to remain effective through challenging circumstances. This idea, similar to the lumberjack effect, states that simply giving human teammates all the information they need to overcome obstacles can harm their team cognition. Several participants made note of this benefit in their focus group interviews:

> *"The AI helped develop our team's common ground because it was someone who would let us know a system is down and that I need to actually communicate with them. We can't just sit there and both do our own thing." (Team 10-ATM-NP)*

> *"Yes [the AI teammate helped contribute to common ground]. I saw that it was telling my teammate that they needed to tell me the distance when my system failed, and then when their system failed, it was telling me that I needed to give them the bearing. So that helped me know what pieces of information were important." (Team 9-ATM-PN)*

As seen in the quotes above from Teams 9 and 10, the participants benefited from learning what information was essential to each role. Specifically, the ATM AI teammate's information during the system failure *"helped me know what pieces of information were important."* The ATM AI teammate also forced critical verbal behaviors supporting situational awareness by reminding teams that they *"can't just sit there and both do our own thing."* As such, the ATM AI teammate's significantly better performance in overcoming situational awareness obstacles and team situational awareness verbal behaviors is supported by the assertion that these teams benefited from engaging in complex cooperation and coordination supported by their AI teammate.

Further still, the participants in the other AI SA attribute conditions spoke on their desire to better understand how their roles overlapped:

> *"The only thing that I think would have helped more is if I knew what my teammate was doing. Because I wasn't really sure how what I was doing affected what they were doing." (Team 22-SA2-PN)*
> *"Even now, I'm still kind of unsure of what exactly my teammate knew about what I was doing and vice versa. ...if we miss a target, so be it." (Team 22-SA2-PN)*
> *"I think I was under the impression that we had the same information, which just wasn't the case. So at the beginning, I was a little bit confused*

about that aspect." (Team 31-Control-PN)

One of the most critical aspects of resilient teams is their ability to understand the needs of their teammates' taskwork, how it integrates with their taskwork, and how to make it coalesce to reach the team goal through effective teamwork behaviors. Not being able to understand *"how what I was doing affected what they were doing"* represents a significant shortcoming in a team's resiliency and ability to build team cognition capable of overcoming obstacles such as system failures. As such, *"if we miss a target, so be it"* and *"I was a little bit confused"* can quickly become a prevailing attitude. These sentiments showcase how the SA2 and control conditions failed to encourage a deeper understanding of how the AVO and PLO roles overlapped that the ATM condition *did* encourage.

The ATM SA attribute was specifically appreciated for its ability to provide additional context to system failures and support for complex coordination:

> *"Yeah, they gave initial information that was necessary but then we have to figure it out human to human, and I think for me it's a lot easier for me to connect with another person." (Team 9-ATM-PN)*
> *"I think it was just like a mediator. That's basically what their job was, and we're connecting things." (Team 6-ATM-PN)*

Teams' saw the ATM AI teammate as a *"mediator,"* and it was up to the human teammates to use the information provided by the AI teammate to *"connect things."* As such, many teams were able to use this support from the AI teammate to successfully overcome their system failure and develop a deeper understanding of the task and roles that allowed them to improve team situational awareness. Many participants also found it *"a lot easier to connect with another person"* rather than coordinating with only the AI teammate to overcome the system failure, which also likely contributed to the significantly higher levels of perceived shared mental model with the AI teammate seen in the ATM SA attribute condition.

5.4.2 The Benefits of AI Participation in Transition Phases are More Substantial Later in A Team's Life Cycle

Participation in transition phases significantly affected performance in the form of target processing efficiency and overcoming the situational awareness roadblock. This outcome appears to largely be driven by what the topic of conversation was in transition phases:

> *"During the first transition period, I'm still getting the hang of things and I didn't fully understand the task yet. So I don't know necessarily what to ask, but then the second transition period, I feel like it went really well. We got through what we needed and talked about how to fix the system crashes and how to help each other." (Team 6-ATM-PN)*

Participants conveyed that the first transition phase was frequently used to understand fundamental aspects of the task as they were *"still getting the hang of things and didn't fully understand the task."* Because of this novelty, teams didn't *"know necessarily what to ask,"* but by the time of the second transition phase, the topics of conversation were significantly more complex and focused on *"how to fix the system crashes and how to help each other."* As such, the second transition period was frequently associated with higher-level coordination and a direct focus on improving resiliency through team situational awareness and shared knowledge.

Because of the nature of the discussions during the second transition phase, the participation of the AI teammate in the second transition allowed the team to ask questions related directly to improving coordination and performance:

> *"The second one because we got to a point where we were more proficient at the task, and because of that, it meant we actually knew what questions to ask. Our questions could be more specific, they could be more tailored to achieving a higher score as opposed to just how do we use this to begin with." (Team 35-Control-NP)*
>
> *"I would prefer it in the second one. There were a few things that I actually would have wanted to ask, which in the first one, I felt like I hadn't even gotten to that point where I knew what questions to ask." (Team 22-SA2-PN)*

Participants saw the second transition period as a point to refine their team's abilities and shared understanding once they had achieved a solid understanding of the basics of the task. Specifically, teams *"knew what questions to ask. We could be more tailored to achieving a higher score."* This sentiment was present across all levels of the two manipulations, as Team 22 stated they would have *"prefer[ed] it [the AI teammate] in the second [transition phase]"* for the same reason the teams' with the AI participating in the second transition phase stated they appreciated having it participate when it did.

The benefits of having the AI teammate participate in transition phases later in the life cycle of the team being clear, the first transition period also held its own unique benefits to establishing accurate shared understanding:

> *"Yeah, I feel like they were the expert in their role and helped us work together at the beginning." (Team 20-SA2-PN)*
>
> *"I think they [the transition phases] were good, especially the first one because I was a little bit worried about making mistakes...but having that reassurance that we were both at that same position of getting used to*

things gave me more confidence going into the later rounds." (Team 22-SA2-PN)

By having the AI teammate available in the first transition phase, these teams were able to establish accurate shared mental models on the basic task. Asking these questions of the AI teammate was especially helpful because *"they were the expert in their role,"* which allowed the teams to leverage that information and improve their ability to *"work together at the beginning"* of the task. Having the AI teammate in the first transition phase also gave *"reassurance that we [the two human teammates] were both at that same position of getting used to things."* Having that reassurance from the first transition phase then had a knock-on effect that gave them *"more confidence going into the later rounds."* Additionally, having the AI teammate available in the first transition phase may induce higher trust in the AI teammate, though this assertion was supported in the quantitative results.

5.4.3 Gradually Sharing More Information with Teammates Improves Information Acceptance

A central question of the current dissertation focuses on human teammates accepting the help offered by AI teammates designed to support the critical constructs of team cognition, such as situational awareness. If the information shared by AI teammates is not utilized, team performance will suffer, and acceptance of AI teammates in general will be negatively impacted for those individuals moving forward. Study 3 continues providing an answer to this question as it demonstrates that utility alone is not the only piece of the puzzle and that the way the AI teammate conveys its utility is similarly critical:

"When I was putting the bearings in for the next target, I knew that I had

to decrease the speed because it reminded me. So that helped a lot because I would not have remembered it." (Team 11-ATM-PN)

"[Improved SA] definitely at the beginning. I think it did by sending the little reminders, especially when the system failed." (Team 23-SA2-PN)

A common theme for teams was that trust and acceptance in the AI teammate was built early on through reminders, which can be designed to convey the utility of an AI teammate unobtrusively and effectively. Team 11 stated that the AI teammate contributed to their situational awareness when the AI teammate *"reminded me [to decrease the speed]. That helped a lot because I would not have remembered it."* The same sentiment was true for Team 23, which was in a different AI SA attribute altogether but conveyed its utility during a challenging time by *"sending little reminders, especially when the system failed."* This statement also emphasizes the need to have the AI teammate step up in some way when the team faces significant roadblocks. Stepping up during roadblocks is especially important given the inherent automation bias in many individuals, who expect automated systems to be competent and perfect to a fault [109].

AI teammates also convey their utility by answering questions that their human teammates have for them while they are learning the taskwork and teamwork required to accomplish the team's shared goals:

"I felt less embarrassed to ask questions to the AI because it's a computer." (Team 6-ATM-PN)

"When I was telling my teammate stuff I felt so annoying, but with the AI I just didn't have to worry about that because I want to be nice, but the AI doesn't care." (Team 11-ATM-PN)

"...It's like you said, you can't be judged because it's a computer; you'll

ask more questions." (Team 6-ATM-PN)

The inherent advantages of working with AI systems do not stop with the technical benefits but extend to social advantages. As all of Team 6 states *"I felt less embarrassed to ask questions to the AI because it's a computer"* distinctly because *"you can't be judged because it's a computer."* This sentiment was also seen in Study 2B, where AI teammates do not discriminate overtly based on protected characteristics (though poor training methods can alter this in some contexts). Because of this advantage, the teams in Study 3 felt comfortable *"ask[ing] more questions"* than they would typically have been. As such, their performance would be improved along with their team cognition, given the AI teammate's accurate information about the task.

Despite these advantages offered by the AI teammate from a technical and social standpoint, the AI teammate and their contributions to situational awareness were not immediately accepted without scrutiny:

> *"I felt like I would trust a human more than an AI, but after working with it my trust in it was the same." (Team 21-SA2-NP)*

There was a sentiment from teams that their initial trust in the AI teammate was at a disadvantage. As Team 21 stated *"I felt like I would trust a human more than an AI,"* and the quantitative data also showcased a similar sentiment as trust fell slightly at the beginning of the teams' life cycle. However, just as the quantitative data suggested a marginally significant trend, the qualitative data supports this trend as *"after working with it [the AI teammate] my trust in it was the same [as their trust in the human teammate]."* This sentiment Shows that the acceptance of contributions by the AI teammate can overcome initial inequity through effective design and meaningful contributions such as reminders.

5.4.4 Qualitative Results Summary

The qualitative results of Study 3 provide excellent supplemental insights to the quantitative findings, specifically by diving deeper into the lived experiences behind the most prominent findings of the quantitative analysis. These experiences focused on three themes: 1) participants provided valuable reasoning behind the higher performance of the ATM SA attribute; 2) the benefits of the two transition phases and the AI teammate's role in them were discussed in detail; and 3) the nature of how human teammates accept the contributions to situational awareness provided by AI teammates. The first theme was borne out by the fact that the ATM SA attribute directly supported the human teammates' complex cooperation. This support led the human teammates to understand the task and their roles within it better. The second theme found that the benefits of an AI teammate participating in transition phases early or late in a team's life cycle can vary widely. While the first transition phase provides an excellent opportunity to build accurate mental models and learn the basics of the task, the second transition phase discusses complex coordination, leading to higher scores. Lastly, the final theme showcased how the design of AI teammates can contribute to how their human counterparts accept the information they provide. For example, the theme discussed the utility of reminders from the AI teammate to foster the rapid growth of trust in the AI teammate. Taken altogether, these three themes provide critical insight into the thought processes behind the quantitative findings of those same participants, allowing the results to be all the more impactful to human-AI teaming research.

5.5 Study 3: Discussion

Study 3 is the culmination of the current dissertation that puts several AI design features together to support situational awareness in practice. These AI SA attributes were tested using an experimental platform designed for complex team interactions over an extended period through several action stages or missions. This study also introduced transition phases for the first time in human-AI teaming research, which was especially critical given their importance to goal setting, strategy formation, and team improvement, which are all essential aspects of effective team cognition [195]. There were three study-specific RQs addressed by Study 3: 1) RQ1.2 asked if human teammates accept contributions to situational awareness from AI teammates in a complex hands-on task; 2) RQ3.3 asked if AI participation in a transition phase was more beneficial earlier or later in the teams' life cycle; and 3) RQ4.3 asked which AI SA attribute best supported the development and sustainment of team situational awareness in human-AI teams best. For RQ1.2, Study 3 found that participants were still hesitant to trust the AI teammate but that the utility of the teammate won out over time, and the qualitative data found that simple reminders provided early on in the task were especially beneficial. The current study answered R3.3 by finding that while both transition phases were helpful, participation in the first transition phase benefits essential task learning. However, the second transition phase discussions typically focus on and benefit higher-level skills with a more tangible impact on performance and resiliency. Lastly, the evidence answering RQ4.3 found that participants' team situational awareness benefited most from the ATM SA attribute, as this AI teammate supported more complex team interactions and built more natural resiliency. This advantage stemmed from its contribution to overcoming the situational awareness roadblock, which required the team to identify and share the relevant information themselves. The following discussion critically examines these findings in light of existing research. It provides interpretations and highlights future research needs to build upon them and further enhance the understanding of team cognition in human-AI teams.

5.5.1 Supporting Team Cognition in Human-AI Teams Means Supporting Complex Coordination Among Human Teammates

A prominent result of the current study was the performance of the ATM SA attribute over the control and even the SA2 attribute, primarily driven by its ability to enhance human teammates' understanding of their two roles. This result is somewhat surprising given the design of the SA2 attribute, which, during situational awareness roadblocks, alerted the team to the failure and provided the information needed to mitigate the failure directly to the team. Based on Study 3's results, it is clear that making meaningful contributions to situational awareness involves more than simply giving teammates the solution. However, there is precedent for such a finding in the automation literature, as the lumberjack effect is somewhat related to this assertion. 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 [237, 289, 242]. While Study 3 does not examine autonomy failures, it does involve a system's failure, which the AI teammate is attempting to help mitigate. When trying to overcome the situational awareness 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 SA attribute. The ATM SA attribute pointed out the information the teammate experiencing a failure needed to continue operating effectively, which improved the human teammates' understanding of their roles and tasks. This information then enhanced the teams' resiliency to future system failures naturally by teaching teammates what information was essential to which teammate, when they needed that information, and how they needed that information, resulting in teams capable of well-coordinated actions stemming from effective decision-making processes [335]. Alternatively, the SA2 attribute essentially completed that entire process for the teammates. As such, it is clear that improving team situational awareness and team cognition broadly should focus on supporting more complex cooperation among human teammates. This assertion builds upon existing theoretical frameworks in human-AI teamwork that suggest AI teammates can play a role in learning, training, and human-AI teaming implementation [75, 277]. Furthermore, the existing research on processes in the form of team verbal behaviors (e.g., information pushing and pulling) is given new context. The ATM and SA2 SA attributes resulted in significantly more verbal behaviors supporting team situational awareness (CAST) and pushing verbal behaviors indicative of effective teaming. Past research on human-AI teams working in the CERTT UAS-STE found that those working with AI teammates struggled to engage in effective pushing verbal behaviors [212, 68]. While the current study is not a one-to-one comparison, the positive performance of these two AI SA attributes would suggest that this shortcoming can be overcome with more effective AI teammate designs.

Accepting situational awareness information from the AI teammate played a role in whether or not teams successfully overcame the situational awareness roadblock. An important finding of Study 3 was the participants' varying attitudes toward the AI teammate and the information it provided. Throughout the experiment, the AI teammate always performed their role flawlessly, and the two AI SA attributes provided even more information to benefit the team's situational awareness. However, the ATM SA attribute resulted in a significantly better perceived shared mental model with the AI teammate than the control condition, and the ATM teams were more likely to overcome the situational awareness roadblocks. This result could be attributed to the lack of explainability in the SA2 condition [216], which was noted in the qualitative results as a feature that improved participants' receptivity to the AI teammate, but there is also the possibility that the system failures affected trust in the AI teammate. Prior research on system-wide trust shows that a failure in one automated system can influence trust in a separate but related system [184]. The marginally significant results on trust in the AI teammate provide some evidence for this assertion as trust was reduced after M3 (this finding is despite being explicitly told the system failure is part of the CERTT UAS-STE simulation), which is when a different type of system failure was implemented. As such, even though AI teammates are distinct entities from the system users interact with to accomplish their team goal, researchers and developers should account for this possibility and have the AI teammate provide as much support as possible to mitigate the failure [185]. Regardless of the participant's trust in the AI teammate, the ATM SA attribute was shown to satisfy a desire for planning ahead, which had a significant perceived and objective benefit to the teams' situational awareness and shared understanding. The ability to send messages providing information about future ROZs improved these teams' perceived shared mental models and conveyed critical utility and an ability to plan and anticipate future events. This ability to anticipate future events is vital to ideal human-AI teams, as shown in all components of Study 2 and related research [342]. These findings showcase how AI teammates can best benefit situational awareness by providing information that supports more complex human cooperation and communication that showcases an ability to anticipate and help teammates plan for future events. These actions support effective team situational awareness, perceived shared mental models, and improve the acceptance of AI teammates and the information they bring to the table.

5.5.2 Transition Phases are Critical to Human-AI Teams but the Benefits of AI Participation Vary

The AI teammates' participation in the transition periods significantly affected how these teams engaged with one another and how the individual teammates perceived one another. Transition periods are critical to effective teaming in all forms, human-only and human-AI [195, 160]. The current study examined how the presence of an AI teammate early or late in the team's life cycle played a significant role in the team's ability to develop situational awareness, effective team processes, and performance. The results overwhelmingly supported the importance of AI participation in transition phases later in the team's life cycle. This echos research pointing towards the different types of discussions occurring earlier and later in teams' life cycles [106]. In the case of Study 3, it was clear that both transition phases were beneficial for the teams; however, the two phases were helpful for different reasons. The qualitative data points to the benefits of the first transition phase, providing critical support to human teammates' basic understanding of the task. In contrast, the second transition phase offered tangible benefits to their ability to engage in more complex team behaviors directly related to team performance. This finding has precedent in human-only teaming as transition phases change their purpose based on when in the life cycle they take place and how similar or dissimilar the following action phase task may be [24]. As such, the benefits of AI teammate participation early in the life cycle are real but do not have the same impact on team processes and performance as participation in later transition phases. This assertion is borne out in all aspects of the results as the qualitative data shows participants knew what questions to ask in the second transition phase but also appreciated the ability to ask basic questions in the first phase. The quantitative data then points to the enhanced team performance of those AI teammates participating in the second transition phase. As such, human-AI teams mirror human-only teams' dependence on team cognition to have highly effective transition phases [24], and these transition phases are a direct avenue for teams' emergent states to act as inputs on future team effectiveness. As the role of human-AI teams increases, the importance of these transition periods will increase, and as such, the criticality of team cognition will be further emphasized.

That being said, the topic of transition phases in human-AI teaming is essentially unanswered in any capacity, despite its importance to effective teaming. The lack of research on transition periods in human-AI teaming is somewhat surprising given the potential opportunities to implement additional aspects of AI explainability and transparency [216, 18, 85]. The need to conduct this type of research is bolstered firmly by the current study that found AI participation in these transition periods significantly affects team performance and the perceptions human teammates have towards their fellow human and AI teammates. Study 3 found that participants' perception of a shared mental model with their human teammate was better when the AI teammate participated in the first transition phase for all AI SA attributes except for control. This finding displays that AI teammates can improve the shared mental models between human team members when they actively participate and provide SA information during transition phase discussions. Alternatively, the teams with the AI teammate participating in the second transition phase were more likely to overcome the situational awareness roadblocks, showcasing enhanced team situational awareness. These findings are a first for team cognition research in human-AI teams [238], representing a primary motivating goal of the current dissertation. As such, these findings must be taken as a motivating call for researchers to begin integrating transition phases in their human-AI teaming research, given these direct benefits to team cognition and the potential for improving AI transparency features [216].

5.6 Study 3: Design Recommendations

The results of Study 3 can be reviewed to continue the goal of providing practical design recommendations that practitioners, developers, and researchers can utilize to tangibly improve human-AI teams. The following design recommendations focus on the most impactful results of Study 3, which include the effectiveness of ATM SA attribute information and participation in transition phases.

5.6.1 AI Teammates Must Be Proactive in Transition Phase Discussions

Often, there is a period in transition phases or any discussion that requires breaking the ice to jump-start talks. This jump-start for communication was an aspect found and suggested in Study 1 of the current dissertation [278]. The current design recommendation suggests that the AI teammate jump-starts communication in the transition phase but should also not sit back throughout the discussion, waiting to be leveraged. Study 3 showed that many participants did not even know what questions they should ask in the first transition phase, meaning the utility of the AI teammate could be further enhanced by designing it to not passively wait to be asked a question. AI teammates should monitor communication between human teammates for inaccurate statements and misunderstandings. Additionally, AI teammates should be capable of bringing in recommendations from past performances to provide critiques to strategy, goal assignment, and goal specification [195]. This design recommendation could result in teams that can reap multiple benefits from transition phases at any stage of the teams' life cycle. For example, the benefits of asking simple questions to the AI teammate early in the life cycle and the benefits of more complex suggestions made by the AI teammate that would typically be made in the later stages of the teams' life cycle. With that being said, the ability to have AI teammates participate in transition phase discussions at all would represent a massive improvement over other implementations of AI teammates. Further, while the current study displayed more tangible benefits of participation in the later stages of the teams' life cycle, it would be best practice to have AI teammates participate in all transition phase discussions if possible, especially as there were no adverse effects to participation measured in the current study.

5.6.2 Design AI Teammates to Augment Team Memory First

Study 3 provided overwhelming support for the benefits of developing AI teammates that support situational awareness by augmenting team memory. AI teammates should be designed to support two aspects of memory augmenting, with the first being the support of more complex cooperation between human teammates. As the ATM SA attribute's response was designed to help teammates understand what information was most important to whom and when necessary, it allowed the team to develop shared knowledge and situational awareness naturally. AI teammates should be designed to identify this information and convey it to human teammates, whether that be during a roadblock such as a system failure, during transition phase discussions, or just as a reminder during action phases (as long as it does not prove to be a distraction to the team). Second, AI teammates should be designed to support the team's future actions by providing information relevant to planning. This information showcases the AI teammate's ability to anticipate future events, which is a necessary feature of AI teammates, as stated in recent research [342]. Augmenting team memory also provides transparency by conveying an understanding of future tasks relevant to the team's shared goal. It instills confidence and trust in the AI teammate if they convey that understanding and helps human teammates reach the same understanding rapidly if they lag behind. For example, one easy way to build that understanding and trust was used in the current study, where the AI teammate provided information about the speed needed for an upcoming restricted zone and then reminded the team of that change when it became relevant. As such, augmenting team memory represents an accessible and easily understood method of improving situational awareness in human-AI teams, along with the host of other benefits noted in the current study, such as improving perceived shared mental models with all teammate types.

5.7 Study 3: Limitations and Future Work

The results of Study 3 have significant benefits to the state of human-AI teaming, though it is not without limitations. Specifically, the findings on team processes include the messages from the AI teammate, which varied based on the AI SA attribute condition. Though this is a similar practice to past CERTT UAS-STE experiments [212, 68, 211], future studies should examine how AI teammates information-sharing for situational awareness can increase situational awareness of verbal behaviors by human teammates. Another limitation of the study is the sample, which involved college students with varying levels of familiarity with AI and video game interfaces. As such, several questions remain on the impact of individual differences on human teammates' willingness to accept information from their AI teammates. Finally, the current study did not collect shared mental model structure, only content, meaning Study 3 cannot make claims on shared mental model similarity levels such as those made in Study 1. Future research studies should collect shared mental model structure and content alongside the processes making up interactive team cognition measures so claims can be made comparing the two perspectives of team cognition.

5.8 Study 3: Conclusion

Study 3 contributes significantly to the current understanding of human-AI teams and how aspects of team cognition, such as team situational awareness, can develop with the influence of AI teammates. The present study accomplished this by examining how AI participation in transition phases and AI SA attributes affected several measures of team effectiveness, situational awareness, and shared knowledge. The ATM AI teammate was found to have the most significant benefit to perceived and objective team cognition and situational awareness measures. Specifically, the perceived shared mental model participants' had with the AI teammate, team situational awareness verbal behaviors (CAST), and the teams' likelihood of overcoming the situational awareness roadblock. The high performance of the ATM SA attribute is likely a result of its ability to improve human teammates understanding of the task, their roles within it, and how their roles overlap. This sentiment was reflected in the qualitative findings, where participants discussed a desire to learn what information was essential and how the ATM SA attribute helped them learn those aspects of the task. The AI teammate's participation in transition phase discussions also played a significant role in the team's perception of their teammates and their performance. If the AI teammate participated in the first transition phase and was not the control AI SA attribute, then participants' perceived shared mental model with their human teammate improved compared to those AI participating in the second transition phase. However, those teams with the AI teammate participating in the second transition phase were more likely to overcome the situational awareness roadblock and had greater target processing efficiency. These results indicate that participation in transition phase discussions have different benefits depending on when they occur in a team's life cycle, with early discussions focusing on a basic understanding of the task and later discussions concentrating on more complex cooperation strategies that influence team performance. Taken altogether, the findings of Study 3 provide a comprehensive and deep understanding of team cognition in human-AI teams, with specific attention focused on what human teammates want from their AI teammates to support team cognition. The study also shows that AI teammates *can* be designed to support better team cognition in the form of team situational awareness, but that not all designs to this end are equal as augmenting team memory was found to benefit the teams' situational awareness more by giving human teammates a more complete understanding of the team and their shared goal. With these findings, researchers and practitioners can benefit in designing their AI teammates to extend what is possible for team situational awareness within human-AI teams, making them more resilient and effective than ever.

Chapter 6

Discussion & Conclusion

The current dissertation has produced a series of contributions that converge on two main research fields: team cognition and human-AI teaming. In doing so, it addresses the many dissertation-wide and study-specific research questions the current dissertation poses. It provides tangible benefits to application and research through design recommendations and findings. The present chapter reviews these contributions and is organized into several overarching sections that begin with: 1) revisiting the dissertation-wide research questions; 2) a review of the dissertation's contributions to team cognition and human-AI teaming before ending with future work; and 3) closing remarks.

6.1 Revisiting the Dissertation Wide Research Questions

The current dissertation proposed four dissertation-wide research questions broadly addressing AI teammates' effect on team cognition. Each of these D-RQs can be addressed in part by one or more of the three studies of the dissertation individually, but to address these questions in full requires all three studies to be examined together. The current section discusses how these studies holistically answer the D-RQs by revisiting each and discussing how the relevant studies address them.

6.1.1 D-RQ1: How Do AI Teammates Influence the Development and Sustainment of Team Cognition in Teams?

Two of the three studies presented in the current dissertation addressed D-RQ1, which asked how the inclusion of AI teammates influences team cognition in teams. Specifically, Study 1 and Study 3 provided insight into this question (see Figure 6.1). Study 1 utilized qualitative and quantitative data to understand how including one or two AI teammates affected human teammates' shared mental models, trust, and perceived team cognition. Study 3 examined how various AI SA attributes designed to share situational awareness information and the participation of those AI teammates in transition phases affected human teammates' team cognition.

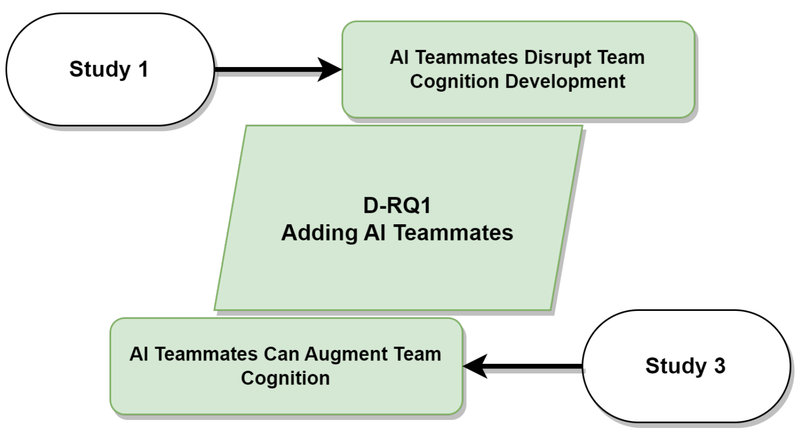


Figure 6.1: Study Findings Relationship to D-RQ1

Study 1 found that participants tend to have an adverse reaction to working with an AI teammate and that adding two AI teammates exacerbated this reaction. The participants working with two AI teammates had lower perceived team cognition and trust levels than those with only one AI teammate. These perceptions were all despite the AI teammate performing at a high level and those teams having the best performance of the three compositions tested. The qualitative data also spoke to D-RQ1 as the teams highlighted how important action-related communication was from the AI teammate and the importance of explicit shared goals to develop more effective team cognition.

In Study 3, the ability of AI teammates to significantly impact the development and sustainment of team cognition was highlighted. The AI SA attribute focusing on augmenting team memory profoundly benefited teams' perceived shared mental models with human and AI teammates, their level of information pushing team verbal behaviors, team situational awareness, and ability to overcome situational awareness roadblocks. This study was the first to show that AI teammates could be designed to impact how team cognition develops for these teams. While the other AI SA attribute studied (awareness of information changes inside and outside the team) also had some significant benefits to team situational awareness, the augmenting team memory SA attribute was consistently above it and the control condition.

Synthesizing these results together to address D-RQ1 fully, adding AI teammates to teams can create challenges to team cognition but do not need to remain a challenge. As the current dissertation has discussed, adding AI teammates to teams challenges existing team processes such as goal monitoring, back-up behaviors, and system monitoring [68]. These disruptions are brought about by the unique nature of teaming with an artificial entity versus a human teammate, as they come with very different skill sets and expectations [342, 327]. These assertions were borne out in Study 1; however, Study 3 showcased that while these challenges still exist when teaming with AI, they do not have to be a constant. Study 3 showcased how teams working with AI teammates specifically designed to support aspects of team situational awareness do benefit from that information to go on and develop more effective levels of team cognition, both perceptually and objectively.

6.1.2 D-RQ2: How Do Attitudes Towards An AI Teammate Affect Team Cognition Development?

To examine how human teammates' attitudes towards their AI teammate affect team cognition development, Study 1 and Study 2 were designed to complement one another (see Figure 6.2). Specifically, Study 1 examines how attitudes toward AI teammates affect team cognition development by manipulating the presence of one or two AI teammates. Alternatively, Study 2 examined how various information-sharing attributes meant to supplement team cognition influenced human teammates' perception of the AI and their subsequent perceptions of team cognition.

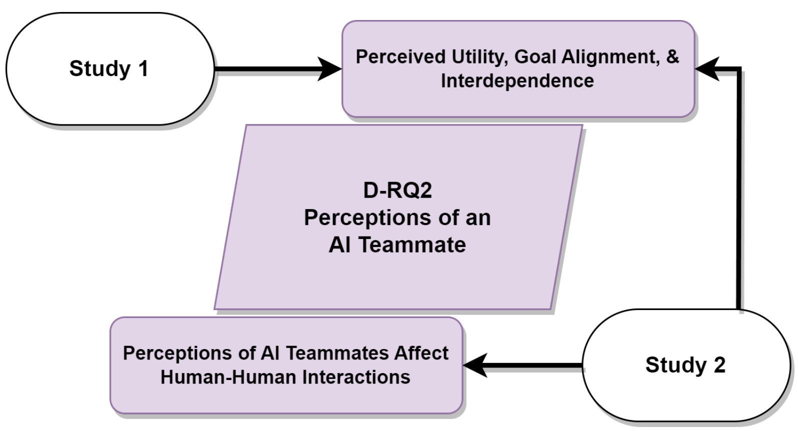


Figure 6.2: Study Findings Relationship to D-RQ2

Throughout Study 1's quantitative and qualitative data, there was overwhelming support for the notion that humans' attitudes towards their AI teammate(s) significantly influence team cognition development. While the participants' shared mental model similarity was not affected, their perceived levels of team cognition did take a significant hit. These adverse effects on their team cognition were coupled with the fact that teams with two AI teammates had significantly worse trust than teams with only one AI teammate. This loss in trust was coupled with the fact that those teams with both teammate types had significantly worse perceived team cognition with the AI teammate compared to their human teammate despite the AI teammate's high performance. These findings make the connection between attitudes and team cognition clear. However, the qualitative data found that attitudes toward AI teammates were positively influenced when the AI teammate conveyed their utility through action-related communication, shared goals, and interdependent actions.

Study 2 utilized SEM, which enabled the study to examine the effect of attitudes on team cognition development directionally. The study found that when AI teammates provided utility to their teammates in the form of information-sharing, their attitudes towards those AI teammates changed significantly. Participants' trust, perceived shared mental model, teammate efficacy, and other related measures increased significantly compared to a control. Furthermore, these attitudinal measures were found to significantly mediate the relationship between perceived team cognition measures and the actual type of situational awareness information-sharing. Interestingly, the AI teammate's information-sharing (as a corrective back-up behavior and AI explainability) also negatively impacted the participants' perceived shared mental model with their human teammate. As such, this SEM made the first direct connection between human teammates' affective attitudes towards their AI teammates and their potential to develop team cognition with their human teammates.

Together, these findings showcase how important it is to introduce and design AI teammates to impact their human teammates positively. Study 1 highlights how difficult it can be to properly introduce AI teammates and have their human counterparts accept them [276, 278, 224]. Thankfully, Study 2 provided positive results for the effect of AI teammates designed to support team cognition through information-sharing, as the AI teammates were overwhelmingly accepted as effective teammates. However, Study 2 also convincingly emphasized the importance of attitudes towards an AI teammate and their effect on team cognition between all teammates. These attitudes could be positively affected by increasing the perceived utility of the AI teammate through more interdependent interactions, action-related communication, discussing shared goals, and increasing the utility of the AI teammate through information-sharing [2].

6.1.3 D-RQ3: What Effect Do Specific AI Information-Sharing Attributes Have on Team Cognition?

D-RQ3 asked what effect various AI information-sharing attributes have on team cognition within human-AI teams, which was addressed by the research conducted in Studies 2 and 3 (see Figure 6.3). Both of the studies utilized mixed methods approaches. However, Study 2's qualitative component was much more extensive than Study 3's, allowing it to act as a standalone study. Study 2 also utilized a factorial survey to gauge participants' perceptions of team cognition and their AI teammates' information-sharing attributes. In contrast, Study 3 used the findings of Study 2 to select two AI SA information-sharing attributes to examine their effect on team cognition in a hands-on experiment.

Study 2 revealed several significant findings highlighting the potential for AI teammates to contribute meaningfully to team cognition in human-AI teams. From the quantitative research conducted as part of Study 2, there was clear evidence for the benefit of AI information-sharing on participants' perceptions of team cognition with their human and AI teammates. Specifically, the augmenting team memory attribute, explainability, and information sharing on changes inside and outside the

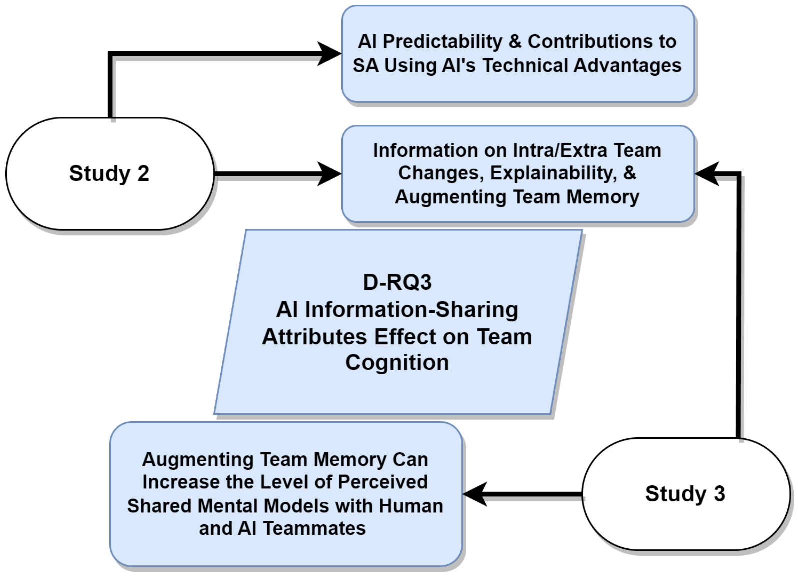


Figure 6.3: Study Findings Relationship to D-RQ3

team. The information-sharing attributes directly related to team cognition had the most significant impact, and the effect of the information-sharing attributes on participants' perception of the AI teammates' contribution to situational awareness was not fully mediated by attitudes towards the AI. The qualitative data reinforces these results as the participants discussed how important the AI teammate was to supporting effective individual and team situational awareness. As such, Study 2 identified situational awareness as the most crucial aspect of team cognition for humans teaming up with AI.

Using the information gathered by Study 2, Study 3, as the final study of the dissertation, directly examined the effect of specific AI information-sharing attributes designed to enhance situational awareness. Study 3 found several significant insights as the augmenting team memory SA attribute outperformed a control and the information-sharing of changes inside and outside the team SA attributes. The participants' perceived shared mental models with their human and AI teammates were improved over the control for the augmenting team memory SA attribute when the AI teammate participated in the first transition phase. The augmenting team memory SA attribute also markedly improved team situational awareness. These teams also excelled in overcoming the situational awareness roadblock, and their performance benefited from the AI teammate's participation in the later transition phase discussion.

Taking these results together to develop a response to D-RQ3, it is clear that AI information-sharing attributes can significantly impact team cognition. Study 2 identified team situational awareness as the most critical component of team cognition in human-AI teams and also highlighted that information-sharing attributes directly relating to situational awareness had the greatest effect on perceptions of team cognition. Study 3 took these results a step further by implementing them within an extended research study and comparing them to a control condition. Specifically, the augmenting team memory SA attribute was found to have significant positive effects on team cognition objectively and perceptually. These effects could also be strengthened with the AI teammate's participation in the team's transition phases, which relate heavily to perceptions and action phase performance [195]. These studies provide strong evidence that information-sharing is a prime candidate for AI teammates to make meaningful contributions to developing and sustaining team cognition within human-AI teams. Specifically, the information-sharing attributes that directly support situational awareness at the individual and team levels, such as augmenting team memory, which related literature has suggested [85].

6.1.4 D-RQ4: How Can An AI Teammate Be Designed to Contribute, Support, and Encourage Team Cognition Development and Sustainment in Teams?

The final D-RQ addressed by the current dissertation focused on how an AI teammate can be designed to contribute, support, and encourage team cognition in human-AI teams, which Studies 2 and 3 concentrate upon (see Figure 6.4. Study 2 tackled this D-RQ by examining how various information-sharing attributes affected perceptions of team cognition. The second study also included a qualitative component investigating what attributes humans wanted their AI teammates to have to support team cognition. Lastly, Study 3 directly examined the effect of AI SA attributes on team cognition development and sustainment. These AI SA attributes for Study 3 were selected based on the results of the quantitative and qualitative components of Study 2. This saw the augmenting team memory and sharing information changes within and outside the team chosen as the AI SA attributes used in Study 3.

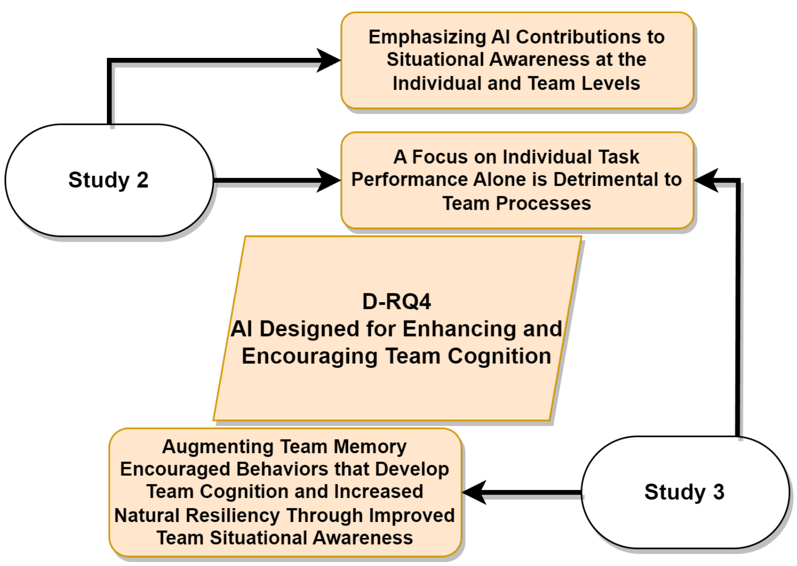


Figure 6.4: Study Findings Relationship to D-RQ4

Starting with the quantitative component of Study 2, which found a set of clear winners when it came to enhancing participants' perception of team cognition. Specifically, the augmenting team memory and sharing information changes within and outside the team. These two AI SA attributes each significantly improved perceptions of team cognition within the team (though many of these effects were mediated by the attitude-related variables). The qualitative component of Study 2 provides overwhelming evidence for the assertion that AI teammates should emphasize predictability in their actions. Participants also wanted AI teammates to convey their utility by leveraging their inherent technical advantages as part of their information-sharing attributes. These results informed the final study of the dissertation and also stand on their own to point out that human teammates will only accept an AI teammate if the AI can justify its membership on the team, which requires AI to be carefully designed to accommodate these expectations.

The third and final study of the current dissertation directly examined how these AI SA attributes influenced team cognition development in human-AI teams. Based on the results of Study 3, it was clear that the augmenting team memory AI SA attribute outperformed the information-sharing of changes inside and outside the team and the control conditions. The augmenting team memory SA attribute surpassed the other two attributes because it was the only attribute to *encourage* team cognition development. This attribute was also the only one to support ongoing team verbal behaviors that help develop team cognition. Specifically, the augmenting team memory SA attribute informed human team members what was important in a system failure, when it was important, and who needed that information. These actions helped teams develop a deeper understanding of the task and the verbal behaviors necessary for supporting effective team cognition. These assertions were also supported by the qualitative data supplementing the quantitative data, where participants discussed wanting a deeper understanding of the task to improve their teamwork.

Looking at the results of Studies 2 and 3, these findings can be considered together to provide an answer to D-RQ4. Specifically, any increase in the utility of an AI teammate will be perceived positively over a control teammate (that does only its task and fundamental aspects of teamwork). This perceived utility is also essential to improving AI teammate acceptance [342, 278]. These studies also reiterated the importance of situational awareness at the individual and team levels for human-AI teams [85]. As such, any additional utility added to an AI teammate is well suited to contributing towards situational awareness within the team, but AI's technical advantages should enable these contributions. AI teammates truly *can* encourage better team cognition by teaching human teammates what information is important, when it is important, and who needs that information. Augmenting team memory was found to engage in this task most effectively, as teams working with this AI SA attribute were found to engage in better team situational awareness verbal behaviors, develop more natural resiliency to system failures, and improve their information pushing team verbal behaviors.

6.2 Contributions of the Dissertation

The following section overviews the contributions of the current dissertation concerning the two main topics covered throughout the dissertation. Specifically, the contributions of each study and the dissertation as a whole to the issues of team cognition and human-AI teaming.

6.2.1 Contributions to Team Cognition

The intersection of team cognition and technology was a central theme of the dissertation, which showed in the dissertation-wide research questions and the study-specific research questions. Mainly how adding AI teammates influenced the development and sustainment of team cognition, how adding more AI teammates affected team cognition, how team cognition may be different conceptually for these teams, and how they relate to team outcomes. Each of these questions was part of the dissertation at some level, and each of these questions had not been satisfactorily addressed in the literature to this point [227]. Thus, the current body of work contributes significantly to the practical understanding of team cognition and future research.

6.2.1.1 Contributions from Study 1

Study 1 contributed to team cognition research in three ways, with the first advancing an emerging research agenda on collaborative activities occurring within human-AI teams and their impacts on humans and agents. Specifically, this work provides the first empirical analysis and comparison of shared mental models in traditional human-only teams and human-AI teams while also investigating the effects of team composition on team cognition and its outcomes in human-AI teams. Second, this study also expands existing research on team cognition by shedding light on the nature of the construct in modern human-AI teams and how to improve human experiences of team cognition in such teams [238]. The study also enhances the field's understanding of developing team cognition in disadvantaged communication environments, a notoriously tricky arena to develop team cognition in the literature [153]. In doing so, Study 1 provided an in-depth examination of how the inclusion of a single AI teammate influenced team cognition, attitudes, and the perception of team cognition. The study also highlighted how the inclusion of multiple AI teammates had a massive adverse effect on trust attitudes and team cognition perceptions. Such results drive improved theory and practice regarding team cognition in human-AI teams and highlighted **the importance of shared goals and information sharing attributes to team cognition development**, which became a significant focus of Study 2.

6.2.1.2 Contributions from Study 2

The contributions of Study 2 develop team cognition by going a step further to understand what aspects of team cognition are essential to human-AI teams. Given the significant differences between human-only teams and human-AI teams [212, 224], the idea of what is vital within team cognition (i.e., team situational awareness, shared mental models, transactive memory systems) will likely differ between the two. As such, it would be rash to move forward in designing AI teammates to support team cognition without first asking this question and working to understand what those differences are if they exist. Study 2 found through its mixed-methods approach that **team situational awareness was on the forefront of participants' minds when engaging in a human-AI team**. More related still, the concept of **AI explainability was even more important, and participants considered AI predictability to be a massive component of their perceptions of situational awareness**. This knowledge enables future research on team cognition in human-AI teams to ensure they do not overlook these crucial concepts. It also provides several design recommendations to enhance how AI teammates offer these features. These practical contributions center around developing AI teammates capable of serving as a transactive memory system hub for aspects of the team task best suited for AI, such as highly dynamic information given AI's better processing speeds [285].

6.2.1.3 Contributions from Study 3

Study 3 contributes a significant understanding of the nature of team cognition in human-AI teams to the literature by being the penultimate study of the current dissertation. Specifically, by directly investigating what implementations of AI design features for team cognition are most effective at supporting team situational awareness. Study 3 found that the AI teammate with the SA attribute designed to augment team memory significantly improved team situational awareness, perceived shared mental models, ability to overcome roadblocks, and information pushing team verbal behaviors. The augmenting team memory SA attribute significantly outperformed the situational awareness of information changes and control SA attributes. Furthermore, the benefits of AI participation in transition phase discussions early and late in team life cycles were defined as AI participation in the later transition phase was found to significantly enhance team performance and ability to overcome situational awareness roadblocks, which human-AI teams have struggled to overcome in the past [68]. These results contribute to team cognition research by allowing researchers to **understand the effects of AI design features on team cognition** ***and*** **how transition phases interact with that development over time**. This research also contributes to team cognition research and practice by moving the field from an exploratory approach to team cognition in human-AI teams [238] to a more hands-on approach by testing various implementations. In doing so, Study 3 found that humans want their AI teammate to support their ability to engage in complex coordination through transition phase discussions and augmenting the teams' memory. These contributions' impact on team cognition is significant as they showcase how an artificial entity can support extremely human constructs such as team cognition, proving that AI teammates do not need to be a burden to the typical human processes teams engage in to be effective.

6.2.1.4 Contributions of the Dissertation as a Whole

The current dissertation provides a significant start toward a comprehensive understanding of how team cognition operates and functions in human-AI teams. The applications of these findings to practice are numerous as developers can utilize this knowledge to design AI teammates that engage in the most critical aspects of team cognition, such as team situational awareness [85, 227]. These improvements to team situational awareness will enable human-AI teams to develop better processes in transition and action phases and strengthen human teammates' attitudes toward their AI teammates. These attitudes are necessary as they encourage human teammates to accept contributions from their AI teammates, which will likely include things outside team cognition as AI technologies advance [98, 280, 135]. Such attitudes will also **help in the design of AI capable of encouraging more team cognition behaviors from their fellow human teammates, improving team cognition even further for all team members**. This appears to have been supported by the augmenting team memory SA attribute in Study 3, which saw teams develop a better shared understanding of the task, which supported a comprehensive and naturally resilient level of team situational awareness. As such, the current dissertation provides meaningful insight into how AI teammates can directly impact team cognition in human-AI teams. Future researchers will also benefit as they move on to understanding how other aspects of team cognition influence human-AI teaming outcomes and design better conceptual models for the theory in human-AI teams [238].

6.2.2 Contributions to Human-AI Teaming

The contributions to human-AI teaming throughout the dissertation focus on better understanding the relationship between human and AI teammates and what is expected of each. Existing literature had already investigated whether human teammates tend to accept AI teammates and found that the cohesion between the two was generally relatively low. This often led to poor team cognition and subsequent poor outcomes in tasks demanding high levels of coordination [212, 333, 90]. What exactly did human teammates want from their AI teammates to support team cognition and support more positive attitudes? These are questions encompassed in the research questions posed by the current dissertation, and this body of work contributes to answering them.

6.2.2.1 Contributions from Study 1

Study 1 contributes to human-AI teaming research specifically in two significant ways: 1) advancing the existing research on human-agent teaming by shedding light on the relationship between humans and agents operating in collaborative environments; and 2) advancing real-world design recommendations that promote more human-centered teaming AI and better integrate the two. This is especially important as human-AI teams are poised to become a significant part of the global workforce in the coming decades [238, 277]. By conducting this research, those utilizing human-AI teams will be able to actualize more enjoyable and overall positive experiences for future workforces. This research also expands knowledge of human-AI teams as it was one of the first studies to investigate human teammates' reaction to being the only human member of a human-AI team, which this study found to be a negative experience for those participants. As such, the helpful knowledge to human-AI teaming practitioners was expanded to include an **understanding that having a single human team member on a human-AI team requires additional care to implement in practice**. It also encourages other researchers to understand the reasons behind the negative experience and work to uncover ways to mitigate those adverse reactions to increase the applicability and generalizability of human-AI teams in the real world, as many teams with such a configuration may be necessary [327].

6.2.2.2 Contributions from Study 2

The contributions of Study 2 were focused on understanding what aspects of an AI teammate were desired by human teammates to support those high levels of coordination and positive attitudes towards AI teammates. This study found that human teammates want genuine AI teammates who can provide reassurance and encouragement to support positive attitudes. Furthermore, AI teammates were requested to be in defined roles and to have a significant degree of agency to make meaningful contributions to the overall team goal, which builds upon the discussion of what AI's role should be in teams [214, 291]. These findings develop our understanding of human-AI teams by emphasizing the importance of AI teammate design *and* their placement within these teams. AI teammates should never be called a teammate unless they can live up to the connotations of the word teammate. Study 2 was also essential to **define what exactly human members of human-AI teams want from their AI teammates to foster positive attitudes and support team cognition**. Human-AI teaming benefits from this knowledge by learning what is expected from AI teammates for these crucial aspects of teaming and how they might manipulate them in future studies or design them in practice. Study 2 also highlighted that AI design features elicit a positive response from human teammates.

6.2.2.3 Contributions from Study 3

Study 3 contributes to the research on human-AI teams by better defining the role of the transition phase to human-AI teaming outcomes, highlighting the importance of AI to team situational awareness, and the possibility of implementing AI teammate designs to benefit aspects of human-AI teams. The role of transition phases in human-AI teams has not yet been explored, and this is a glaring omission from the human-AI teaming literature [238]. This study has begun rectifying that gap by directly examining how AI teammates' participation in transition phase processes early versus late in the team life cycle influences human-AI teams' affective and objective performance outcomes, similar to human-only team research [106]. These outcomes were found to benefit from AI participation in both transition phases, but those benefits had different impacts as participation early in the team's life cycle improved shared understanding of the basic aspects of the task. Alternatively, participation later in the teams' life cycle focused on more complex topics that directly influenced team performance and resiliency to unexpected system failures. Study 3 also directly examined how AI teammates impact processes supporting team situational awareness in the form of team verbal behaviors, which found that augmenting team memory enhanced team communication during situational awareness roadblocks and improved teammates' information pushing verbal behaviors. As such, **the study directly highlights how implementations of AI teammate designs for supporting and contributing to team cognition perform in real human-AI teams that work together over an extended period**. Such results enhance the understanding of how AI teammates might be designed to improve team cognition in human-AI teams and several other related emergent states, such as trust, cohesion, and confidence, which also influence team cognition development [93, 159].

6.2.2.4 Contributions of the Dissertation as a Whole

The current dissertation contributes to human-AI teaming in two main thrusts. First, the design of AI teammates will receive a significant wealth of knowledge that precisely measures the benefits of various design features such as AI explainability and situational awareness updates of intra/extra team information changes while measuring workload and performance. Thus, the benefits of these design features can be weighed against the costs they pose to development time and the perceived increase in workload or complexity by the users. Future research will also benefit by gaining a **better understanding of how various design features influence affective and objective outcomes in human-AI teams and their relationship to perceived complexity, allowing for even better potential designs down the road**. Second, this work provides a better understanding of what humans working in human-AI teams want. Similar to the recent research on the ideal AI teammate [342], this research specifically focuses on what human teammates want their AI teammates to do to support team cognition and improved attitudes (i.e., trust and cohesion). Furthermore, this information on human teammates' requests of an AI teammate for team cognition was empirically examined in Study 3, which provided actualized and practical results to those quantitative and qualitative findings from Study 2. Namely, the significant benefits of designing AI teammates to augment team memory by helping them plan ahead and supporting responses to unexpected system failures. Ultimately, these contributions advance the field of human-AI teaming by promoting knowledge on how the needs of human teammates can be practically met by AI teammates to improve coordination.

6.3 Future Work and Generalizability

As is the nature of research, there remain several unanswered questions for future research. The first is what scenario is appropriate for an AI teammate to directly provide teams with the information they need without any context. Study 3 found that pointing out *what* information was needed was better than simply giving human teammates the answer they need; however, it is unknown whether this is always the case. There are likely several scenarios where there is not enough time to do anything other than give human teammates the answer [46]. Future research could also examine if explainability could be used to explain why that specific information was provided, allowing AI teammates to have the best of both worlds. There is also the guarantee that several other information-sharing attributes would significantly improve aspects of team cognition outside of situational awareness. As such, future research should continue examining the effects of AI teammates making contributions to team cognition through information-sharing and how they affect teams in various contexts.

The topic of transition phases is also incredibly under-explored in human-AI teams. One question that should be examined is whether or not AI teammates can be designed to drive transition phase discussions, especially when teams do not know what to talk about and what questions to ask. However, developing AI teammates to drive transition phase discussions risks alienating human teammates, making additional research necessary to understand how AI leadership may be perceived [97]. Transition phases should also be researched for their ability to improve training sequences. As teams go through the training process with their AI teammate, transition phases could be introduced to enhance humans understanding of the AI, their role in the task, and what information is vital to each teammate. Transition phases in training would also allow human teammates to ask fundamental questions about the task to the AI teammate, getting to the benefits of the later transition phases shown in Study 3 sooner.

The current dissertation did not examine how team cognition between an AI and a human teammate developed. However, this remains a significant question in the team cognition and human-AI teaming fields [281]. Whether AI teammates can develop team cognition with their teammates at a conceptual and fundamental level aside, there is a growing need for AI teammates to emulate a personal understanding of individual teammates by adapting to their needs. For example, AI teammates could adjust the information-sharing attributes demonstrated in the current dissertation. The information shared by the AI teammate could change based on the teammate's needs, the current situation the team is facing, or critiques made by the human teammates in a transition phase. Adaptive autonomy will be a significant area of future research within the realm of human-AI team cognition given its ability to emulate another component of human-only team cognition [134, 135], that is, becoming less rigid by adapting actions to match the expectations of their human teammates.

6.3.1 Generalizing Research Findings Across Team Types

Applying the findings of the current dissertation has notable benefits for practitioners and researchers of human-AI teams, but these findings will not be a one-to-one fit for all types of teams. Discussing the generalizability of the aforementioned research findings is important, as the types of teams examined throughout this dissertation's three studies were all small action-based teams. These action-based teams all had interdependent roles, but the nature of their teamwork was reciprocal, which meant their teamwork often required two-way interaction between team members to coordinate and accomplish their shared task [266]. The ability of these findings to generalize to other styles of interdependence in teams, such as pooled (teammates work independently together), sequential (taskwork completed in order one teammate at a time), or team (all teammates interact simultaneously to complete taskwork together) is something that cannot be directly stated [266]. The other types of team interdependence have vastly different levels of complexity depending on the level of coordination overhead and the tasks they must accomplish, let alone teams that change their styles to adapt to changing circumstances. That being said, reciprocal interdependence is the most applicable to action-based teams given its dependence on specific roles [266] and, as such, will allow this research to have the most generalizability to practical and theoretical applications. However, the differences among team types also include the various tasks that teams must accomplish, ranging from creative problem-solving and ideation tasks to those with physical tasks [145].

Tasks requiring creative problem-solving differ considerably from the well-defined problem-solving tasks utilized in the current dissertation studies. Teams that engage with well-defined problem-solving tasks are frequently utilized, but so too are the teams that come together to solve ill-defined problems creatively [326]. As such, it is important to keep in mind the generalizability of these findings to the team types that live in the creative problem-solving space [245]. The question of how AI teammates may contribute to the team cognition of these teams is unknown, and the recommendations made by the current dissertation are unlikely to directly apply, especially given the focus on information-sharing for situational awareness. The role of situational awareness in these creative ideation teams is considerably reduced as these teams do not have well-defined task spaces requiring constant monitoring to enable team success. It is far more likely that AI teammates would be able to contribute meaningful information in these teams by acting as a transactive memory hub, which the current dissertation did not directly examine. Alternatively, these findings contribute meaningfully to teams with more well-defined problems and task spaces, which make up a great deal of the teams in practice today, giving the current dissertation considerable generalizability.

6.4 Closing Remarks

I have been fascinated by the science of teams and the integration of technology since before I started my graduate education. In fact, I told my adviser that I wanted to study team cognition within human-AI teams the first month I started, as I had been aware of how critical team cognition was to effective teaming. Team research is so appealing to me because it represents an incredibly complex field and human dynamic [200]. Teams' are the most frequent modality used by humanity to solve the biggest questions we have faced as a society. However, effective teaming does not happen without good practices and proper development. As incredible as teams can be, there are nearly just as many bad teams that hinder progress and stymie the progression of entire fields. Understanding what separates effective teams from ineffective ones is incredibly important, especially as we begin the process of implementing new and exciting technologies such as AI to enhance the abilities of teams further.

As much as I have been interested in the science of teams, I have been equally enthralled by technology and its role in our society. As such, the opportunity to complete a dissertation focused on how AI can be successfully integrated into teams and enhance aspects of team cognition was ideal for me. I began this dissertation seeking to contribute fundamental knowledge to the field of team cognition and human-AI teaming, and this goal is reflected in the studies I have completed and reported here. All three studies discussed improve our fundamental understanding of a human construct in light of the integration of rapidly advancing technology, just as the field of human-centered computing teaches [124]. Team cognition and even teams themselves are significantly altered by the introduction of AI teammates, and their effect is not necessarily a good one, as Study 1 and others showcase [68]. Studying this adverse effect and seeking to overcome it is also a cornerstone of the field of human-centered computing. Understanding how technology influences the lives of humans and designing that technology to complement and not disrupt lives is a critical aspect of this field.

I am confident that this dissertation presents strong evidence that AI teammates *can* complement team cognition within human-AI teams. As Studies 2 and 3 displayed, AI teammates can provide meaningful contributions to shared knowledge in the form of situational awareness and also improve the team processes supporting team cognition as verbal team behaviors. These improvements were made by developing information-sharing attributes based on previous findings from the literature and studies conducted as part of this dissertation. As such, this evidence can be used to develop more effective AI teammates that improve teams' outcomes across various contexts. A core component of the dissertation was also to develop a series of design recommendations that ensure the research remains relevant to the applied components of human-AI teaming. These design recommendations were developed across all studies of the dissertation and can be used to improve AI teammate design and human-AI team training or deployment.

Moving forward, I have learned much from this experience as a researcher and individual. First, while completing a dissertation is a solitary task, in theory, that is never truly the case. To accomplish significant tasks requires the support and care of others, including co-workers, family members, and friends. Always be willing to help others, and they will more than return the favor one way or another. Second, the pursuit of research is not meant to be a selfish endeavor. Contributing to the state of knowledge is for the betterment of humanity as a whole, and the work we conduct must be capable of standing on its own regardless of any name attached to it, as those are forgotten in time. Third and finally, there is simply no replacement for hard work. There is something to be said for setting a goal and working to achieve it. I have been extremely fortunate to have incredible role models of this in my personal and professional lives who have taught me the value of hard work. As for the research presented in my dissertation, I am confident that AI teammates have a bright future working alongside humans.

Appendices

Appendix A Surveys

* Denotes Reverse Coding

Study 1 Demographics
Enter your Age: (Number Entry)
What is your sex? (Male, Female, Other)
Please rate your familiarity with artificial intelligence. Artificial intelligence being defined as follows "The ability of a machine or a computer program to think and learn. The concept of artificial intelligence is based on the idea of building machines capable of thinking, acting, and learning like humans". A very common example of artificial intelligence would be Siri or Google Assistant. (Not at all familiar, Slightly familiar, Somewhat familiar, Moderately familiar, Extremely familiar)
What is your highest level of education achieved currently? (High school diploma or equivalent, Some undergrad, B.S. or B.A., Some graduate school, Masters degree, Ph.D)
Is English your first and primary language? (Yes, No)
Do you have previous experience working in teams? (Yes, No)
How many teams have you been apart of? (1, 2, 3, 4, 5+)
How comfortable are you working on a team? (Very uncomfortable, Uncomfortable, Neither comfortable nor uncomfortable, Comfortable, Very comfortable)
How familiar are you with the other participants in the room? (Not at all familiar, Slightly familiar, Somewhat familiar, Moderately familiar, Extremely familiar)
How are you familiar with the participant(s)? (Text Entry)
Have you ever worked on a team with the other participants in the room? (Yes, No)

Table 1: Study 1 Demographics.

Study 1 Task Mental Model

Below are several characteristics associated with the interactions that typically occur between team members. Please rate how related each of these characteristics is to the other characteristic in completing the team task you have just completed with your teammates. The relationship goes both ways. The characteristics are as follows:

- Familiarizing with the simulation layout
- Determine which resources are at your individual disposal
- Determine location of event
- Send resource to event if available
- Learn what resources your teammates have
- Recall resources
- Determine resource allocation based on event importance
- Send resources in the correct order for critical events

(-4 Negatively related a high degree of one requires a low degree of the other, -3, -2, -1, 0 Totally Unrelated, 1, 2, 3, 4 Positively related a high degree of one requires a high degree of the other)

- Familiarizing with the simulation layout
- Determine which resources are at your individual disposal

- Familiarizing with the simulation layout
- Determine the location of event

- Familiarizing with the simulation layout
- Send resource to event if available

- Familiarizing with the simulation layout
- Learn what resources your teammates have

- Familiarizing with the simulation layout
- Recall resources

- Familiarizing with the simulation layout
- Determine resource allocation based on event importance

- Familiarizing with the simulation layout
- Send resources in the correct order for critical events

- Determine which resources are at your individual disposal
- Determine location of event

- Determine which resources are at your individual disposal
- Learn what resources your teammates have

- Determine which resources are at your individual disposal
- Recall resources

- Determine which resources are at your individual disposal
- Determine resource allocation based on event importance

- Determine which resources are at your individual disposal
- Send resources in the correct order for critical events

- Determine location of event
- Send resource to event if available

- Determine location of event
- Learn what resources your teammates have

- Determine location of event
- Recall resources

- Determine location of event
- Determine resource allocation based on event importance

- Determine location of event
- Send resources in the correct order for critical events

- Send resource to event if available
- Learn what resources your teammates have

- Send resource to event if available
- Recall resources

- Send resource to event if available
- Determine resource allocation based on event importance

- Send resource to event if available
- Send resources in the correct order for critical events

- Learn what resources your teammates have
- Recall resources

- Learn what resources your teammates have
- Determine resource allocation based on event importance

- Learn what resources your teammates have
- Send resources in the correct order for critical events

- Recall resources
- Determine resource allocation based on event importance

- Recall resources
- Send resources in the correct order for critical events

- Determine resource allocation based on event importance
- Send resources in the correct order for critical events

Table 2: Study 1 Task Mental Model.

Study 1 Team Mental Model

Below are several characteristics associated with the interactions tat typically occur between team members. Please rate how related each of these characteristics is to the other characteristic in completing the team task you just completed with your teammates. The relationship goes both ways. The characteristics are as follows:

- Amount of Information-The amount of information existing within a team
- Quality of information-The general ability to use the information within the team
- Role/Responsibility-Usual or expected function of a given team member, the tasks for which a team member is accountable
- Interaction Patterns-Common communication between or joint activity involving team members
- Communication Channels-Ways (modes) that the team uses to communicate
- Role Interdependencies-Relying on mutual assistance, support, cooperation, or interaction among team members' roles
- Teammates' Skill-General team members' ability to do something well, usually gained through experience and training
- Teammates' Attitudes-Team members' opinion or general feeling about something
- Teammates' Preferences-Team members' views that a particular course of action is more desirable than another

(-4 Negatively related a high degree of one requires a low degree of the other, -3, -2, -1, 0 Totally Unrelated, 1, 2, 3, 4 Positively related a high degree of one requires a high degree of the other)

- Amount of information
- Quality of information

- Amount of information
- Role/Responsibility

- Amount of information
- Interaction patterns

- Amount of information
- Communication channels

- Amount of information
- Role interdependencies

- Amount of information
- Teammates' skill

- Amount of information
- Teammates' attitudes

- Amount of information
- Teammates' preferences

- Quality of information
- Role/Responsibility

- Quality of information
- Interaction patterns

- Quality of information
- Communication channels

- Quality of information
- Role interdependencies

- Quality of information
- Teammates' skill

- Quality of information
- Teammates' attitudes

- Quality of information
- Teammates' preferences

- Role/Responsibility
- Interaction patterns

- Role/Responsibility
- Communication channels

- Role/Responsibility
- Role interdependencies

- Role/Responsibility
- Teammates' skill

- Role/Responsibility
- Teammates' attitudes

- Role/Responsibility
- Teammates' preferences

- Interaction patterns
- Communication channels

- Interaction patterns
- Role interdependencies

- Interaction patterns
- Teammates' skill

- Interaction patterns
- Teammates' attitudes

- Interaction patterns
- Teammates' preferences

- Communication channels
- Role interdependencies

- Communication channels
- Teammates' skill

- Communication channels
- Teammates' attitudes

- Communication channels
- Teammates' preferences

- Role interdependencies
- Teammates' skill

- Role interdependencies
- Teammates' attitudes

- Role interdependencies
- Teammates' preferences

- Teammates' skill
- Teammates' attitudes

- Teammates' skill
- Teammates' preferences

- Teammates' attitudes
- Teammates' preferences

Table 3: Study 1 Team Mental Model.

Study 1 Team Member Schema (Self)

Think about what teamwork means to you. Think about teamwork as it may occur on any team. In other words, try not to think about any specific team, but rather think about teams and teamwork in general. Thinking about teamwork in this way, please read each of the following statements. When reading these statements, think about how important these behaviors are to your meaning of teamwork. Ask yourself, "Does this behavior tell me anything about the meaning of teamwork?" When considering the importance of each item, keep in mind your view or meaning of teamwork.

For example, consider the statement "Team members do not interrupt each other." This behavior could have a variety of meanings regarding teamwork (e.g., politeness, lack of assertiveness). Keeping in mind what the statement means to you about teamwork, **rate its importance to your meaning of teamwork**. Thus, if the above statement indicates politeness to you, and you think politeness is not very important to your idea of teamwork, then you should rate "Team members do not interrupt each other" as not important.

PLEASE READ EACH ITEM CAREFULLY and respond using only the following scale. (7-point Likert, Extremely Unimportant $\Longleftrightarrow$ Extremely Important).

The team continuously re-evaluates its strategy
The team makes decisions with information provided provided by movements
The team expects to make mistakes
Personal preferences are compromised to meet the team goals
Team members know when a mistake has been made
Team members are aware of the task at hand
Each player makes the decision on which objectives they are responsible for
Team members have various roles & tasks based on that role
The team adapts to each new situation based on previous experience
The team supports one another in the completion of the task
Team members are open to adapting their movements
Team members are active participants
Members prompt one another to certain strategies or paths through moves on the board
Team members focus on the overall team effort

Table 4: Study 1 Team Member Schema (Self).

Study 1 Team Member Schema (Human Teammate)

Think about what teamwork means to your human partners. Think about teamwork as it may occur on any team. In other words, try not to think about any specific team, but rather think about teams and teamwork in general. Thinking about teamwork in this way, please read each of the following statements. When reading these statements, think about how important these behaviors are to your meaning of teamwork. Ask yourself, "Does this behavior tell my human partners anything about the meaning of teamwork?" When considering the importance of each item, keep in mind your partner's view or meaning of teamwork.

For example, consider the statement "Team members do not interrupt each other." This behavior could have a variety of meanings regarding teamwork (e.g., politeness, lack of assertiveness). Keeping in mind what the statement means to your human partners about teamwork, **rate its importance to his or her meaning of teamwork**. Thus, if the above statement indicates politeness to your human partners, and you think politeness is not very important to your human partner's idea of teamwork, then you should rate "Team members do not interrupt each other" as not important.

PLEASE READ EACH ITEM CAREFULLY and respond using only the following scale. (7-point Likert, Extremely Unimportant $\Longleftrightarrow$ Extremely Important).

The team continuously re-evaluates its strategy

The team makes decisions with information provided provided by movements

The team expects to make mistakes

Personal preferences are compromised to meet the team goals

Team members know when a mistake has been made

Team members are aware of the task at hand

Each player makes the decision on which objectives they are responsible for

Team members have various roles & tasks based on that role
The team adapts to each new situation based on previous experience
The team supports one another in the completion of the task
Team members are open to adapting their movements
Team members are active participants
Members prompt one another to certain strategies or paths through moves on the board
Team members focus on the overall team effort

Table 5: Study 1 Team Member Schema (Human Teammate).

Study 1 Team Member Schema (AI Teammate)

Think about what teamwork means to your artificial intelligence (AI) partner. Think about teamwork as it may occur on any team. In other words, try not to think about any specific team, but rather think about teams and teamwork in general. Thinking about teamwork in this way, please read each of the following statements. When reading these statements, think about how important these behaviors are to your meaning of teamwork. Ask yourself, "Does this behavior tell my AI partner anything about the meaning of teamwork?" When considering the importance of each item, keep in mind your partner's view or meaning of teamwork.

For example, consider the statement "Team members do not interrupt each other." This behavior could have a variety of meanings regarding teamwork (e.g., politeness, lack of assertiveness). Keeping in mind what the statement means to your AI partner about teamwork, **rate its importance to their meaning of teamwork**. Thus, if the above statement indicates politeness to your AI partner, and you think politeness is not very important to your AI partner's idea of teamwork, then you should rate "Team members do not interrupt each other" as not important.

PLEASE READ EACH ITEM CAREFULLY and respond using only the following scale. (7-point Likert, Extremely Unimportant $\Longleftrightarrow$ Extremely Important).

The team continuously re-evaluates its strategy

The team makes decisions with information provided provided by movements

The team expects to make mistakes

Personal preferences are compromised to meet the team goals

Team members know when a mistake has been made

Team members are aware of the task at hand

Each player makes the decision on which objectives they are responsible for

Team members have various roles & tasks based on that role

The team adapts to each new situation based on previous experience

The team supports one another in the completion of the task

Team members are open to adapting their movements

Team members are active participants

Members prompt one another to certain strategies or paths through moves on the board

Team members focus on the overall team effort

Table 6: Study 1 Team Member Schema (AI Teammate).

Study 1 Trust in AI Teammate
Please read each of the following items carefully and respond using only the following scale. (5-point Likert, Strongly Disagree $\Longleftrightarrow$ Strongly Agree).
Did you trust the artificial intelligence that you worked with? (5-point Likert, Definitely Not $\Longleftrightarrow$ Definitely Yes)
Did you feel confident in the AI you just worked with?
Did you feel that you had to monitor the AI's actions during the game?*
Did you feel that the AI had harmful motives in the game?*
Did you feel fearful, paranoid, or skeptic of the AI during the game?*
Did you feel that the AI allowed joint problem solving in the game?

Table 7: Study 1 Trust in AI Teammate.

Study 1 and Study 3 Perceived Team Performance Scale

Indicate the degree to which you agree with each statement (5-point Likert, Strongly Disagree $\Longleftrightarrow$ Strongly Agree).
Team members 'carried their weight' during the task
Members were highly committed to the team during the task
The researcher will be satisfied with the team product
People outside of the team would give the team positive feedback about this work today
The researcher would be satisfied with the team's performance
Team members worked better together at the end of the task than at the beginning
Team members were more aware of group dynamics at the end of the task than when they began the task
Being a part of this team helped members appreciate different types of people

Table 8: Study 1 and Study 3 Perceived Team Performance Scale.

Study 1 Qualitative Text Response Questions
For the following questions think back on your experience completing the team task simulation and provide responses to the following open ended questions:
Do you feel that everyone on your team thought about cooperating and responding to events the same? If not, why?
Do you feel team cognition was established within your team? (Team cognition is the shared understanding between team members of team member resources, their roles, and how the team is supposed to operate and respond to their shared tasks)

Did that shared understanding or team cognition happen in the early games or the later games?
If possible, can you provide a specific example of a time where you felt the team displayed an act of shared understanding or team cognition? (Example: A team member helping out with an event you could not immediately dispatch a resource to)
Do you believe you trusted your autonomous teammate(s)? Why or why not?
Do you feel like having a real time spatial map of the task space helped or hindered your team in developing a shared understanding or team cognition?
Do you feel like you and your other human teammate paid much attention to your one autonomous teammate? Why or why not?
How would you describe your experience interacting with your autonomous teammate?

Table 9: Study 1 Qualitative Text Response Questions.

Study 2 Demographics
Please enter your current age: (Number Entry)
Please specify your identified gender: (Male, Female, Non-binary/third gender, Prefer not to say, Prefer to specify-Text Entry)
What is the highest level of school you have completed or the highest degree you have received? (Less than high school degree, High school graduate or equivalency, Some college but no degree, Associate degree in college, Bachelor's degree in college, Master's degree, Doctoral degree, Professional degree)
Choose one or more races that you consider yourself to be: (White, Black or African American, American Indian or Alaska Native, Asian, Native Hawaiian or Pacific Islander, Latino or Hispanic, Other-Text Entry)

Is English your native language? (Yes, No)
How often do you play multiplayer digital games? (On A Daily Basis, On A Weekly Basis, Several Times A Month, Once Every Month, Several Times Per Year, Never)
How much experience do you have with playing video games? (None at All, A Little, A Moderate Amount, A Lot, A Great Deal)
What genre video games do you primarily play? (First Person Shooter, Role Playing Game, Massively Online Multiplayer Role Playing Game, Strategy, Sports, Multiplayer Online Battle Arena, Other-Text Entry)
What video games do you frequently play specifically? (Text Entry)
Which platform do you play video games on typically? (PC, Console, Other-Text Entry)
What input device (i.e., controller, mouse, keyboard) do you use when playing video games? (Mouse and Keyboard, Controller, Other-Text Entry)
Please list the video games you currently play that you feel you're an expert at (i.e., have over 100 hours playing)? (Text Entry)
Do you play video games where you work with artificial intelligence (AI) based teammates? (Yes, No)
How many times per week do you play video games with AI teammates? (Text Entry)
What game mode do you typically play when working with AI teammates? (Capture The Flag, Territory Capture, Death Match, Other-Text Entry)

Table 10: Study 2 Demographics.

Study 2 Single Item Measures
Please rate your agreement with the following statements: (7-point Likert, Strongly Disagree $\Longleftrightarrow$ Strongly Agree).

AI NAME has valuable skills that benefit the team. (**AI Teammate Skill**)
AI NAME improves the team's understanding of the current situation. (**Perceived Situational Awareness**)
Working with *AI NAME* overcomplicates the team's task. (**AI Task Disruption**)
I believe *HUMAN NAME* (Teammate B) would be a good teammate. (**Human Teammate Rating**)
I believe *AI NAME* (Teammate C) would be a good teammate. (**AI Teammate Rating**)
I trust *AI NAME*. (**Trust in AI Teammate**)
I trust *HUMAN NAME*. (**Trust in Human Teammate**)

Table 11: Study 2 Single Item Measures.

Study 2 Perceived Shared Mental Model
Please indicate your level of agreement to the following statements about your *AI/human* teammate in the human-AI teaming scenario described above: (7-point Likert, Strongly Disagree $\Longleftrightarrow$ Strongly Agree).
I believe *AI OR HUMAN NAME* and I have a similar understanding about specific strategies for completing the task in the scenario.
I believe *AI OR HUMAN NAME* and I have a similar understanding about how to communicate with each other in the scenario.
I believe *AI OR HUMAN NAME* and I have a similar understanding about sharing information with the team in the scenario.

Table 12: Study 2 Perceived Shared Mental Model.

Study 2 Perceived Influence Over the Team
Please indicate whether your AI teammate (Alpha) or you are more aligned with the following questions: (5-point Likert, Definitely the AI Teammate $\Longleftrightarrow$ Definitely Myself).
Who felt they had the most influence on what happened in this situation? Who felt they had the most influence on the action that was taken? Who felt they had the least influence on what happened in the situation? Who did you feel had the least influence on the action carried out?

Table 13: Study 2 Perceived Influence Over the Team.

Study 2 Perceived Information Certainty
Please indicate your level of agreement to the following statements regarding the human-AI teaming scenario described above: (7-point Likert, Strongly Disagree $\Longleftrightarrow$ Strongly Agree).
I believe I understand how my action would affect *AI NAME*. I believe I know what *AI NAME* is planning in this situation. I believe I am informed about *AI NAME* planned action in this situation. I believe I know why *AI NAME* prefers a certain action.

Table 14: Study 2 Perceived Information Certainty.

Study 3 Demographics

Please enter your current age: (Number Entry)
Please specify your identified gender: (Male, Female, Non-binary/third gender, Prefer not to say, Prefer to specify-Text Entry)
What is the highest level of school you have completed or the highest degree you have received? (Less than high school degree, High school graduate or equivalency, Some college but no degree, Associate degree in college, Bachelor's degree in college, Master's degree, Doctoral degree, Professional degree)
Choose one or more races that you consider yourself to be: (White, Black or African American, American Indian or Alaska Native, Asian, Native Hawaiian or Pacific Islander, Latino or Hispanic, Other-Text Entry)
Is English your native language, or do you have English fluency? (Yes, No)
How much experience do you have with playing video games? (None at All, A Little, A Moderate Amount, A Lot, A Great Deal)
How familiar are you with the other participants in the room? (Not at All Familiar, Slightly Familiar, Somewhat Familiar, Moderately Familiar, Extremely Familiar)
What is your current class standing? (Freshman, Sophomore, Junior, Senior, Graduate Student, Not Enrolled)
Please enter your major field of study: (Text Entry) *If "Not Enrolled" was selected, then this question did not appear.*

Table 15: Study 3 Demographic Questions.

Study 3 Trust Survey
Please answer the following questions in regards to the artificial teammate you worked with in the last (3rd) game (5-point Likert).

In general, I trusted the *AI OR HUMAN* teammate I just worked with.

I felt like I had to monitor my *AI OR HUMAN* teammate's actions during the game.

I felt like my *AI OR HUMAN* teammate had harmful motives in the task.*

I felt confident in the *AI OR HUMAN* teammate I just worked with.*

I felt like my *AI OR HUMAN* teammate allowed joint problem-solving in the task.

I felt fearful, paranoid, and or skeptical of my *AI OR HUMAN* teammate during the game.*

Table 16: Study 3 Teammate Trust Survey.

Study 3 Perceived Shared Mental Model

Please answer the following questions regarding your hypothetical teammate (7-point Likert scale, Strongly Disagree $\Longleftrightarrow$ Strongly Agree). For each question, "My *AI/HUMAN* teammate and I have a similar understanding about..."

Execution

Specific strategies for completing various tasks.

How to deal with the task.

How best to perform our tasks.

The relationships between tasks.

Interaction

How to communicate with each other.

Sharing information with each other.

How we should interact with each other.

Temporal

Our deadlines.
How quickly we need to work.
Appropriately timing our work.
Coordinating the timing of our work.

Table 17: Study 3 Perceived Shared Mental Model.

Study 3 Situational Awareness Rating Technique (SART)
Please answer the following questions in regards to your experience with your team in the previous mission. (7-point Likert scale, Very Low $\Longleftrightarrow$ Very High)).
How changeable is the situation? Is the situation highly unstable and likely to change suddenly (High) or is it very stable and straightforward (Low)?
How complicated is the system? Is it complex with many interrelated components (High) or is it simple and straightforward (Low)?
How many variables are changing within the situation? Are there a large number of factors varying (High) or are there very few variables changing (Low)?
How aroused are you in the situation? Are you alert and ready for activity (High) or do you have a low degree of alertness (Low)?
How much are you concentrating on the situation? Are you concentrating on many aspects of the situation (High) or focused on only one (Low)?
How much is your attention divided in the situation? Are you concentrating on many aspects of the situation (High) or focused on only one (Low)?
How much mental capacity do you have to spare in the situation? Do you have sufficient resources to attend to many variables (High) or nothing to spare at all (Low)?

How much information have you gained about the situation? Have you received and understood a great deal of knowledge (High) or very little (Low)?	
How familiar are you with the situation? Do you have a great deal of relevant experience (High) or is it a new situation (Low)?	

Table 18: Study 3 Situational Awareness Rating Technique (SART).

Bibliography

[1] Martín Abadi, Paul Barham, Jianmin Chen, Zhifeng Chen, Andy Davis, Jeffrey Dean, Matthieu Devin, Sanjay Ghemawat, Geoffrey Irving, Michael Isard, et al. {TensorFlow}: a system for {Large-Scale} machine learning. In *12th USENIX symposium on operating systems design and implementation (OSDI 16)*, pages 265–283, 2016.

[2] Saleema Amershi, Dan Weld, Mihaela Vorvoreanu, Adam Fourney, Besmira Nushi, Penny Collisson, Jina Suh, Shamsi Iqbal, Paul N Bennett, Kori Inkpen, et al. Guidelines for human-ai interaction. In *Proceedings of the 2019 chi conference on human factors in computing systems*, pages 1–13, 2019.

[3] Mandy M Archibald, Rachel C Ambagtsheer, Mavourneen G Casey, and Michael Lawless. Using zoom videoconferencing for qualitative data collection: perceptions and experiences of researchers and participants. *International Journal of Qualitative Methods*, 18:1609406919874596, 2019.

[4] Katrin Auspurg and Thomas Hinz. *Factorial survey experiments*, volume 175. Sage Publications, 2014.

[5] John R Austin. *Knowing what other people know: Situated expertise and assistance seeking behavior in cross-functional teams*. Boston College, 1998.

[6] John R Austin. Transactive memory in organizational groups: the effects of content, consensus, specialization, and accuracy on group performance. *Journal of applied psychology*, 88(5):866, 2003.

[7] Angela J Bahns, Christian S Crandall, Omri Gillath, and Kristopher J Preacher. Similarity in relationships as niche construction: Choice, stability, and influence within dyads in a free choice environment. *Journal of Personality and Social Psychology*, 112(2):329, 2017.

[8] Gagan Bansal, Besmira Nushi, Ece Kamar, Walter S Lasecki, Daniel S Weld, and Eric Horvitz. Beyond accuracy: The role of mental models in human-ai team performance. In *Proceedings of the AAAI Conference on Human Computation and Crowdsourcing*, volume 7, pages 2–11, 2019.

[9] Gagan Bansal, Besmira Nushi, Ece Kamar, Daniel S Weld, Walter S Lasecki, and Eric Horvitz. Updates in human-ai teams: Understanding and addressing the performance/compatibility tradeoff. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 33, pages 2429–2437, 2019.

[10] Bernard M Bass, Bruce J Avolio, Dong I Jung, and Yair Berson. Predicting unit performance by assessing transformational and transactional leadership. *Journal of applied psychology*, 88(2):207, 2003.

[11] S Gayle Baugh and George B Graen. Effects of team gender and racial composition on perceptions of team performance in cross-functional teams. *Group & Organization Management*, 22(3):366–383, 1997.

[12] Daniel J Beal, Robin R Cohen, Michael J Burke, and Christy L McLendon. Cohesion and performance in groups: a meta-analytic clarification of construct relations. *Journal of applied psychology*, 88(6):989, 2003.

[13] Lisa Beeler, Alex R Zablah, and Adam Rapp. Ability is in the eye of the beholder: How context and individual factors shape consumer perceptions of digital assistant ability. *Journal of Business Research*, 148:33–46, 2022.

[14] Herbert H Bell and Don R Lyon. Using observer ratings to assess situation awareness. 2000.

[15] Pernille Bjørn, Morten Esbensen, Rasmus Eskild Jensen, and Stina Matthiesen. Does distance still matter? revisiting the cscw fundamentals on distributed collaboration. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 21(5):1–26, 2014.

[16] Cheryl A Bolstad and Mica R Endsley. Measuring shared and team situation awareness in the army's future objective force. In *Proceedings of the human factors and ergonomics society annual meeting*, volume 47, pages 369–373. Sage Publications Sage CA: Los Angeles, CA, 2003.

[17] Louis J Bourgeois III. Performance and consensus. *Strategic management journal*, 1(3):227–248, 1980.

[18] Michael W Boyce, Jessie YC Chen, Anthony R Selkowitz, and Shan G Lakhmani. Effects of agent transparency on operator trust. In *Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction Extended Abstracts*, pages 179–180, 2015.

[19] Ralph Allan Bradley and Milton E Terry. Rank analysis of incomplete block designs: I. the method of paired comparisons. *Biometrika*, 39(3/4):324–345, 1952.

[20] Erin Bradner and Gloria Mark. Why distance matters: effects on cooperation, persuasion and deception. In *Proceedings of the 2002 ACM conference on Computer supported cooperative work*, pages 226–235, 2002.

[21] David P Brandon and Andrea B Hollingshead. Transactive memory systems in organizations: Matching tasks, expertise, and people. *Organization science*, 15(6):633–644, 2004.

[22] Virginia Braun and Victoria Clarke. Thematic analysis. 2012.

[23] Michael J Burtscher, Michaela Kolbe, Johannes Wacker, and Tanja Manser. Interactions of team mental models and monitoring behaviors predict team performance in simulated anesthesia inductions. *Journal of Experimental Psychology: Applied*, 17(3):257, 2011.

[24] John T Bush, Jeffery A LePine, and Daniel W Newton. 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, 2018.

[25] David F. Caldwell and Charles A. O'Reilly. The Determinants of Team-Based Innovation in Organizations: The Role of Social Influence. *Small Group Research*, 34(4):497–517, August 2003. Publisher: SAGE Publications Inc.

[26] Gloria Calhoun. Adaptable (not adaptive) automation: Forefront of human–automation teaming. *Human Factors*, 64(2):269–277, 2022.

[27] Michael A Campion, Gina J Medsker, and A Catherine Higgs. Relations between work group characteristics and effectiveness: Implications for designing effective work groups. *Personnel psychology*, 46(4):823–847, 1993.

[28] J. A. Cannon-Bowers, E. Salas, and S. A. Converse. Cognitive psychology and team training: Shared mental models in complex systems. In *5th Annual Conference ofthe Society for Industrial and Organizational Psychology, Miami, florida*, 1990.

[29] J. A. Cannon-Bowers, E. Salas, and Sharolyn Converse. Shared mental models in expert team decision making. *Individual and group decision making: Current issues*, 221:221–46, 1993. Publisher: Hillsdale.

[30] Janis A Cannon-Bowers and Eduardo Salas. Reflections on shared cognition. *Journal of Organizational Behavior: The International Journal of Industrial, Occupational and Organizational Psychology and Behavior*, 22(2):195–202, 2001.

[31] Lorenzo Barberis Canonico, Christopher Flathmann, and Dr. Nathan McNeese. Collectively intelligent teams: Integrating team cognition, collective intelligence, and ai for future teaming. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 63(1):1466–1470, 2019.

[32] Albert V Carron, Lawrence R Brawley, and William N Widmeyer. The measurement of cohesiveness in sportgroups. 1998.

[33] Jessie Y Chen, Katelyn Procci, Michael Boyce, Julia L Wright, Andre Garcia, and Michael Barnes. Situation awareness-based agent transparency. Technical report, ARMY RESEARCH LAB ABERDEEN PROVING GROUND MDUNIVERSITY OF CENTRAL FLORIDA ..., 2014.

[34] Jessie YC Chen and Michael J Barnes. Supervisory control of robots using roboleader. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, volume 54, pages 1483–1487. SAGE Publications Sage CA: Los Angeles, CA, 2010.

[35] Jessie YC Chen, Michael J Barnes, Anthony R Selkowitz, and Kimberly Stowers. Effects of agent transparency on human-autonomy teaming effectiveness. In *2016 IEEE international conference on Systems, man, and cybernetics (SMC)*, pages 001838–001843. IEEE, 2016.

[36] Jessie YC Chen, Shan G Lakhmani, Kimberly Stowers, Anthony R Selkowitz, Julia L Wright, and Michael Barnes. Situation awareness-based agent transparency and human-autonomy teaming effectiveness. *Theoretical issues in ergonomics science*, 19(3):259–282, 2018.

[37] Shih-Yi Chien, Michael Lewis, Katia Sycara, Jyi-Shane Liu, and Asiye Kumru. The effect of culture on trust in automation: reliability and workload. *ACM Transactions on Interactive Intelligent Systems (TiiS)*, 8(4):1–31, 2018.

[38] Erin K Chiou and John D Lee. Trusting automation: Designing for responsivity and resilience. *Human factors*, page 00187208211009995, 2021.

[39] Leah Chong, Guanglu Zhang, Kosa Goucher-Lambert, Kenneth Kotovsky, and Jonathan Cagan. Human confidence in artificial intelligence and in themselves: The evolution and impact of confidence on adoption of ai advice. *Computers in Human Behavior*, 127:107018, 2022.

[40] K Chowdhary. Natural language processing. *Fundamentals of artificial intelligence*, pages 603–649, 2020.

[41] Mashrur Chowdhury and Adel W Sadek. Advantages and limitations of artificial intelligence. *Artificial intelligence applications to critical transportation issues*, 6(3):360–375, 2012.

[42] EVEV CLARK. 15 common ground. *The handbook of language emergence*, page 328, 2015.

[43] Herbert H Clark. *Using language*. Cambridge university press, 1996.

[44] Herbert H Clark and Susan E Brennan. Grounding in communication. 1991.

[45] Jacob Cohen. *Statistical power analysis for the behavioral sciences*. Academic press, 1988.

[46] Stanley C Collyer and Gerald S Malecki. Tactical decision making under stress: History and overview. 1998.

[47] Gregorio Convertino, Helena M Mentis, Mary Beth Rosson, Aleksandra Slavkovic, and John M Carroll. Supporting content and process common ground in computer-supported teamwork. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 2339–2348, 2009.

[48] Nancy J Cooke. Team cognition as interaction. *Current directions in psychological science*, 24(6):415–419, 2015.

[49] Nancy J Cooke, Mustafa Demir, and Nathan McNeese. Synthetic teammates as team players: Coordination of human and synthetic teammates. Technical report, Cognitive Engineering Research Institute Mesa United States, 2016.

[50] Nancy J Cooke, Jamie Gorman, Harry Pedersen, Jennifer Winner, Jasmine Duran, Amanda Taylor, Polemnia G Amazeen, Dee H Andrews, and Leah Rowe. Acquisition and retention of team coordination in command-and-control. Technical report, COGNITIVE ENGINEERING RESEARCH INST MESA AZ, 2007.

[51] Nancy J Cooke, Jamie C Gorman, Jasmine L Duran, and Amanda R Taylor. Team cognition in experienced command-and-control teams. *Journal of Experimental Psychology: Applied*, 13(3):146, 2007.

[52] Nancy J Cooke, Jamie C Gorman, Christopher W Myers, and Jasmine L Duran. Interactive team cognition. *Cognitive science*, 37(2):255–285, 2013.

[53] Nancy J Cooke, Jamie C Gorman, and Jennifer L Winner. Team cognition. 2007.

[54] Nancy J Cooke, Preston A Kiekel, and Erin E Helm. Measuring team knowledge during skill acquisition of a complex task. *International Journal of Cognitive Ergonomics*, 5(3):297–315, 2001.

[55] Nancy J Cooke, Preston A Kiekel, Eduardo Salas, Renée Stout, Clint Bowers, and Janis Cannon-Bowers. Measuring team knowledge: A window to the cognitive underpinnings of team performance. *Group Dynamics: Theory, Research, and Practice*, 7(3):179, 2003.

[56] Nancy J Cooke, Eduardo Salas, Janis A Cannon-Bowers, and Renee J Stout. Measuring team knowledge. *Human factors*, 42(1):151–173, 2000.

[57] Nancy J Cooke and Steven M Shope. Synthetic task environments for teams: Certt's uav-ste. In *Handbook of human factors and ergonomics methods*, pages 476–483. CRC Press, 2004.

[58] Chris W Coultas, Tripp Driskell, C Shawn Burke, and Eduardo Salas. A conceptual review of emergent state measurement: Current problems, future solutions. *Small Group Research*, 45(6):671–703, 2014.

[59] Haydee M Cuevas, Stephen M Fiore, Barrett S Caldwell, and Laura Strater. Augmenting team cognition in human-automation teams performing in complex operational environments. *Aviation, space, and environmental medicine*, 78(5):B63–B70, 2007.

[60] Kevin Daniels, Leslie De Chernatony, and Gerry Johnson. Validating a method for mapping managers' mental models of competitive industry structures. *Human Relations*, 48(9):975–991, 1995.

[61] Ewart J De Visser, Marieke MM Peeters, Malte F Jung, Spencer Kohn, Tyler H Shaw, Richard Pak, and Mark A Neerincx. Towards a theory of longitudinal trust calibration in human–robot teams. *International journal of social robotics*, 12(2):459–478, 2020.

[62] Leslie A DeChurch and Jessica R Mesmer-Magnus. The cognitive underpinnings of effective teamwork: a meta-analysis. *Journal of applied psychology*, 95(1):32, 2010.

[63] William DeLone, J Alberto Espinosa, Gwanhoo Lee, and Erran Carmel. Bridging global boundaries for is project success. In *Proceedings of the 38th Annual Hawaii International Conference on System Sciences*, pages 48b–48b. IEEE, 2005.

[64] Mustafa Demir and Nancy J Cooke. Human teaming changes driven by expectations of a synthetic teammate. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, volume 58, pages 16–20. SAGE Publications Sage CA: Los Angeles, CA, 2014.

[65] Mustafa Demir, Nancy J Cooke, and Polemnia G Amazeen. A conceptual model of team dynamical behaviors and performance in human-autonomy teaming. *Cognitive Systems Research*, 52:497–507, 2018.

[66] Mustafa Demir, Aaron D Likens, Nancy J Cooke, Polemnia G Amazeen, and Nathan J McNeese. Team coordination and effectiveness in human-autonomy teaming. *IEEE Transactions on Human-Machine Systems*, 49(2):150–159, 2018.

[67] Mustafa Demir, Nathan J McNeese, and Nancy J Cooke. Team situation awareness within the context of human-autonomy teaming. *Cognitive Systems Research*, 46:3–12, 2017.

[68] Mustafa Demir, Nathan J McNeese, and Nancy J Cooke. The impact of perceived autonomous agents on dynamic team behaviors. *IEEE Transactions on Emerging Topics in Computational Intelligence*, 2(4):258–267, 2018.

[69] Mustafa Demir, Nathan J McNeese, and Nancy J Cooke. Understanding human-robot teams in light of all-human teams: Aspects of team interaction and shared cognition. *International Journal of Human-Computer Studies*, 140:102436, 2020.

[70] Mustafa Demir, Nathan J McNeese, Nancy J Cooke, Jerry T Ball, Christopher Myers, and Mary Frieman. Synthetic teammate communication and coordination with humans. In *Proceedings of the human factors and ergonomics society annual meeting*, volume 59, pages 951–955. SAGE Publications Sage CA: Los Angeles, CA, 2015.

[71] Mustafa Demir, Nathan J McNeese, Manrong She, and Nancy J Cooke. Team coordination of team situation awareness in human-autonomy teaming. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, volume 63, pages 146–147. SAGE Publications Sage CA: Los Angeles, CA, 2019.

[72] Terry L Dickinson and Robert M McIntyre. A conceptual framework for teamwork measurement. In *Team performance assessment and measurement*, pages 31–56. Psychology Press, 1997.

[73] Paul Dourish and Victoria Bellotti. Awareness and coordination in shared workspaces. In *Proceedings of the 1992 ACM conference on Computer-supported cooperative work*, pages 107–114, 1992.

[74] Jeff Druce, James Niehaus, Vanessa Moody, David Jensen, and Michael L Littman. Brittle ai, causal confusion, and bad mental models: Challenges and successes in the xai program. *arXiv preprint arXiv:2106.05506*, 2021.

[75] Alpana Dubey, Kumar Abhinav, Sakshi Jain, Veenu Arora, and Asha Puttaveerana. Haco: a framework for developing human-ai teaming. In *Proceedings of the 13th Innovations in Software Engineering Conference on Formerly known as India Software Engineering Conference*, pages 1–9, 2020.

[76] Francis T Durso, Carla A Hackworth, Todd R Truitt, Jerry Crutchfield, Danko Nikolic, and Carol A Manning. Situation awareness as a predictor of performance for en route air traffic controllers. *Air Traffic Control Quarterly*, 6(1):1–20, 1998.

[77] Clarence A Ellis, Simon J Gibbs, and Gail Rein. Groupware: some issues and experiences. *Communications of the ACM*, 34(1):39–58, 1991.

[78] Mica Endsley and William M Jones. Situation awareness information dominance & information warfare. Technical report, LOGICON TECHNICAL SERVICES INC DAYTON OH, 1997.

[79] Mica R Endsley. Sagat: A methodology for the measurement of situation awareness (nor doc 87-83). *Hawthorne, CA: Northrop Corporation*, 18, 1987.

[80] Mica R Endsley. Design and evaluation for situation awareness enhancement. In *Proceedings of the Human Factors Society annual meeting*, volume 32, pages 97–101. Sage Publications Sage CA: Los Angeles, CA, 1988.

[81] Mica R Endsley. Situation awareness global assessment technique (sagat). In *Proceedings of the IEEE 1988 national aerospace and electronics conference*, pages 789–795. IEEE, 1988.

[82] Mica R Endsley. Toward a theory of situation awareness in dynamic systems. In *Situational awareness*, pages 9–42. Routledge, 1995.

[83] Mica R. Endsley. From here to autonomy: lessons learned from human–automation research. *Human factors*, 59(1):5–27, 2017.

[84] Mica R Endsley. Situation awareness. *Handbook of human factors and ergonomics*, pages 434–455, 2021.

[85] Mica R Endsley. Supporting human-ai teams: Transparency, explainability, and situation awareness. *Computers in Human Behavior*, page 107574, 2022.

[86] Mica R Endsley, Betty Bolté, and Debra G Jones. *Designing for situation awareness: An approach to user-centered design*. CRC press, 2003.

[87] Alberto Espinosa, Jonathan Cadiz, Luis Rico-Gutierrez, Robert Kraut, William Scherlis, and Glenn Lautenbacher. Coming to the wrong decision quickly: why awareness tools must be matched with appropriate tasks. In *Proceedings of the*

SIGCHI conference on Human factors in computing systems, pages 392–399, 2000.

[88] Neta Ezer, Sylvain Bruni, Yang Cai, Sam J Hepenstal, Christopher A Miller, and Dylan D Schmorrow. Trust engineering for human-ai teams. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, volume 63, pages 322–326. SAGE Publications Sage CA: Los Angeles, CA, 2019.

[89] Xiaocong Fan, Shuang Sun, Michale McNeese, and John Yen. 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*, pages 945–952, 2005.

[90] Xiaocong Fan and John Yen. 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, 2010.

[91] Franz Faul, Edgar Erdfelder, Albert-Georg Lang, and Axel Buchner. G* power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior research methods*, 39(2):175–191, 2007.

[92] Rosemarie Fernandez, Sachita Shah, Elizabeth D Rosenman, Steve WJ Kozlowski, Sarah Henrickson Parker, and James A Grand. Developing team cognition: a role for simulation. *Simulation in healthcare: journal of the Society for Simulation in Healthcare*, 12(2):96, 2017.

[93] Stephen M Fiore, Eduardo Salas, and Janis A Cannon-Bowers. Group dynamics and shared mental model development. In *How people evaluate others in organizations*, pages 335–362. Psychology Press, 2001.

[94] Stephen M Fiore, Eduardo Salas, Haydee M Cuevas, and Clint A Bowers. Distributed coordination space: toward a theory of distributed team process and performance. *Theoretical Issues in Ergonomics Science*, 4(3-4):340–364, 2003.

[95] Paul M Fitts. Human engineering for an effective air-navigation and traffic-control system. 1951.

[96] Christopher Flathmann, Nathan McNeese, and Lorenzo Barberis Canonico. Using human-agent teams to purposefully design multi-agent systems. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, volume 63, pages 1425–1429. SAGE Publications Sage CA: Los Angeles, CA, 2019.

[97] Christopher Flathmann, Beau G Schelble, and Nathan J McNeese. Fostering human-agent team leadership by leveraging human teaming principles. In *2021 IEEE 2nd International Conference on Human-Machine Systems (ICHMS)*, pages 1–6. IEEE, 2021.

[98] Christopher Flathmann, Beau G Schelble, Rui Zhang, and Nathan J McNeese. Modeling and guiding the creation of ethical human-ai teams. In *Proceedings of the 2021 AAAI/ACM Conference on AI, Ethics, and Society*, pages 469–479, 2021.

[99] Edwin A Fleishman and Stephen J Zaccaro. Toward a taxonomy of team performance functions. 1992.

[100] Jakob Foerster, Richard Y Chen, Maruan Al-Shedivat, Shimon Whiteson, Pieter Abbeel, and Igor Mordatch. Learning with opponent-learning awareness. In *Proceedings of the 17th International Conference on Autonomous Agents and MultiAgent Systems*, pages 122–130. International Foundation for Autonomous Agents and Multiagent Systems, 2018.

[101] US Air Force. Autonomous horizons. *Washington, DC: US Air Force Office of the Chief Scientist*, 2015.

[102] Guo Freeman and Donghee Yvette Wohn. Understanding esports team formation and coordination. *Computer supported cooperative work (CSCW)*, 28:95–126, 2019.

[103] Helen Gavin. Thematic analysis. *Understanding research methods and statistics in psychology*, pages 273–282, 2008.

[104] Darren Gergle, Robert E. Kraut, and Susan R. Fussell. Using visual information for grounding and awareness in collaborative tasks. *Human–Computer Interaction*, 28(1):1–39, 2013. Publisher: Taylor & Francis.

[105] Fabiola H Gerpott, Daniel Balliet, Simon Columbus, Catherine Molho, and Reinout E de Vries. How do people think about interdependence? a multidimensional model of subjective outcome interdependence. *Journal of personality and social psychology*, 115(4):716, 2018.

[106] Connie JG Gersick. Time and transition in work teams: Toward a new model of group development. *Academy of Management journal*, 31(1):9–41, 1988.

[107] Deborah L Gladstein. Groups in context: A model of task group effectiveness. *Administrative science quarterly*, pages 499–517, 1984.

[108] Gene V Glass, Percy D Peckham, and James R Sanders. Consequences of failure to meet assumptions underlying the fixed effects analyses of variance and covariance. *Review of educational research*, 42(3):237–288, 1972.

[109] Kate Goddard, Abdul Roudsari, and Jeremy C Wyatt. Automation bias: a systematic review of frequency, effect mediators, and mitigators. *Journal of the American Medical Informatics Association*, 19(1):121–127, 2012.

[110] Paul S Goodman, Elizabeth Ravlin, and Marshall Schminke. Understanding groups in organizations. *Research in organizational behavior*, 1987.

[111] Jamie C Gorman, Nancy J Cooke, and Polemnia G Amazeen. Training adaptive teams. *Human Factors*, 52(2):295–307, 2010.

[112] Jamie C Gorman, Nancy J Cooke, Harry K Pederson, and Janie A DeJoode. Coordinated awareness of situation by teams (cast): Measuring team situation awareness of a communication glitch. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, volume 49, pages 274–277. SAGE Publications Sage CA: Los Angeles, CA, 2005.

[113] Jamie C Gorman, David A Grimm, Ronald H Stevens, Trysha Galloway, Ann M Willemsen-Dunlap, and Donald J Halpin. Measuring real-time team cognition during team training. *Human factors*, 62(5):825–860, 2020.

[114] Lia M Gray, Gina Wong-Wylie, Gwen R Rempel, and Karen Cook. Expanding qualitative research interviewing strategies: Zoom video communications. *The Qualitative Report*, 25(5):1292–1301, 2020.

[115] Terri L Griffith and Margaret A Neale. 8. information processing in traditional, hybrid, and virtual teams: From nascent knowledge to transactive memory. *Research in organizational behavior*, 23:379–421, 2001.

[116] Tom Gross. Supporting effortless coordination: 25 years of awareness research. *Computer Supported Cooperative Work (CSCW)*, 22(4):425–474, 2013.

[117] Jonathan Grudin. Why cscw applications fail: problems in the design and evaluationof organizational interfaces. In *Proceedings of the 1988 ACM conference on Computer-supported cooperative work*, pages 85–93, 1988.

[118] Jonathan Grudin. Computer-supported cooperative work: History and focus. *Computer*, 27(5):19–26, 1994.

[119] Stephanie Guerlain, Greg A Jamieson, Peter Bullemer, and Ronald Blair. The mpc elucidator: A case study in the design for human-automation interaction. *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans*, 32(1):25–40, 2002.

[120] Carl Gutwin and Saul Greenberg. A descriptive framework of workspace awareness for real-time groupware. *Computer Supported Cooperative Work (CSCW)*, 11(3):411–446, 2002.

[121] Carl Gutwin and Saul Greenberg. The importance of awareness for team cognition in distributed collaboration. 2004.

[122] Carl Gutwin, Saul Greenberg, and Mark Roseman. Workspace awareness in real-time distributed groupware: Framework, widgets, and evaluation. In *People and Computers XI*, pages 281–298. Springer, 1996.

[123] Carl Gutwin, Mark Roseman, and Saul Greenberg. A usability study of awareness widgets in a shared workspace groupware system. In *Proceedings of the 1996 ACM conference on Computer supported cooperative work*, pages 258–267, 1996.

[124] Mark Guzdial. Human-centered computing: a new degree for licklider's world. *Communications of the ACM*, 56(5):32–34, 2013.

[125] JR Hackman. The design of work teams. in handbook of organizational behavior, ed. jw lorsch, 1987.

[126] Feyza Merve Haf? ızoğlu and Sandip Sen. The effects of past experience on trust in repeated human-agent teamwork. In *Proceedings of the 17th International Conference on Autonomous Agents and MultiAgent Systems*, pages 514–522, 2018.

[127] Feyza Merve Hafizoglu and Sandip Sen. Reputation Based Trust In Human-Agent Teamwork Without Explicit Coordination. In *Proceedings of the 6th International Conference on Human-Agent Interaction*, pages 238–245, 2018.

[128] Katherine Hamilton, Vincent Mancuso, Dev Minotra, Rachel Hoult, Susan Mohammed, Alissa Parr, Gaurav Dubey, Eric McMillan, and Michael McNeese. Using the neocities 3.1 simulation to study and measure team cognition. In *Proceedings of the human factors and ergonomics society annual meeting*, volume 54, pages 433–437. SAGE Publications Sage CA: Los Angeles, CA, 2010.

[129] Peter A Hancock, Deborah R Billings, Kristin E Schaefer, Jessie YC Chen, Ewart J De Visser, and Raja Parasuraman. A meta-analysis of factors affecting trust in human-robot interaction. *Human factors*, 53(5):517–527, 2011.

[130] Peter A Hancock, Richard J Jagacinski, Raja Parasuraman, Christopher D Wickens, Glenn F Wilson, and David B Kaber. Human-automation interaction research: Past, present, and future. *Ergonomics in Design*, 21(2):9–14, 2013.

[131] Nader Hanna and Deborah Richards. 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*, pages 1485–1486, 2014.

[132] Nader Hanna and Deborah Richards. 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, pages 1777–1778, 2015.

[133] Sogand Hasanzadeh, Behzad Esmaeili, and Michael D Dodd. Measuring construction workers' real-time situation awareness using mobile eye-tracking. In *Construction research congress 2016*, pages 2894–2904, 2016.

[134] Allyson I Hauptman, Beau G Schelble, and Nathan J McNeese. Adaptive autonomy as a means for implementing shared ethics in human-ai teams. 2021.

[135] Allyson I Hauptman, Beau G Schelble, Nathan J McNeese, and Kapil Chalil Madathil. Adapt and overcome: Perceptions of adaptive autonomous agents for human-ai teaming. *Computers in Human Behavior*, 138:107451, 2023.

[136] Yugo Hayashi and Kosuke Wakabayashi. Can ai become reliable source to support human decision making in a court scene? In *Companion of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing*, pages 195–198, 2017.

[137] Jun He, Brian S Butler, and William R King. Team cognition: Development and evolution in software project teams. *Journal of Management Information Systems*, 24(2):261–292, 2007.

[138] Matthias Heinrich, Franz Josef Grüneberger, Thomas Springer, Philipp Hauer, and Martin Gaedke. Gawi: A comprehensive workspace awareness library for collaborative web applications. In *International Conference on Web Engineering*, pages 482–485. Springer, 2013.

[139] D Benjamin Hellar and David L Hall. Neocities: an experimental test-bed for quantifying the effects of cognitive aids on team performance in c2 situations. In *Modeling and Simulation for Military Operations IV*, volume 7348, page 73480K. International Society for Optics and Photonics, 2009.

[140] D Benjamin Hellar and Michael McNeese. Neocities: A simulated command and control task environment for experimental research. In *Proceedings of the human factors and ergonomics society annual meeting*, volume 54, pages 1027–1031. SAGE Publications Sage CA: Los Angeles, CA, 2010.

[141] Kevin Anthony Hoff and Masooda Bashir. Trust in automation: Integrating empirical evidence on factors that influence trust. *Human factors*, 57(3):407–434, 2015.

[142] Sean D Holcomb, William K Porter, Shaun V Ault, Guifen Mao, and Jin Wang. Overview on deepmind and its alphago zero ai. In *Proceedings of the 2018 international conference on big data and education*, pages 67–71, 2018.

[143] Andrea B Hollingshead. Communication, learning, and retrieval in transactive memory systems. *Journal of experimental social psychology*, 34(5):423–442, 1998.

[144] Christian Homburg, Harley Krohmer, and John P Workman. Jr. Strategic consensus and performance: the role of strategy type and market-related dynamism. *Strategic management journal*, 20(4):339–357, 1999.

[145] Christopher Honts, Matthew Prewett, John Rahael, and Michael Grossenbacher. The importance of team processes for different team types. *Team performance management: an international journal*, 18(5/6):312–327, 2012.

[146] Li-tze Hu and Peter M Bentler. Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural equation modeling: a multidisciplinary journal*, 6(1):1–55, 1999.

[147] William Ickes and Richard Gonzalez. " social" cognition and social cognition: From the subjective to the intersubjective. *Small group research*, 25(2):294–315, 1994.

[148] Daniel R Ilgen, John R Hollenbeck, Michael Johnson, and Dustin Jundt. Teams in organizations. *Annual review of psychology*, 56:517–543, 2005.

[149] Florian Jentsch, John Barnett, Clint A Bowers, and Eduardo Salas. Who is flying this plane anyway? what mishaps tell us about crew member role assignment and air crew situation awareness. *Human Factors*, 41(1):1–14, 1999.

[150] Leila J Johannesen, Richard I Cook, and David D Woods. Cooperative communications in dynamic fault management. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, volume 38, pages 225–229. SAGE Publications Sage CA: Los Angeles, CA, 1994.

[151] Douglas H Johnson. The insignificance of statistical significance testing. *The journal of wildlife management*, pages 763–772, 1999.

[152] Philip Nicholas Johnson-Laird. *Mental models: Towards a cognitive science of language, inference, and consciousness*. Number 6. Harvard University Press, 1983.

[153] Brennan Jones, Anthony Tang, and Carman Neustaedter. Remote communication in wilderness search and rescue: implications for the design of emergency distributed-collaboration tools for network-sparse environments. *Proceedings of the ACM on human-computer interaction*, 4(GROUP):1–26, 2020.

[154] Rashaad ET Jones, Michael D McNeese, Erik S Connors, Tyrone Jefferson Jr, and David L Hall Jr. A distributed cognition simulation involving homeland security and defense: The development of neocities. In *Proceedings of the human factors and ergonomics society annual meeting*, volume 48, pages 631–634. SAGE Publications Sage CA: Los Angeles, CA, 2004.

[155] David B Kaber and Mica R Endsley. Team situation awareness for process control safety and performance. *Process Safety Progress*, 17(1):43–48, 1998.

[156] Prasert Kanawattanachai and Youngjin Yoo. The impact of knowledge coordination on virtual team performance over time. *MIS quarterly*, pages 783–808, 2007.

[157] Steven J Kass, Kerstan S Cole, and Claudia J Stanny. Effects of distraction and experience on situation awareness and simulated driving. *Transportation Research Part F: Traffic Psychology and Behaviour*, 10(4):321–329, 2007.

[158] Franz W Kellermanns, Jorge Walter, Christoph Lechner, and Steven W Floyd. The lack of consensus about strategic consensus: Advancing theory and research. *Journal of Management*, 31(5):719–737, 2005.

[159] Janice R Kelly and Sigal G Barsade. Mood and emotions in small groups and work teams. *Organizational behavior and human decision processes*, 86(1):99–130, 2001.

[160] Deanna M Kennedy and Sara A McComb. When teams shift among processes: Insights from simulation and optimization. *Journal of Applied Psychology*, 99(5):784, 2014.

[161] Martin Kilduff, Reinhard Angelmar, and Ajay Mehra. Top management-team diversity and firm performance: Examining the role of cognitions. *Organization science*, 11(1):21–34, 2000.

[162] Young Ji Kim, David Engel, Anita Williams Woolley, Jeffrey Yu-Ting Lin, Naomi McArthur, and Thomas W Malone. What makes a strong team? using collective intelligence to predict team performance in league of legends. In *Proceedings of the 2017 ACM conference on computer supported cooperative work and social computing*, pages 2316–2329, 2017.

[163] Richard Klimoski and Susan Mohammed. Team mental model: Construct or metaphor? *Journal of management*, 20(2):403–437, 1994.

[164] Bart P Knijnenburg, Niels JM Reijmer, and Martijn C Willemsen. Each to his own: how different users call for different interaction methods in recommender systems. In *Proceedings of the fifth ACM conference on Recommender systems*, pages 141–148, 2011.

[165] Bart P Knijnenburg and Martijn C Willemsen. Evaluating recommender systems with user experiments. *Recommender systems handbook*, pages 309–352, 2015.

[166] Steve WJ Kozlowski and Daniel R Ilgen. Enhancing the effectiveness of work groups and teams. *Psychological science in the public interest*, 7(3):77–124, 2006.

[167] Kurt Kraiger and Lucy H Wenzel. Conceptual development and empirical evaluation of measures of shared mental models as indicators of team effectiveness. *Team performance assessment and measurement: Theory, methods, and applications*, 63:84, 1997.

[168] Thomas Kude, Sunil Mithas, Christoph T Schmidt, and Armin Heinzl. How pair programming influences team performance: The role of backup behavior, shared mental models, and task novelty. *Information Systems Research*, 30(4):1145–1163, 2019.

[169] Janice Langan-Fox, Sharon Code, and Kim Langfield-Smith. Team mental models: Techniques, methods, and analytic approaches. *Human factors*, 42(2):242–271, 2000.

[170] Alex Leavitt, Brian C Keegan, and Joshua Clark. Ping to win? non-verbal communication and team performance in competitive online multiplayer games. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, pages 4337–4350, 2016.

[171] John Lee and Neville Moray. Trust, control strategies and allocation of function in human-machine systems. *Ergonomics*, 35(10):1243–1270, 1992.

[172] John D Lee and Neville Moray. Trust, self-confidence, and operators' adaptation to automation. *International journal of human-computer studies*, 40(1):153–184, 1994.

[173] John D Lee and Katrina A See. Trust in automation: Designing for appropriate reliance. *Human factors*, 46(1):50–80, 2004.

[174] Miyoung Lee and Tristan E Johnson. Understanding the effects of team cognition associated with complex engineering tasks: Dynamics of shared mental models, task-smm, and team-smm. *Performance Improvement Quarterly*, 21(3):73–95, 2008.

[175] Herman B Leonard and Arnold M Howitt. Katrina as prelude: Preparing for and responding to katrina-class disturbances in the united states–testimony to us senate committee, march 8, 2006. *Journal of Homeland Security and Emergency Management*, 3(2), 2006.

[176] Jeffery A LePine, Ronald F Piccolo, Christine L Jackson, John E Mathieu, and Jessica R Saul. A meta-analysis of teamwork processes: tests of a multidimensional model and relationships with team effectiveness criteria. *Personnel psychology*, 61(2):273–307, 2008.

[177] Kyle Lewis. Measuring transactive memory systems in the field: scale development and validation. *Journal of applied psychology*, 88(4):587, 2003.

[178] Mengyao Li and John D Lee. Modeling goal alignment in human-ai teaming: A dynamic game theory approach. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, volume 66, pages 1538–1542. SAGE Publications Sage CA: Los Angeles, CA, 2022.

[179] Tianyi Li, Mihaela Vorvoreanu, Derek DeBellis, and Saleema Amershi. Assessing human-ai interaction early through factorial surveys: A study on the guidelines for human-ai interaction. *ACM Transactions on Computer-Human Interaction*, 2022.

[180] Claire Liang, Julia Proft, Erik Andersen, and Ross A Knepper. Implicit communication of actionable information in human-ai teams. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, pages 1–13, 2019.

[181] Shu-Hsien Liao. Expert system methodologies and applications—a decade review from 1995 to 2004. *Expert systems with applications*, 28(1):93–103, 2005.

[182] Beng-Chong Lim and Katherine J Klein. Team mental models and team performance: A field study of the effects of team mental model similarity and accuracy. *Journal of Organizational Behavior: The International Journal of Industrial, Occupational and Organizational Psychology and Behavior*, 27(4):403–418, 2006.

[183] Jessica Lipnack and Jeffrey Stamps. Virtual teams: The new way to work. *Strategy & Leadership*, 1999.

[184] Jeremy Lopez and Richard Pak. Does one bad machine spoil the bunch?: A review of trust in multiple-component systems. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, volume 64, pages 1546–1550. SAGE Publications Sage CA: Los Angeles, CA, 2020.

[185] Jeremy Lopez, Heather Watkins, and Richard Pak. Enhancing component-specific trust with consumer automated systems through humanness design. *Ergonomics*, 66(2):291–302, 2023.

[186] Fabrice Lumineau. How contracts influence trust and distrust. *Journal of management*, 43(5):1553–1577, 2017.

[187] Kevin J Lynch, Joel M Snyder, Douglas R Vogel, and William K McHenry. The arizona analyst information system: Supporting collaborative research on international technological trends. In *Proceedings of the IFIP WG 8.4 confernece on Multi-user interfaces and applications*, pages 159–174. 1990.

[188] Joseph B Lyons and Svyatoslav Y Guznov. Individual differences in human–machine trust: A multi-study look at the perfect automation schema. *Theoretical Issues in Ergonomics Science*, 20(4):440–458, 2019.

[189] Joseph B Lyons, Nhut T Ho, William E Fergueson, Garrett G Sadler, Samantha D Cals, Casey E Richardson, and Mark A Wilkins. Trust of an automatic ground collision avoidance technology: A fighter pilot perspective. *Military Psychology*, 28(4):271–277, 2016.

[190] Joseph B Lyons, Kevin T Wynne, Sean Mahoney, and Mark A Roebke. Trust and human-machine teaming: A qualitative study. In *Artificial intelligence for the internet of everything*, pages 101–116. Elsevier, 2019.

[191] Ruiqi Ma and David B Kaber. Situation awareness and driving performance in a simulated navigation task. *Ergonomics*, 50(8):1351–1364, 2007.

[192] Merce Mach, Simon Dolan, and Shay Tzafrir. The differential effect of team members' trust on team performance: The mediation role of team cohesion. *Journal of Occupational and Organizational Psychology*, 83(3):771–794, 2010.

[193] Jean MacMillan, Elliot E Entin, and Daniel Serfaty. Communication overhead: The hidden cost of team cognition. 2004.

[194] Jane T Malin, Debra L Schreckenghost, David D Woods, Scott S Potter, Leila Johannesen, Matthew Holloway, and Kenneth D Forbus. Making intelligent systems team players: Case studies and design issues. volume 1: Human-computer interaction design. 1991.

[195] Michelle A Marks, John E Mathieu, and Stephen J Zaccaro. A temporally based framework and taxonomy of team processes. *Academy of management review*, 26(3):356–376, 2001.

[196] Michelle A Marks and Frederick J Panzer. The influence of team monitoring on team processes and performance. *Human Performance*, 17(1):25–41, 2004.

[197] Michelle A Marks, Mark J Sabella, C Shawn Burke, and Stephen J Zaccaro. The impact of cross-training on team effectiveness. *Journal of Applied Psychology*, 87(1):3, 2002.

[198] Shannon L Marlow, Christina N Lacerenza, and Eduardo Salas. Communication in virtual teams: A conceptual framework and research agenda. *Human Resource Management Review*, 27(4):575–589, 2017.

[199] Luis L Martins, Lucy L Gilson, and M Travis Maynard. Virtual teams: What do we know and where do we go from here? *Journal of management*, 30(6):805–835, 2004.

[200] John Mathieu, M Travis Maynard, Tammy Rapp, and Lucy Gilson. Team effectiveness 1997-2007: A review of recent advancements and a glimpse into the future. *Journal of management*, 34(3):410–476, 2008.

[201] John E Mathieu, Tonia S Heffner, Gerald F Goodwin, Janis A Cannon-Bowers, and Eduardo Salas. Scaling the quality of teammates' mental models: Equifinality and normative comparisons. *Journal of Organizational Behavior: The International Journal of Industrial, Occupational and Organizational Psychology and Behavior*, 26(1):37–56, 2005.

[202] John E Mathieu, Tonia S Heffner, Gerald F Goodwin, Eduardo Salas, and Janis A Cannon-Bowers. The influence of shared mental models on team process and performance. *Journal of applied psychology*, 85(2):273, 2000.

[203] David Maulsby, Saul Greenberg, and Richard Mander. Prototyping an intelligent agent through wizard of oz. In *Proceedings of the INTERACT'93 and CHI'93 conference on Human factors in computing systems*, pages 277–284, 1993.

[204] Roger C Mayer, James H Davis, and F David Schoorman. An integrative model of organizational trust. *Academy of management review*, 20(3):709–734, 1995.

[205] Joseph E McGrath, Holly Arrow, and Jennifer L Berdahl. The study of groups: Past, present, and future. In *Personality and social psychology review*, pages 95–105. Psychology Press, 2000.

[206] Joseph Edward McGrath. *Social psychology: A brief introduction*. Holt, Rinehart and Winston, 1964.

[207] Joseph Edward McGrath. *Groups: Interaction and performance*, volume 14. Prentice-Hall Englewood Cliffs, NJ, 1984.

[208] Michael D McNeese, Vincent F Mancuso, Nathan J McNeese, Tristan Endsley, and Pete Forster. An integrative simulation to study team cognition in emergency crisis management. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, volume 58, pages 285–289. SAGE Publications Sage CA: Los Angeles, CA, 2014.

[209] Nathan McNeese, Nancy J Cooke, Mike Fedele, and Rob Gray. Perspectives on team cognition and team sports. In *Sport and Exercise Psychology Research*, pages 123–141. Elsevier, 2016.

[210] Nathan McNeese, Mustafa Demir, Erin Chiou, Nancy Cooke, and Giovanni Yanikian. Understanding the role of trust in human-autonomy teaming. In *Proceedings of the 52nd Hawaii international conference on system sciences*, 2019.

[211] Nathan J McNeese, Mustafa Demir, Erin K Chiou, and Nancy J Cooke. Trust and team performance in human–autonomy teaming. *International Journal of Electronic Commerce*, 25(1):51–72, 2021.

[212] Nathan J McNeese, Mustafa Demir, Nancy J Cooke, and Christopher Myers. Teaming with a synthetic teammate: Insights into human-autonomy teaming. *Human factors*, 60(2):262–273, 2018.

[213] Nathan J McNeese, Mustafa Demir, Nancy J Cooke, and Manrong She. Team situation awareness and conflict: A study of human–machine teaming. *Journal of Cognitive Engineering and Decision Making*, 15(2-3):83–96, 2021.

[214] Nathan J McNeese, Christopher Flathmann, Thomas A O'Neill, and Eduardo Salas. Stepping out of the shadow of human-human teaming: Crafting a unique identity for human-autonomy teams. *Computers in Human Behavior*, 148:107874, 2023.

[215] Nathan J. McNeese, Beau G. Schelble, Lorenzo Barberis Canonico, and Mustafa Demir. Who/what is my teammate? team composition considerations in human–ai teaming. *IEEE Transactions on Human-Machine Systems*, 51(4):288–299, 2021.

[216] Joseph E Mercado, Michael A Rupp, Jessie YC Chen, Michael J Barnes, Daniel Barber, and Katelyn Procci. Intelligent agent transparency in human–agent teaming for multi-uxv management. *Human Factors*, 58(3):401–415, 2016.

[217] Tim R Merritt, Kian Boon Tan, Christopher Ong, Aswin Thomas, Teong Leong Chuah, and Kevin McGee. Are artificial team-mates scapegoats in computer games. In *Proceedings of the ACM 2011 conference on Computer supported cooperative work*, pages 685–688, 2011.

[218] Jessica Mesmer-Magnus, Ashley A Niler, Gabriel Plummer, Lindsay E Larson, and Leslie A DeChurch. The cognitive underpinnings of effective teamwork: A continuation. *Career Development International*, 2017.

[219] Michael D Michalisin, Steven J Karau, and Charnchai Tangpong. The effects of performance and team cohesion on attribution: A longitudinal simulation. *Journal of Business Research*, 57(10):1108–1115, 2004.

[220] Christopher A Miller, Harry Funk, Robert Goldman, John Meisner, and Peggy Wu. Implications of adaptive vs. adaptable uis on decision making: Why "automated adaptiveness" is not always the right answer. In *Proceedings of the 1st international conference on augmented cognition*, pages 22–27, 2005.

[221] Susan Mohammed, Lori Ferzandi, and Katherine Hamilton. Metaphor no more: A 15-year review of the team mental model construct. *Journal of management*, 36(4):876–910, 2010.

[222] Susan Mohammed, Richard Klimoski, and Joan R Rentsch. The measurement of team mental models: We have no shared schema. *Organizational Research Methods*, 3(2):123–165, 2000.

[223] Bonnie M Muir. Trust between humans and machines, and the design of decision aids. *International journal of man-machine studies*, 27(5-6):527–539, 1987.

[224] Geoff Musick, Thomas A O'Neill, Beau G Schelble, Nathan J McNeese, and Jonn B Henke. What happens when humans believe their teammate is an ai? an investigation into humans teaming with autonomy. *Computers in Human Behavior*, 122:106852, 2021.

[225] Geoff Musick, Rui Zhang, Nathan J McNeese, Guo Freeman, and Anurata Prabha Hridi. Leveling up teamwork in esports: Understanding team cognition in a dynamic virtual environment. *Proceedings of the ACM on Human-Computer Interaction*, 5(CSCW1):1–30, 2021.

[226] Clifford Nass, BJ Fogg, and Youngme Moon. Can computers be teammates? *International Journal of Human-Computer Studies*, 45(6):669–678, 1996.

[227] Engineering National Academies of Sciences and Medicine. Human-AI teaming: State of the art and research needs, 2021.

[228] Stefanos Nikolaidis, Przemyslaw Lasota, Ramya Ramakrishnan, and Julie Shah. 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, 2015.

[229] Stefanos Nikolaidis and Julie Shah. Human-robot cross-training: computational formulation, modeling and evaluation of a human team training strategy. In *2013 8th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, pages 33–40. IEEE, 2013.

[230] Ashley A Niler, Jessica R Mesmer-Magnus, Lindsay E Larson, Gabriel Plummer, Leslie A DeChurch, and Noshir S Contractor. Conditioning team cognition: A meta-analysis. *Organizational Psychology Review*, 11(2):144–174, 2021.

[231] Niina Nurmi and Pamela J Hinds. Job complexity and learning opportunities: A silver lining in the design of global virtual work. *Journal of International Business Studies*, 47(6):631–654, 2016.

[232] Debra L O'Connor, Tristan E Johnson, and Mohammed K Khalil. Measuring team cognition: Concept mapping elicitation as a means of constructing team shared mental models in an applied setting. 2004.

[233] Kenya Freeman Oduor and Eric N Wiebe. The effects of automated decision algorithm modality and transparency on reported trust and task performance. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, volume 52, pages 302–306. SAGE Publications Sage CA: Los Angeles, CA, 2008.

[234] Gary M Olson and Judith S Olson. Distance matters. *Human–computer interaction*, 15(2-3):139–178, 2000.

[235] Judith S Olson, Stephanie Teasley, Lisa Covi, and Gary Olson. The (currently) unique advantages of collocated. *Distributed work*, page 113, 2002.

[236] Christopher Ong, Kevin McGee, and Teong Leong Chuah. Closing the human-ai team-mate gap: how changes to displayed information impact player behavior towards computer teammates. In *Proceedings of the 24th Australian Computer-Human Interaction Conference*, pages 433–439, 2012.

[237] Linda Onnasch, Christopher D Wickens, Huiyang Li, and Dietrich Manzey. Human performance consequences of stages and levels of automation: An integrated meta-analysis. *Human factors*, 56(3):476–488, 2014.

[238] Thomas O'Neill, Nathan McNeese, Amy Barron, and Beau Schelble. Human–autonomy teaming: A review and analysis of the empirical literature. *Human Factors*, page 0018720820960865, 2020.

[239] Matthew Pacailler, Sarah Yahoodik, Tetsuya Sato, Jeremiah G Ammons, and Jeremiah Still. Human-centered artificial intelligence: Beyond a two-dimensional framework. In *International Conference on Human-Computer Interaction*, pages 471–482. Springer, 2022.

[240] Rohan Paleja, Muyleng Ghuy, Nadun Ranawaka Arachchige, Reed Jensen, and Matthew Gombolay. The utility of explainable ai in ad hoc human-machine teaming. *Advances in Neural Information Processing Systems*, 34:610–623, 2021.

[241] Laura J Pape. *The effect of personality characteristics on team member schema similarity.* PhD thesis, Wright State University, 1998.

[242] Raja Parasuraman and Victor Riley. Humans and automation: Use, misuse, disuse, abuse. *Human factors*, 39(2):230–253, 1997.

[243] Raja Parasuraman, Thomas B Sheridan, and Christopher D Wickens. A model for types and levels of human interaction with automation. *IEEE Transactions on systems, man, and cybernetics-Part A: Systems and Humans*, 30(3):286–297, 2000.

[244] Hee Sun Park. The effects of shared cognition on group satisfaction and performance: Politeness and efficiency in group interaction. *Communication Research*, 35(1):88–108, 2008.

[245] Paul Paulus. Groups, teams, and creativity: The creative potential of idea-generating groups. *Applied psychology*, 49(2):237–262, 2000.

[246] R Pew. An introduction to the concept of situation awareness. *Situational awareness in complex systems*, pages 17–23, 1994.

[247] Mark S Pfaff and Michael D McNeese. Effects of mood and stress on distributed team cognition. *Theoretical Issues in Ergonomics Science*, 11(4):321–339, 2010.

[248] Anne Powell, Gabriele Piccoli, and Blake Ives. Virtual teams: a review of current literature and directions for future research. *ACM SIGMIS Database: the DATABASE for Advances in Information Systems*, 35(1):6–36, 2004.

[249] Jane S Prichard and Melanie J Ashleigh. The effects of team-skills training on transactive memory and performance. *Small group research*, 38(6):696–726, 2007.

[250] Carolyn Prince and Eduardo Salas. Training and research for teamwork in the military aircrew. 1993.

[251] Gregory E Prussia and Angelo J Kinicki. A motivational investigation of group effectiveness using social-cognitive theory. *Journal of Applied Psychology*, 81(2):187, 1996.

[252] Martin Ragot, Nicolas Martin, and Salomé Cojean. Ai-generated vs. human artworks. a perception bias towards artificial intelligence? In *Extended abstracts of the 2020 CHI conference on human factors in computing systems*, pages 1–10, 2020.

[253] Molly Inhofe Rapert, Anne Velliquette, and Judith A Garretson. The strategic implementation process: evoking strategic consensus through communication. *Journal of business research*, 55(4):301–310, 2002.

[254] Devaki Rau. The influence of relationship conflict and trust on the transactive memory: Performance relation in top management teams. *Small group research*, 36(6):746–771, 2005.

[255] Richard E Redding, John R Cannon, and Thomas L Seamster. Expertise in air traffic control (atc): What is it, and how can we train for it? In *Proceedings of the Human Factors Society Annual Meeting*, volume 36, pages 1326–1330. SAGE Publications Sage CA: Los Angeles, CA, 1992.

[256] Roberto Refinetti. Demonstrating the consequences of violations of assumptions in between-subjects analysis of variance. *Teaching of Psychology*, 23(1):51–54, 1996.

[257] Joan Rentsch, Michael D McNeese, Laura J Pape, Dawn D Burnett, and Darcy M Menard. Testing the effects of team processes on team member schema similarity and team performance: Examination of the team member schema similarity model. Technical report, WRIGHT STATE UNIV DAYTON OH DEPT OF PSYCHOLOGY, 1998.

[258] Joan R Rentsch and Rosalie J Hall. Members of great teams think alike: A model of team effectiveness and schema similarity among team members. 1994.

[259] Joan R Rentsch and Richard J Klimoski. Why do 'great minds' think alike?: Antecedents of team member schema agreement. *Journal of Organizational Behavior: The International Journal of Industrial, Occupational and Organizational Psychology and Behavior*, 22(2):107–120, 2001.

[260] Christian J Resick, Marcus W Dickson, Jacqueline K Mitchelson, Leslie K Allison, and Malissa A Clark. Team composition, cognition, and effectiveness: Examining mental model similarity and accuracy. *Group Dynamics: Theory, Research, and Practice*, 14(2):174, 2010.

[261] Victor Riley. A general model of mixed-initiative human-machine systems. In *Proceedings of the Human Factors Society Annual Meeting*, volume 33, pages 124–128. Sage Publications Sage CA: Los Angeles, CA, 1989.

[262] Mark Roseman and Saul Greenberg. Building real-time groupware with groupkit, a groupware toolkit. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 3(1):66–106, 1996.

[263] William B Rouse and Nancy M Morris. On looking into the black box: Prospects and limits in the search for mental models. *Psychological bulletin*, 100(3):349, 1986.

[264] Anna L Rowe and Nancy J Cooke. Measuring mental models: Choosing the right tools for the job. *Human resource development quarterly*, 6(3):243–255, 1995.

[265] Simone Rozzi and Paola Amaldi. Organizational and inter-organizational precursors to problematic automation in safety critical domains. In *Proceedings of the 2nd International Conference on Application and Theory of Automation in Command and Control Systems*, pages 98–106, 2012.

[266] Richard Saavedra, P Christopher Earley, and Linn Van Dyne. Complex interdependence in task-performing groups. *Journal of applied psychology*, 78(1):61, 1993.

[267] Parya Saberi. Research in the time of coronavirus: continuing ongoing studies in the midst of the covid-19 pandemic. *AIDS and Behavior*, 24(8):2232–2235, 2020.

[268] Eduardo Salas, Nancy J Cooke, and Michael A Rosen. On teams, teamwork, and team performance: Discoveries and developments. *Human factors*, 50(3):540–547, 2008.

[269] Eduardo Salas, Terry L Dickinson, Sharolyn A Converse, and Scott I Tannenbaum. Toward an understanding of team performance and training. 1992.

[270] Eduardo Salas, Carolyn Prince, David P Baker, and Lisa Shrestha. Situation awareness in team performance: Implications for measurement and training. *Human Factors*, 37(1):123–136, 1995.

[271] Lelyn D Saner, Cheryl A Bolstad, Cleotilde Gonzalez, and Haydee M Cuevas. Measuring and predicting shared situation awareness in teams. *Journal of cognitive engineering and decision making*, 3(3):280–308, 2009.

[272] Catarina Marques Santos, Sjir Uitdewilligen, and Ana Margarida Passos. A temporal common ground for learning: The moderating effect of shared mental models on the relation between team learning behaviours and performance improvement. *European Journal of Work and Organizational Psychology*, 24(5):710–725, 2015.

[273] Paul M Satchell. *Cockpit monitoring and alerting systems*. Routledge, 1993.

[274] Kristin E Schaefer, Jessie YC Chen, James L Szalma, and Peter A Hancock. A meta-analysis of factors influencing the development of trust in automation: Implications for understanding autonomy in future systems. *Human factors*, 58(3):377–400, 2016.

[275] Kristin E Schaefer, Jean Oh, Derya Aksaray, and Daniel Barber. Integrating context into artificial intelligence: research from the robotics collaborative technology alliance. *Ai Magazine*, 40(3):28–40, 2019.

[276] Beau Schelble, Lorenzo-Barberis Canonico, Nathan McNeese, Jack Carroll, and Casey Hird. Designing human-autonomy teaming experiments through reinforcement learning. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, volume 64, pages 1426–1430. SAGE Publications Sage CA: Los Angeles, CA, 2020.

[277] Beau G Schelble, Christopher Flathmann, and Nathan McNeese. Towards meaningfully integrating human-autonomy teaming in applied settings. In *Proceedings of the 8th International Conference on Human-Agent Interaction*, pages 149–156, 2020.

[278] Beau G Schelble, Christopher Flathmann, Nathan J McNeese, Guo Freeman, and Rohit Mallick. 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, 2022.

[279] Beau G Schelble, Christopher Flathmann, Nathan J McNeese, Thomas O'Neill, Richard Pak, and Moses Namara. Investigating the effects of perceived teammate artificiality on human performance and cognition. *International Journal of Human–Computer Interaction*, pages 1–16, 2022.

[280] Beau G Schelble, Jeremy Lopez, Claire Textor, Rui Zhang, Nathan J McNeese, Richard Pak, and Guo Freeman. Towards ethical ai: Empirically investigating dimensions of ai ethics, trust repair, and performance in human-ai teaming. *Human Factors*, page 00187208221116952, 2022.

[281] Matthias Scheutz, Scott A DeLoach, and Julie A Adams. A framework for developing and using shared mental models in human-agent teams. *Journal of Cognitive Engineering and Decision Making*, 11(3):203–224, 2017.

[282] Kjeld Schmidt and Liam Bannon. Constructing cscw: The first quarter century. *Computer supported cooperative work (CSCW)*, 22(4):345–372, 2013.

[283] Tjeerd AJ Schoonderwoerd, Emma M van Zoelen, Karel van den Bosch, and Mark A Neerincx. Design patterns for human-ai co-learning: A wizard-of-oz evaluation in an urban-search-and-rescue task. *International Journal of Human-Computer Studies*, 164:102831, 2022.

[284] Amber N Schroeder, Patrick J Rosopa, Julia H Whitaker, Ian N Fairbanks, and Phoebe Xoxakos. Heteroscedasticity in organizational research. *Research Methods in Human Research Management: Toward Valid Research-Based Inferences*, page 67, 2020.

[285] John Schulman, Filip Wolski, Prafulla Dhariwal, Alec Radford, and Oleg Klimov. Proximal policy optimization algorithms. *arXiv preprint arXiv:1707.06347*, 2017.

[286] Roger W Schvaneveldt. *Pathfinder associative networks: Studies in knowledge organization.* Ablex Publishing, 1990.

[287] Douglas Schwartz. Training for situational awareness. In *Proceedings of the 5th International Symposium on Aviation Psychology*, volume 46, pages 18–20. International Civil Aviation Organization, 1992.

[288] Julie Shah, Been Kim, and Stefanos Nikolaidis. Human-inspired techniques for human-machine team planning. In *2012 AAAI Fall Symposium Series*, 2012.

[289] Thomas B Sheridan and Raja Parasuraman. Human vs. automation in responding to failures: An expected-value analysis. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, volume 44, pages 1–4. SAGE Publications Sage CA: Los Angeles, CA, 2000.

[290] Donghee Shin. The effects of explainability and causability on perception, trust, and acceptance: Implications for explainable ai. *International Journal of Human-Computer Studies*, 146:102551, 2021.

[291] Ben Shneiderman. Bridging the gap between ethics and practice: guidelines for reliable, safe, and trustworthy human-centered ai systems. *ACM Transactions on Interactive Intelligent Systems (TiiS)*, 10(4):1–31, 2020.

[292] Ben Shneiderman. Human-centered artificial intelligence: Reliable, safe & trustworthy. *International Journal of Human–Computer Interaction*, 36(6):495–504, 2020.

[293] Ben Shneiderman. Human-centered artificial intelligence: Three fresh ideas. *AIS Transactions on Human-Computer Interaction*, 12(3):109–124, 2020.

[294] Indramani L Singh, Robert Molloy, and Raja Parasuraman. Individual differences in monitoring failures of automation. *The Journal of general psychology*, 120(3):357–373, 1993.

[295] Mark K Singley, Moninder Singh, Peter Fairweather, Robert Farrell, and Steven Swerling. Algebra jam: supporting teamwork and managing roles in a collaborative learning environment. In *Proceedings of the 2000 ACM conference on Computer supported cooperative work*, pages 145–154, 2000.

[296] JA Smith and M Osborn. Chapter 4: Interpretive phenomenological analysis. *Qualitative psychology: A practical guide to methods*, pages 53–80, 2003.

[297] Robert Stalnaker. Common ground. *Linguistics and philosophy*, 25(5/6):701–721, 2002.

[298] Ivan Dale Steiner. *Group process and productivity*. Academic press New York, 1972.

[299] Jan E Stets and Peter J Burke. Identity theory and social identity theory. *Social psychology quarterly*, pages 224–237, 2000.

[300] Renée J Stout, Janis A Cannon-Bowers, Eduardo Salas, and Dana M Milanovich. Planning, shared mental models, and coordinated performance: An empirical link is established. *Human factors*, 41(1):61–71, 1999.

[301] Anselm L Strauss. *Qualitative analysis for social scientists*. Cambridge university press, 1987.

[302] Aaquib Tabrez, Matthew B Luebbers, and Bradley Hayes. A survey of mental modeling techniques in human–robot teaming. *Current Robotics Reports*, 1(4):259–267, 2020.

[303] Richard M Taylor. Situational awareness rating technique (sart): The development of a tool for aircrew systems design. In *Situational awareness*, pages 111–128. Routledge, 1990.

[304] Richard M Taylor. Situational awareness rating technique (sart): The development of a tool for aircrew systems design. In *Situational awareness*, pages 111–128. Routledge, 2017.

[305] Nathan L Tenhundfeld, Ewart J de Visser, Anthony J Ries, Victor S Finomore, and Chad C Tossell. Trust and distrust of automated parking in a tesla model x. *Human factors*, 62(2):194–210, 2020.

[306] Claire Textor, Rui Zhang, Jeremy Lopez, Beau G Schelble, Nathan J McNeese, Guo Freeman, Richard Pak, Chad Tossell, and Ewart J de Visser. Exploring the relationship between ethics and trust in human–artificial intelligence teaming: A mixed methods approach. *Journal of cognitive engineering and decision making*, 16(4):252–281, 2022.

[307] Zachary O Toups, Andruid Kerne, William Hamilton, and Alan Blevins. Emergent team coordination: From fire emergency response practice to a non-mimetic simulation game. In *Proceedings of the ACM 2009 international conference on Supporting group work*, pages 341–350, 2009.

[308] Anthony M Townsend, Samuel M DeMarie, and Anthony R Hendrickson. Virtual teams: Technology and the workplace of the future. *Academy of Management Perspectives*, 12(3):17–29, 1998.

[309] Jeroen Vaes, Maria Paola Paladino, Luigi Castelli, Jacques-Philippe Leyens, and Anna Giovanazzi. On the behavioral consequences of infrahumanization: The implicit role of uniquely human emotions in intergroup relations. *Journal of personality and social psychology*, 85(6):1016, 2003.

[310] Jandre J Van Rensburg, Catarina M Santos, Simon B de Jong, and Sjir Uitdewilligen. The five-factor perceived shared mental model scale: a consolidation of items across the contemporary literature. *Frontiers in Psychology*, 12, 2021.

[311] Christiane Vandenplas-Holper. Nursery school teacher's control beliefs, beliefs about development and education, and educational action. *Journal of Applied Developmental Psychology*, 17(3):453–473, 1996.

[312] Diane Vaughan. *The Challenger launch decision: Risky technology, culture, and deviance at NASA*. University of Chicago press, 1996.

[313] George Veletsianos. Cognitive and affective benefits of an animated pedagogical agent: Considering contextual relevance and aesthetics. *Journal of Educational Computing Research*, 36(4):373–377, 2007.

[314] Catherine E Volpe, Janis A Cannon-Bowers, Eduardo Salas, and Paul E Spector. The impact of cross-training on team functioning: An empirical investigation. *Human factors*, 38(1):87–100, 1996.

[315] James C Walliser, Patrick R Mead, and Tyler H Shaw. The perception of teamwork with an autonomous agent enhances affect and performance outcomes. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, volume 61, pages 231–235. SAGE Publications Sage CA: Los Angeles, CA, 2017.

[316] Dayong Wang, Aditya Khosla, Rishab Gargeya, Humayun Irshad, and Andrew H Beck. Deep learning for identifying metastatic breast cancer. *arXiv preprint arXiv:1606.05718*, 2016.

[317] Kelly D Wason, Michael J Polonsky, and Michael R Hyman. Designing vignette studies in marketing. *Australasian Marketing Journal*, 10(3):41–58, 2002.

[318] Daniel M Wegner. Transactive memory: A contemporary analysis of the group mind. In *Theories of group behavior*, pages 185–208. Springer, 1987.

[319] Daniel M Wegner, Toni Giuliano, and Paula T Hertel. Cognitive interdependence in close relationships. In *Compatible and incompatible relationships*, pages 253–276. Springer, 1985.

[320] Rina R Wehbe, Edward Lank, and Lennart E Nacke. Left them 4 dead: Perception of humans versus non-player character teammates in cooperative gameplay. In *Proceedings of the 2017 Conference on Designing Interactive Systems*, pages 403–415, 2017.

[321] A Rodney Wellens. Group situation awareness and distributed decision making: From military to civilian applications. *Individual and group decision making: Current issues*, pages 267–291, 1993.

[322] Jacob Westfall. Pangea: Power analysis for general anova designs. *Unpublished manuscript. Available at http://jakewestfall. org/publications/pangea. pdf*, 4, 2015.

[323] Susan A Wheelan. *Group processes: A developmental perspective.* Allyn & Bacon, 1994.

[324] Carolin Wienrich, Astrid Carolus, David Roth-Isigkeit, and Andreas Hotho. Inhibitors and enablers to explainable ai success: A systematic examination of explanation complexity and individual characteristics. *Multimodal Technologies and Interaction*, 6(12):106, 2022.

[325] RAND Wilcox. Robustness in anova. 1993.

[326] Jessica L Wildman, Amanda L Thayer, Michael A Rosen, Eduardo Salas, John E Mathieu, and Sara R Rayne. Task types and team-level attributes: Synthesis of team classification literature. *Human Resource Development Review*, 11(1):97–129, 2012.

[327] H James Wilson and Paul R Daugherty. Collaborative intelligence: humans and ai are joining forces. *Harvard Business Review*, 96(4):114–123, 2018.

[328] John R Wilson and Andrew Rutherford. Mental models: Theory and application in human factors. *Human Factors*, 31(6):617–634, 1989.

[329] Marcel Woide, Dina Stiegemeier, Stefan Pfattheicher, and Martin Baumann. Measuring driver-vehicle cooperation: development and validation of the human-machine-interaction-interdependence questionnaire (hmii). *Transportation research part F: traffic psychology and behaviour*, 83:424–439, 2021.

[330] David D Woods. Coping with complexity: the psychology of human behaviour in complex systems. In *Tasks, errors, and mental models*, pages 128–148. 1988.

[331] David D Woods. Decomposing automation: Apparent simplicity, real complexity. In *Automation and human performance: Theory and applications*, pages 3–17. CRC Press, 1996.

[332] Julia L Wright, Jessie YC Chen, Michael J Barnes, and Peter A Hancock. The effect of agent reasoning transparency on automation bias: An analysis of response performance. In *International conference on virtual, augmented and mixed reality*, pages 465–477. Springer, 2016.

[333] Melanie C Wright and David B Kaber. Effects of automation of information-processing functions on teamwork. *Human Factors*, 47(1):50–66, 2005.

[334] Anna Wu, Xiaolong Zhang, Gregorio Convertino, and John M Carroll. Civil: support geo-collaboration with information visualization. In *Proceedings of the ACM 2009 international conference on Supporting group work*, pages 273–276, 2009.

[335] Edgar Toshiro Yano, Welton de Abreu, Per M Gustavsson, and Rose-Mharie Åhlfeldt. A framework to support the development of cyber resiliency with situational awareness capability. In *20th international command and control research and technology symposium, june*, pages 16–19, 2015.

[336] Nick Yee. Beyond 50/50: Breaking down the percentage of female gamers by genre. *Quantic Foundry*, 19:2017, 2017.

[337] John Yen, Xiaocong Fan, Shuang Sun, Timothy Hanratty, and John Dumer. Agents with shared mental models for enhancing team decision makings. *Decision Support Systems*, 41(3):634–653, 2006.

[338] Jianwen Yin, Michael S Miller, Thomas R Ioerger, John Yen, and Richard A Volz. A knowledge-based approach for designing intelligent team training systems. In *Proceedings of the fourth international conference on Autonomous agents*, pages 427–434, 2000.

[339] Ming Yin, Jennifer Wortman Vaughan, and Hanna Wallach. Understanding the effect of accuracy on trust in machine learning models. In *Proceedings of the 2019 chi conference on human factors in computing systems*, pages 1–12, 2019.

[340] Sung Wook Yoon, Alan Fern, and Robert Givan. Learning Heuristic Functions from Relaxed Plans. In *ICAPS*, volume 2, page 3, 2006.

[341] Tom Young, Devamanyu Hazarika, Soujanya Poria, and Erik Cambria. Recent trends in deep learning based natural language processing. *ieee Computational intelligenCe magazine*, 13(3):55–75, 2018.

[342] Rui Zhang, Nathan J McNeese, Guo Freeman, and Geoff Musick. " an ideal human" expectations of ai teammates in human-ai teaming. *Proceedings of the ACM on Human-Computer Interaction*, 4(CSCW3):1–25, 2021.

[343] Zikai Zhou and Pilar Pazos. Empirical perspectives of transactive memory systems: a meta-analysis. *Team Performance Management: An International Journal*, 2020.