TOPICS
TOPICS IN COGNITIVE SCIENCE

Topics in Cognitive Science 17 (2025) 217–247

ISSN: 1756-8765 online
DOI: 10.1111/tops.12681

This article is part of the topic "Building the Socio-Cognitive Architecture of COHUMAIN: Collective Human-Machine Intelligence," Cleotilde Gonzalez, Henny Admoni, Scott Brown and Anita Williams Woolley (Topic Editors).

# Self-beliefs, Transactive Memory Systems, and Collective Identification in Teams: Articulating the Socio-Cognitive Underpinnings of COHUMAIN

Ishani Aggarwal,[a] Gabriela Cuconato,[b] Nüfer Yasin Ateş,[c] Nicoleta Meslec[d]

[a]*Brazilian School of Public and Business Administration, Fundação Getulio Vargas*
[b]*Department of Organizational Behavior, Weatherhead School of Management, Case Western Reserve University*
[c]*Sabancı Business School, Sabancı University*
[d]*Department of Organisation Studies, Tilburg University*

Received 31 May 2022; received in revised form 20 June 2023; accepted 20 June 2023

**Abstract**

Socio-cognitive theory conceptualizes individual contributors as both enactors of cognitive processes and targets of a social context's determinative influences. The present research investigates how contributors' metacognition or self-beliefs, combine with others' views of themselves to inform collective team states related to learning about other agents (i.e., transactive memory systems) and forming social attachments with other agents (i.e., collective team identification), both important teamwork states that have implications for team collective intelligence. We test the predictions in a longitudinal study with 78 teams. Additionally, we provide interview data from industry experts in human–artificial intelligence teams. Our findings contribute to an emerging socio-cognitive architecture for *COllective HUman-MAchine INtelligence* (i.e., COHUMAIN) by articulating its underpinnings in individual and collective cognition and metacognition. Our resulting model has implications for the

Correspondence should be sent to Gabriela Cuconato, Weatherhead School of Management—CWRU, 11119 Bellflower Rd, Room 42, Cleveland, OH 44106, USA. E-mail: gabriela.cuconato@case.edu

critical inputs necessary to design and enable a higher level of integration of human and machine teammates.

*Keywords:* Collective intelligence; Self-beliefs; Human–AI teams; Socio-cognitive theory; Teams; Transactive memory systems; Collective team identification

## 1. Introduction

One way in which humans attempt to solve complex, multifaceted problems is by working in collectives; and like other species such as bees, ants, and wolves, human beings cooperate and can make group decisions that are better than those made by individuals (Couzin, 2009). In organizations, the importance of working collectively is well-recognized, and work teams are often used as a common way of structuring taskwork (Hong & Page, 2004; van Knippenberg & Schippers, 2007; Wuchty, Jones, & Uzzi, 2007). Accordingly, we know a lot about human–human team interaction and that some teams are more collectively intelligent than others, that is, they have the ability for sustained performance when faced with complex changes over time (Engel, Woolley, Jing, Chabris, & Malone, 2014; Gupta & Woolley, 2021; Woolley, Chabris, Pentland, Hashmi, & Malone, 2010).

However, we do not know as much about human–machine teams. Artificial intelligence (AI), that is, the machine in human–machine teams, represents a highly capable and complex technology that aims to simulate human intelligence for problem-solving (Glikson & Woolley, 2020). AIs solve problems that humans cannot easily—or at all—solve, faster and cheaper (Carter-Browne et al., 2021). It is well-acknowledged that human–machine teams outperform solely human teams or solely AI since humans and AI can complement and offset each other's strengths and weaknesses (Malone, Bernstein, Atkins-Burnett, & Xue, 2018; Song et al., 2022). For example, Centaur chess, a combination of humans and AIs working together, makes up a better player than either alone (Cassidy, 2014). Yet, we still do not know how the socio-cognitive processes in human teams unfold in human–machine teams.

Extant literature investigating the performance of human–machine (automation, AI, digital agent) interactions (e.g., Hancock et al., 2013; van de Merwe, Oprins, Eriksson, & van der Plaat, 2012) recommend that AIs meant to work in human teams should be designed with teamwork skills (i.e., making the team work in unity) such as the ability to adapt to team dynamics rather than only taskwork skills (i.e., fulfilling team goals; e.g., Carter-Browne et al., 2021; Tausch, Kluge, & Adolph, 2020). Going beyond solely being a tool and being a teammate instead calls for an understanding of team socio-cognitive processes (Phillips, Ososky, Grove, & Jentsch, 2011; Seeber et al., 2020). Accordingly, there have been recent calls for research to "build a science of cooperative AI" (Dafoe et al., 2021) and to understand the *COllective HUman-MAchine INtelligence* (COHUMAIN). This need sets the stage for integration between human teams and human–machine teams literature, where the knowledge from human teams can be extended and contextualized to human–machine interaction research.

Accordingly, in an attempt to articulate the socio-cognitive underpinnings of COHUMAIN, we investigate how human teams' cognitive and social aspects impact team effectiveness and offer insights for extensions to human–machine teams. To this aim, we theorize and empirically test a model of team effectiveness in human teams. This model is built on the tenets of the socio-cognitive theory that emphasizes that team members are not mere respondents to external events and conditions, but they also are equipped with a sense of agency (Bandura, 1999)—which primarily distinguishes them from the machines. Specifically, we study whether and how team members' creative self-efficacy (CSE)—that is, the belief an agent has in its ability to produce novel and useful outcomes (Tierney & Farmer, 2002)—informs collective social and metacognitive states that are necessary for sustained effectiveness. We probe how these self-beliefs help individuals learn about other team members (i.e., transactive memory systems [TMS] or who knows what within a team) and develop social attachments with them (i.e., collective team identification). We test this model in a longitudinal study with 78 human teams.

Then we extend the insights from the human teams to human–AI teams. To facilitate such theorization and provide further anecdotal insights into the observed patterns from our empirical model, we conducted interviews with nine industry experts who work in human–AI teams in two organizations. Theorizing the peculiarities of our empirical results in the new COHUMAIN setting extends the socio-cognitive theory of humans toward a more insightful understanding of the socio-cognitive architecture of collective human–machine intelligence. We discuss how the underlying dynamics of the socio-cognitive processes in human teams apply to human–AI teaming and propose prominent future research directions.

## 2. Theoretical background

### *2.1. Human–machine teams literature*

In line with the increasing pervasiveness of digital technology in organizations, research has been geared toward understanding how the interaction between humans and machines (automation, digital agents, AI, robotic processes, etc.) takes place such that it leads to effective collaboration and high performance (Hancock et al., 2013; van de Merwe et al., 2012). While there is an established body of research about human teams (for literature reviews, see Guzzo & Dickson, 1996; Hackman & Morris, 1975; Ilgen, Hollenbeck, Johnson, & Jundt, 2005; Mathieu, Hollenbeck, van Knippenberg, & Ilgen, 2017), research on human–AI teams is at its burgeoning phase (Carter-Browne et al., 2021). A vast majority of this work tends to focus on the AI side of the interface, such as AI's autonomy (Shneiderman, 2020), functional/non-functional requirements (Voas, 2004), representation (e.g., Glikson & Woolley, 2020; Strohkorb et al., 2016), learning methods (e.g., Hoi, Sahoo, Lu, & Zhao, 2021; Shane, 2019), and capacity to observe human behavior (e.g., Russell, 2019). However, there is relatively limited research on how AI participates in teamwork (Carter-Browne et al., 2021).

Within this underexplored research area of human–AI teamwork, issues of coordination, trust, and shared mental models have received particular attention. For effective coordination, team members must be observable, predictable, and directable (Christoffersen & Woods, 2002; Johnson et al., 2014); therefore, humans and AIs must engage in a set of behaviors that allows team members to accurately predict each other's ability to complete a task. In this regard, transparency—the extent to which a human team member can comprehend the inner processes of an autonomous agent (e.g., knowing whether there is a need to take over)–is argued to be central for coordination. In general, transparency is associated with increased situational awareness (Roth, Schulte, Schmitt, & Brand, 2020; Selkowitz, Lakhmani, & Chen, 2017) and improved decision accuracy of the human team members (Bhaskara et al., 2021; Stowers et al., 2020) while increasing mental workload in some situations (Guznov et al., 2020).

Similarly, the development of shared mental models is seen as a vital teamwork process in human–AI teams. Mental models are the "mechanisms whereby humans can generate descriptions of system purpose and form explanations of system functioning and observed system states, and predictions of future system states" (Rouse & Morris, 1986). Shared mental models produce mutually compatible expectations in each teammate (Jonker, van Riemsdijk, & Vermeulen, 2011), thereby enhancing team collaboration and trust between parties (Gervits et al., 2020). The development of shared mental models is reported to be facilitated by verbal and non-verbal communication between humans and AIs (Hanna & Richards, 2018). Yet, further research to study its antecedents has been called for (Andrews, Lilly, Srivastava, & Feigh, 2022).

Another research stream focuses on the socio-emotional attributes, specifically trust, in human–AI teams (Phillips et al., 2011; Seeber et al., 2020). The switch from seeing AI as a complex tool that solely supports task performance to seeing it as a teammate (Matthews, Lin, Panganiban, & Long, 2020) pivots around trust in AI technology, which sets its role in organizations moving forward (Glikson & Woolley, 2020). Studies have shown that a variety of factors both individual differences (such as negative attitudes toward robots; Matthews et al., 2020) and AI characteristics (such as performance-based factors) influence the level of perceived trust in human–machine interactions (Hancock et al., 2013). A recent review distinguishes between a cognitive (based on perceptions of trustee reliance and competence) and an emotional dimension (based on affect) of trust. It further explicates that transparency, reliability, or predictability (i.e., exhibiting the same and expected behavior over time) and immediacy behaviors (i.e., socially oriented gestures to increase interpersonal closeness) of AI are crucial for developing cognitive trust, and anthropomorphism (or human-likeness) is essential for developing emotional trust (Glikson & Woolley, 2020).

Based on the current state of science, we build our research model around investigating how human members' self-beliefs influence their teamwork processes, in particular, learning about other agents (i.e., shared mental models of specialization, coordination, and trust) as well as forming social attachments with other members (i.e., identification with the collective) to influence taskwork (i.e., team performance), which are reviewed in the next subsections.

### 2.2. Metacognition and individual differences: Creative self-efficacy

Certain cognitive processes enable individuals to understand one another with a higher degree of precision. They are often referred to as mentalizing (Frith & Frith, 2012), which refers to implicit or explicit attribution of mental states to others and self (desires, beliefs) to explain and predict what they will do. A largely implicit form of mentalizing is likely to be involved in perspective-taking and tracking the intentional states of others, and this has been claimed for a variety of social animals as well as humans (e.g., Clayton, Dally, & Emery, 2007). It is only the explicit form of mentalizing that appears to be unique to humans and is closely linked to metacognition: the ability to reflect on one's actions and to think about one's own thoughts (Frith & Frith, 2012). This ability confers significant benefits to human social cognition over and above the contribution from the many powerful implicit processes that we share with other social species (Frith & Frith, 2012).

These metacognitive processes and beliefs are quite important in determining human behavior and surprising as it may be, people's behavior is determined by their beliefs rather than by physical reality, even if this belief happens to diverge from reality (Frith & Frith, 2012). In fact, it has long been argued that a core part of human mentalizing is that we recognize that other agents do not base their decisions directly on the state of the world but rather on an internal representation of the state of the world (Gopnik & Astington, 1988; Leslie, 1987; Rabinowitz et al., 2018; Wellman, 1992). Especially teamwork-related metacognitive individual differences will influence different ways of representing a desire and propensity to work in teams, collaborate, and align with team goals (e.g., Driskell, Salas, & Hughes, 2010; Salas, Sims, & Burke, 2005), all of which are important aspects of successful collaboration whether teaming with humans or AIs (Carter-Browne et al., 2021).

One such metacognitive individual difference is CSE, which is an individual's belief and confidence in mobilizing his or her resources for successful performance (Bandura, 1977), and has been theorized to affect human behavior and interaction with social contexts across a wide range of settings (Wood & Bandura, 1989).

When individuals believe more strongly in their own capabilities, they feel more confident, perceive difficulties as challenges, set higher goals, and exert more effort to accomplish an outcome (Michael, Hou, & Fan, 2011). Self-efficacy further determines whether an individual chooses to engage in certain behavior, and, if so, how much effort is expended on that behavior (Bandura, 1997; Eden & Aviram, 1993; Tasa, Taggar, & Seijts, 2007). Research consistently shows that efficacy beliefs contribute significantly to the level of motivation and performance (Bandura, 1997; Bandura & Locke, 2003; Gist & Mitchell, 1992; Katz-Navon & Erez, 2005; Locke & Latham, 2002; Tasa et al., 2007).

In predicting task-related human performance, specific self-efficacy beliefs—such as CSE, academic self-efficacy, computer self-efficacy, musical self-efficacy, and so forth—are more powerful, compared to global or general self-efficacy beliefs (Gist & Mitchell, 1992; Menold & Jablokow, 2019; Simmons, Payne, & Pariyothorn, 2014; Tasa et al., 2007). For example, academic self-efficacy—students' beliefs about their academic capabilities—is regarded as one of the strongest predictors of students' achievement as this belief becomes an inner resource of their academic engagement and performance (Diseth, Meland, & Breidablik,

2014; Levpušček & Zupančič, 2009; Multon, Brown, & Lent, 1991; Valentine, DuBois, & Cooper, 2004). This is counter-intuitive since one would imagine that the actual capability would be the strongest predictor. We contend that in settings that require the completion of complex, multifaceted tasks through teamwork with novel ways of operating and coordinating, which collectively intelligent behavior entails, CSE is a particularly relevant self-belief.

Scholars working from Bandura's general self-efficacy concept defined CSE as one's belief in one's own ability to produce novel and useful—as opposed to routine—outcomes (Tierney & Farmer, 2002). Different from artistic or scientific creativity—which might be limited to specialized settings—"everyday creativity" (that CSE centers on) refers to self-expression in daily activities, interpersonal style, avocational pursuits, and problem-solving in daily life (Ivcevic, 2007; Richards, Kinney, Lunde, Benet, & Merzel, 1988; Torrance, 1988) and reflects an ability to respond to challenges in the environment, assuring resilience and personal growth (Cropley, 1990; Flach, 1990; Richards & Kinney, 1990). A belief in one's capacity for such creativity is likely to serve as an inner resource, which is likely to determine team members' interaction patterns and the amount of effort one is willing to exert in social contexts, especially about goal or task accomplishment. CSE has been empirically shown to predict various task-related behavior such as job performance, explaining additional variance, compared to job self-efficacy in non-routine contexts (Tierney & Farmer, 2002), creative performance (Gong, Huang, & Farh, 2009; Jaussi & Randel, 2014; Tierney & Farmer, 2002), imagination (Liang & Chang, 2014), and innovative behavior (Farmer & Tierney, 2017; Michael et al., 2011).

What happens when individuals with high (or low) levels of CSE work in a collective? To understand this, we look at team composition, which refers to the configuration of member attributes that are considered inputs that translate into outputs through the medium of team states and processes (Bell, Brown, Colaneri, & Outland, 2018; Marks, Mathieu, & Zaccaro, 2001). Specifically, we argue that high levels in a team's CSE composition enhance learning about other team members (i.e., TMS) and forming social attachments with them (i.e., collective team identification). These two states represent team-level social cognitive processes and both have been shown to have a strong influence on a team's ability to function and perform well (Bezrukova, Jehn, Zanutto, & Thatcher, 2009; Lewis, 2003; Lin, He, Baruch, & Ashforth, 2017; Solansky, 2011; Zhang, Hempel, Han, & Tjosvold, 2007).

### *2.3. Collective socio-cognitive team states*

#### *2.3.1. Learning about other agents: Transactive memory systems*

A TMS refers to the awareness a group has about its members' skills and knowledge (Wegner, 1987). It pertains to how groups process and structure information as a collective mental mind that encodes, stores, retrieves, and communicates group knowledge (Hollingshead, 1998). This socially shared cognitive state is a key factor for working groups to ensure that important information is not forgotten, enabling members to know who is good at what and who knows what (Moreland, Argote, & Krishnan, 1996). As noted by Frith and Frith (2012), arguably the most important and valuable aspect of social cognition is taking account of other

individuals and learning about other agents and the need to keep track of who has the most relevant knowledge to function well as a collective.

TMS works as an external source of information storage, with which team members can locate and retrieve information that might be unavailable to them otherwise (Liang, Moreland, & Argote, 1995). When teams have a high level of TMS, they can jointly specialize in different knowledge areas, give credibility to each other's knowledge, and better coordinate knowledge retrieval processes (Lewis, 2003; Lewis & Herndon, 2011; Moreland et al., 1996; Zhang et al., 2007). Taken together, these three sub-dimensions of TMS, namely, specialization, credibility, and coordination, increase a team's ability to access a larger pool of information while reducing members' cognitive load (Hollingshead, 1998).

Team composition variables have been shown to affect TMS formation. According to the signal-detection perspective, team composition attributes and configurations that facilitate the location of where team members' cognitive resources—by amplifying signals and promoting their detection—facilitate the formation of a team's TMS (Aggarwal & Woolley, 2019). Previous research reports team expertise composition and team cognitive style composition as antecedents of TMS (Aggarwal & Woolley, 2019; Hollingshead, 2001).

Following these studies, we argue that the team's CSE composition facilitates the tracking of who knows what in the team for several reasons. First, the confidence and belief in one's capability in connecting things and creatively working within available resources encourage these team members to display and/or assert their skills and knowledge to other team members more readily. When members willingly reveal their skills, this enhances the team's overall understanding of where team members' knowledge and skills are located within the team and facilitate the formation of TMS. Second, team members high in CSE are also likely to seek an understanding of others' knowledge and skills to gauge the resources that are available within the team to fuel their creative endeavors while fulfilling the tasks. This is especially relevant in a teamwork context where tasks are complex, interdependent, and multifaceted, so drawing on other team members' skills and understanding and combining them are essential resources through which effective teamwork is conducted. Such a need motivates team members to understand where member resources lie, leading to a higher level of TMS. Third, as CSE increases, creativity in problem-solving and approaching the task will be enhanced (Tierney & Farmer, 2002), allowing members to link ideas from multiple sources and areas (Gilson & Shalley, 2004), thus leading teams to be more inclined to access each other's knowledge. Accordingly, we expect that higher levels of CSE in team composition will facilitate TMS, and predict the following:

**Hypothesis 1**. Team-level CSE is positively related to transactive memory systems in teams.

Research on TMS development has shown TMS's positive effects on team outcomes such as goal attainment, knowledge integration, creativity, learning, ambidexterity, and performance on several types of complex and multifaceted tasks (Aggarwal & Woolley, 2019; Argote & Ren, 2012; Austin, 2003; Cabeza-Pullés, Gutierrez-Gutierrez, & Llorens-Montes, 2018; Gino, Argote, Miron-Spektor, & Todorova, 2010; Heavey & Simsek, 2017; Hollingshead, 1998; Huang & Chen, 2018; Lewis & Herndon, 2011; Lewis, Lange, & Gillis, 2005),

which are, broadly, indicators of collective intelligence. Further, the development of TMS in teams ensures that members' skills and knowledge are effectively brought into play (Gupta & Woolley, 2021) and is thus considered valuable especially when knowledge becomes obsolete and current problems change (Ren, Carley, & Argote, 2006). TMS facilitates group adaptation to new tasks (Lewis et al., 2005) and new problems (Miller, Choi, & Pentland, 2014) and enables groups to be more creative (Aggarwal & Woolley, 2019; Gino et al., 2010). When team members know their knowledge domains, believe in the credibility of that knowledge, and coordinate themselves efficiently, they can reshape their knowledge toward superior team outcomes (Argote & Ren, 2012). Thus, we anticipate that by facilitating the development of TMS, team-level CSE will have a positive indirect impact on team effectiveness through the mechanism of TMS. Accordingly, we predict the following:

**Hypothesis 2**. Team's transactive memory systems mediates the relationship between team-level CSE and team effectiveness.

*2.3.2. Social attachments with other agents: Collective team identification*

Collective team identification occurs when individuals have an emotional attachment to the group and recognize themselves as members (Sutton, McDonald, Milne, & Cimperman, 1997) and perceive themselves as sharing a common association with a particular group (Bezrukova et al., 2009; Tajfel, 1978; Van Der Vegt & Bunderson, 2005). It refers to a team's motivational climate to overcome disruptive tendencies and reflects both the motivation to work toward meeting common objectives and the commitment to overcome any difficulties (Kearney, Gebert, & Voelpel, 2009; Van Der Vegt & Bunderson, 2005).

Social identification theory posits identification as a part of an individual's self-concept aroused from a membership to a group (Tajfel, 1978) where a team member experiences the team's successes and failures as her own (Foote, 1951) along with the desire to see the team succeed (Pearsall & Venkataramani, 2015). Therefore, members put more effort and actively contribute to team goals (Mael & Ashforth, 1992), sacrificing their own interests, if necessary, as they see teams' achievements as their own. Collective team identification has been studied in previous research as a team state influencing the relationship between team composition and team effectiveness (Kearney et al., 2009).

We argue that the team's CSE composition will facilitate the formation of this motivational climate toward meeting common objectives and the commitment to overcoming any difficulties. Specifically, individuals high in CSE are likely to mobilize more effort, motivation, and cognitive resources in teams as they believe that they can deal with problems, including the complexity of teamwork itself (Beghetto, 2006; Beghetto & Karwowski, 2017; Gong et al., 2009; Michael et al., 2011; Sangsuk & Siriparp, 2015; Tierney & Farmer, 2002). Moreover, team members with high CSE are less concerned about other team members' potential negative assessments of themselves due to their greater confidence in their abilities and knowledge (less evaluation apprehension; Diehl & Stroebe, 1987). They will be less sensitive to the difficulty of identifying individual contributions in teamwork and thus will be more motivated to see the team's goals as their own goals, further enhancing the levels of collective identification with the team. Also, the mere presence of others biases us toward group-oriented

behavior (Frith & Frith, 2012). Hence, when people are high on this self-belief, they might be more likely to favor and have bonds with the in-group, that is, they might consider the group as having more resourcefulness and hence be more attached to it. Accordingly, we predict that team CSE composition will trigger the development of higher levels of collective team identification.

**Hypothesis 3**. Team-level creative self-efficacy is positively related to collective team identification.

Teams with higher levels of collective team identification have been shown to have higher levels of team effectiveness across settings (Kearney et al., 2009; Lin et al., 2017; Solansky, 2011), also yielding higher satisfaction among team members satisfied (Johnson & Avolio, 2019). Research also reports that the negative effects of team diversity (Bezrukova et al., 2009), as well as social undermining (Duffy, Scott, Shaw, Tepper, & Aquino, 2012), are mitigated in teams with high collective identification. Collective team identification leads to team performance through enhanced coordination within the team (Lin et al., 2017).

Since team members feel attachment toward their teams and interpret the team's successes as their own, collective team identification leads members to dedicate more effort to meeting teams' goals, influencing sustained team performance, which is broadly related to collective intelligence. Thus, a positive relationship between collective team identification and team performance is expected. Furthermore, since team CSE is related to an enhanced dedication of team members toward team goals, social interactions, and effort in tasks, thereby influencing collective team identification positively, we assert that it will foster team effectiveness indirectly through collective team identification. Accordingly, we predict the following:

**Hypothesis 4**. Collective team identification mediates the relationship between team creative self-efficacy composition and team effectiveness.

More comprehensively, we predict that team CSE composition will impact team effectiveness indirectly by simultaneously influencing both the TMS and collective team identification. Accordingly, combing the logic in the previous hypotheses, we predict:

**Hypothesis 5**. Team-level creative self-efficacy will have an indirect relationship with team effectiveness simultaneously through transactive memory systems and collective team identification.

## 3. Method

### *3.1. Participants*

The sample in the study consisted of 312 senior bachelor students randomly assigned to 78 four-person teams as a part of an Integrated Business Administration course. The teams worked together for a semester. Twenty-nine percent of the participants were female.

Forty-nine percent of the participants were majoring in accounting or finance, 22% in organization and strategy, 9% in marketing, and 20% in other areas.

### 3.2. Task

Each team was responsible for managing an athletic footwear company in a simulated business environment called The Business Strategy Game (Thompson & Stappenbeck, 1999). Student teams competed with each other in a given industry. To outperform other teams, students completed multifaceted tasks, encompassing decisions about operations (e.g., capacity decisions, plant upgrades, closings, workforce improvements, shipping), finance (e.g., cash flow, loans, stock buybacks), marketing and sales (e.g., pricing, product variety, celebrity endorsements), and human resources (e.g., wages, bonuses, layoffs) in four different geographical regions (i.e., North America, Asia, Europe, and Latin America). The task necessitates an understanding of the interrelation between different business functions, and students make several functional decisions that must be aligned with an overarching business strategy. The nature of the simulation requires team members to work and make decisions together, as their tasks are highly interdependent.

The teams played for a total of 14 rounds, and the first two rounds were practice rounds (scores reset after the practice rounds). After each round, detailed company performance reports were automatically generated for each company, together with a general industry summary report, so that teams could make informed decisions by considering the general industry trends and their competitors' potential actions.

This experiential learning exercise constituted 50% of the student course grade, providing an incentive for participants' dedication to the game. The same professor administered all teams and industries. Individual consultation was not provided regarding the tasks or decision-making to avoid variance in the information provided to different teams. Participants had not participated in similar games in their undergraduate curriculum previously. This setting has also been used in past research as a representative of business environments (Boies, Lvina, & Martens, 2011; Mathieu & Rapp, 2009).

### 3.3. Procedure and measures

Before the onset of the simulation, participants answered an online survey at the beginning of the course. This survey contained measures of demographic characteristics and CSE. Halfway into the semester, participants completed a second survey, assessing TMS and collective team identification. At the end of the semester, team effectiveness was objectively determined by the simulation system.

#### 3.3.1. Team CSE composition

Team members' CSE was assessed by the eight-item scale from Chen, Gully, and Eden (2001) adapted by Carmeli and Schaubroeck (2007). Participants were asked about the extent to which they believe in their creative abilities on a 7-point Likert-type scale ranging from 1 (*strongly disagree*) to 7 (*strongly agree*). Sample items include "I will be able to achieve

most of the goals I have set for myself in a creative way" and "Even when things are tough, I can perform quite creatively." Following the team personality elevation paradigm (Neuman, Wagner, & Christiansen, 1999) and the additive compositional model (Chan, 1998), a team's CSE was calculated as the average of members' CSE scores. The Cronbach's alpha for this scale was 0.94 (Cronbach, 1951). Cronbach's alpha is referred to as a measure of internal consistency or reliability (Bonett & Wright, 2014), with higher values indicating that the scale items hang well together. It is the most widely used reliability measure in the social and organizational sciences when measures with multiple questionnaire items are used (Bonett & Wright, 2014).

#### *3.3.2. TMS*

A 15-item scale from Lewis (2003) was used to measure teams' TMS. TMS has three dimensions: specialization, credibility, and coordination. Each dimension is measured by five items on a 5-point Likert-type scale ranging from 1 (*strongly disagree*) to 5 (*strongly agree*). Sample items are "Each team member has a specialized knowledge of some aspect of the game" (specialization); "I was comfortable accepting procedural suggestions from other team members" (credibility); and "Our team worked together in a well-coordinated fashion" (coordination). Cronbach's alpha for the three dimensions ranged from 0.79 to 0.84, and the reliability of the whole TMS scale was 0.82.

As a team-level construct, individual answers were aggregated to team level (median $r_{wg(j)}$ value was 0.98; intraclass correlation [ICC](1) = 0.27, ICC(2) = 0.59, $F = 2.47$, $p < .01$; Bliese, 2000; James, Demaree, & Wolf, 1993). The $r_{wg(j)}$ estimates the interrater agreement for a group by comparing the realized group variance with an expected random variance (James et al., 1993). Values higher than 0.7 are deemed satisfactory. The ICC(1) is a measure of the extent to which individual ratings are attributable to team memberships, while ICC(2) demonstrates the reliability of average ratings by a group of evaluators. These statistics support the aggregation to the team level by averaging (Bliese, 2000).

#### *3.3.3. Collective identification*

Collective identification was assessed by a five-item adapted version of the organizational identification scale from van Knippenberg and Schippers (2007) and Mael and Ashforth (1992). On a 5-point Likert-type scale ranging from 1 (*strongly disagree*) to 5 (*strongly agree*), participants answered questions including "When someone criticizes my team, it feels like a personal insult." Cronbach's alpha of this scale was 0.68. Member's individual answers were further aggregated to a team-level construct (median $r_{wg(j)}$ value was 0.96; ICC(1) = 0.10, ICC(2) = 0.31, $F = 1.45$, $p < .05$; Bliese 2000; James et al. 1993).

#### *3.3.4. Team effectiveness*

Team effectiveness was calculated automatically by the simulation system and operationalized as the average of how well each team met the five performance targets set by the simulation: (a) Earnings per share: grow earnings per share at least 7% annually, (b) return on equity: maintain a return on average equity investment of 15% or more annually, (c) credit rating: maintain a B+ or higher credit rating, (d) image rating: achieve an image rating of

70 or higher, and (e) stock price: achieve a stock price gains averaging about 7% annually. This approach is analogous to the balanced scorecard approach where the overall performance score is a composite of different performance indices and is also recommended by the publishers of the simulation (Thompson & Stappenbeck, 1999) because it would capture the effectiveness of different overall strategies (i.e., cost leadership, differentiation, cost focus or differentiation focus strategies) pursued by the companies. The simulation also had built-in bonus points awards, namely, the bull's eye award (awarded annually for the most accurately projected company performance) and the leap-frog award (awarded annually for the most improved overall performance). This objective measurement for company performance was clearly communicated to all students before the game. Effectiveness scores ranged from 49 to 108, on a scale spectrum of 0 to 110.[1]

#### *3.3.5. Control variables*

We controlled for team CSE diversity, which was operationalized as the within-team standard deviation on CSE, following dispersion measures recommended in the literature (Barrick, Stewart, Neubert, & Mount, 1998; Chan, 1998; Harrison & Klein, 2007) to account for within-team differences that average measures do not capture.

### *3.4. Analytical approach*

To test our main effect hypotheses (H1 and H3), we use multiple regression analysis, regressing the dependent variable on the independent variables and controls. Further, mediation hypotheses elucidate how, or more specifically, through which processes or mechanisms, the independent variable of interest exerts its influence on the dependent variable (Preacher & Hayes, 2008). To test our mediation hypotheses (H2, H4, and H5), we used the common practice of generating bootstrapped confidence intervals for the indirect effects by using the statistical routines proposed by Hayes (2017) and Preacher and Hayes (2004, 2008).[2]

## 4. Results

### *4.1. Empirical results from teams study*

Descriptive statistics and correlations among the study measures are presented in Table 1. Table 2 shows the results of Ordinary Least Squares (OLS) regressions used to test the main effect hypotheses (H1 and H3). Table 3 presents the results of bootstrapped confidence intervals used to test the mediation hypotheses (H2 and H4) and the multiple regression model (H5). All models are tested with 5000 bootstrapped samples, controlling for diversity in CSE. We present the standardized results ($\beta$) for ease of comparability across effects in which stronger effects mean proportionally higher values, unstandardized results (b) for interpretability of effects as it indicates how many units are changed in the dependent variable by one unit increase in the independent variables, and the bootstrapped standard errors that are used to calculate the respective $p$-value ($p$) and confidence interval (CI), which indicates the

Table 1
Team descriptives and intercorrelations

| | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| 1. Team Effectiveness | – | | | | |
| 2. Team-level Creative Self-Efficacy (CSE) | 0.24* | (0.94) | | | |
| 3. Transactive Memory Systems (TMS) | 0.45** | 0.26* | (0.82) | | |
| 4. Collective Identification | 0.39** | 0.33** | 0.39** | (0.68) | |
| 5. CSE diversity | −0.11 | −0.21 | −0.11 | −0.14 | – |
| Minimum | 49 | 3.44 | 2.71 | 2.93 | 0.16 |
| Maximum | 108 | 5.66 | 3.93 | 4.15 | 1.72 |
| Mean | 87.12 | 4.59 | 3.41 | 3.52 | 0.88 |
| Standard Deviation | 14.62 | 0.46 | 0.25 | 0.27 | 0.38 |

*Note.* $N = 78$; Cronbach alphas in parentheses.
*$p < .05$; **$p < .01$.

Table 2
Results of OLS regressions

| | TMS | Collective Identification | Team Effectiveness | | | |
|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 |
| Team-level CSE | 0.14* | 0.18** | 7.03* | 3.68 | 3.67 | 2.06 |
| | $\beta = 0.25$ | $\beta = 0.31$ | $\beta = 0.22$ | $\beta = 0.12$ | $\beta = 0.12$ | $\beta = 0.06$ |
| | (0.06) | (0.06) | (3.49) | (3.54) | (3.77) | (3.59) |
| TMS | | | | 24.03** | | 19.84** |
| | | | | $\beta = 0.42$ | | $\beta = 0.35$ |
| | | | | (6.62) | | (7.06) |
| Collective Identification | | | | | 18.23** | 11.96* |
| | | | | | $\beta = 0.34$ | $\beta = 0.22$ |
| | | | | | (5.87) | (5.69) |
| CSE diversity | −0.04 | −0.06 | −2.54 | −1.62 | −1.51 | −1.10 |
| | $\beta = -0.06$ | $\beta = -0.08$ | $\beta = -0.07$ | $\beta = -0.04$ | $\beta = -0.04$ | $\beta = -0.03$ |
| | (0.08) | (0.09) | (4.46) | (3.70) | (3.90) | (3.43) |
| $R^2$ | 0.07 | 0.11 | 0.06 | 0.22 | 0.16 | 0.26 |
| Adjusted $R^2$ | 0.05 | 0.09 | 0.03 | 0.19 | 0.13 | 0.22 |
| Wald test | 6.44 | 9.32 | 4.19 | 14.55 | 15.13 | 26.91 |

*Note*. Unstandardized coefficients presented, $\beta$ = standardized coefficients, bootstrapped standard errors with 5000 samples in parentheses. Two-tailed test. All models tested with CSE diversity as a control.
*$p < .05$; **$p < .01$.

statistical significance of the results in supporting our hypotheses ($p < .05$, 95% CI does not include zero).

Hypothesis 1 predicted a positive relationship between the team's CSE and the team's TMS. The results in Table 2 (Model 1) provide support for this hypothesis ($\beta = 0.25$, $b = 0.14$, $SE$ (Boot) $= 0.06$, $p = .01$).

Table 3
Results of mediation analysis—indirect effects of team CSE composition on team effectiveness

| | | | |
|---|---|---|---|
| Model 1 | | | |
| Mediator: TMS | | | |
| | Indirect Effect | Direct Effect | Total Effect |
| Coefficient | 3.36* | −3.68 | 7.03 |
| | *std* = 0.11 | *std* = −0.12 | *std* = 0.22 |
| Bootstrapped *SE* | 1.54 | 3.49 | 3.59 |
| 95% CI | 0.33 to 6.38 | −3.16 to 10.51 | 0.00 to 14.06 |
| Model 2 | | | |
| Mediator: Collective Identification | | | |
| | Indirect Effect | Direct Effect | Total Effect |
| Coefficient | 3.36* | −3.91 | 7.03 |
| | *std* = 0.11 | *std* = −0.12 | *std* = 0.22 |
| Bootstrapped *SE* | 1.63 | 3.57 | 3.66 |
| 95% CI | 0.17 to 6.55 | −10.92 to 3.09 | −0.14 to 14.21 |
| Model 3 - Multiple Mediation Model: | | | |
| Total Indirect Effect | | | |
| Mediators: TMS and Collective Identification | | | |
| | Indirect Effect | Direct Effect | Total Effect |
| Coefficient | 4.97** | 2.06 | 7.03* |
| | *std* = 0.16 | *std* = 0.07 | |
| Bootstrapped *SE* | 1.83 | 3.61 | 3.54 |
| 95% CI | 1.39 to 8.56 | −5.01 to 9.13 | 0.10 to 13.97 |

*Note*. All models tested with CSE diversity as a control. Bootstrapped standard errors with 5000 samples.
*$p < .05$; **$p < .01$.

There was a positive and significant relationship between TMS and team effectiveness ($\beta = 0.35$, $b = 19.84$, *SE* (Boot) = 7.06, $p = .01$; Table 2, Model 6). Hypothesis 2 predicted that the influence of team CSE on team effectiveness will be mediated by TMS. Results in Table 3 (Model 1) showed a positive and significant indirect effect of team-level CSE on team effectiveness through TMS (indirect effect = 3.36, standardized indirect effect = 0.11, *SE* (Boot) = 1.54, $p = .03$, 95% CI excluded zero [0.33 to 6.38]).

Hypothesis 3 predicted a positive relationship between team-level CSE and collective identification. Results in Table 2 (Model 2) show support for this hypothesis ($\beta = 0.31$, $b = 0.18$, *SE* (Boot) = 0.06, $p = .00$).

There was a positive and significant relationship between collective identification and team effectiveness ($\beta = 0.22$, $b = 11.96$, *SE* (Boot) = 5.69, $p = .04$; Table 2, Model 6). Hypothesis 4 predicted that collective identification will mediate the relationship between team-level CSE

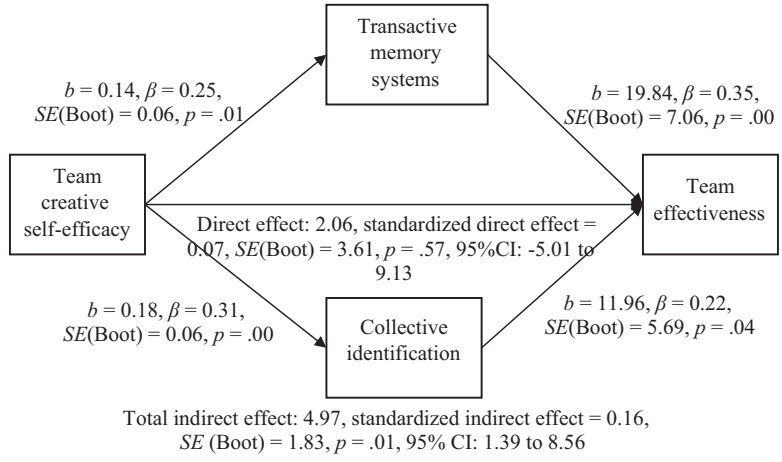


Fig 1. Results of multiple mediation analysis controlling for creative self-efficacy (CSE) diversity. $b$ = unstandardized estimate, $\beta$ = standardized estimate.

and team effectiveness. Results in Table 3 (Model 2) provided support for this hypothesis and showed a positive and significant indirect effect of team CSE level on team effectiveness through collective identification (indirect effect = 3.36, standardized indirect effect = 0.11, *SE* (Boot) = 1.63, $p = .043$, 95% confidence interval excluded zero [0.17 to 6.65]).

To test Hypothesis 5, which predicted that TMS and collective identification will simultaneously mediate the relationship between team-level CSE and team effectiveness, we bootstrapped the multiple mediation model, following Preacher and Hayes's (2008) procedure. This hypothesis was supported. The total indirect effect (Table 3, Model 3) was positive and significant (total indirect effect = 4.97, standardized total indirect effect = 0.16, $p = .006$; *SE* (Boot) = 1.83) with the 95% bias-corrected confidence interval excluding zero (95% CI = [1.39 to 8.56]; Fig. 1).

### *4.2. Qualitative evidence from interviews*

We tested our socio-cognitive model in human teams. To be able to make accurate reflections about human–machine teams, we conducted semi-structured interviews with nine experts (two heads of IT departments, two AI engineers, five human–AI team members) from two industrial organizations, namely, Brisa and Kordsa.[3] The purpose of these interviews was not to build or test the new theory but to acquire further insights into the empirical results derived from human teams (Ateş, Tarakci, Porck, van Knippenberg, & Groenen, 2020) and what they would mean in human–AI teams.

Both organizations have internally developed their own AI systems that comprise robotic process automation, chatbot, and optical character recognition features. What was once started as a curious experimentation with the new AI technology almost half a decade ago has quickly become a central pillar of their business operations. The AI systems are actively used

in more than 50 functional and cross-functional core business processes in each organization and save thousands of work hours while being cheaper, faster, and more accurate. Starting with a trend toward delegating only the routine, menial, and time-consuming tasks to AI, the organizations currently run their complex, multifaceted, difficult, and attention-demanding critical tasks with AI–human teams.

Robi (AI agent in Brisa) and Jojo (AI agent in Kordsa) are called *metal-collar* colleagues and are listed as a "team member" in the company communication platform Microsoft Teams. Employees communicate with their machine teammates by typing on the chat screen or via voice comments. Both AIs have robot avatars and animations. The AI engineers and corporate communications department paid particular attention to making the AI agents more human-like to enhance their spread and utilization. Beyond receiving commands from their human teammates, the AI agents also proactively engage with their human teammates. For instance, Jojo has 13 proactive functions from simply asking how things are going to suggesting new songs and popular YouTube videos, very much like a real teammate. It also checks its human teammates' calendars and alerts them if there are too many consecutive meetings or if they will miss lunchtime due to the meetings. Consequently, one human–AI team member said that "*well, we all eventually perceived it* [the AI] *as a real teammate*." Therefore, these two organizations provide a good context to explore the implications of our socio-cognitive model (Yin, 2013).

All interviewees were well informed about the AI agents and involved in the development of AI projects to varying extents either as designers, engineers, business analysts, project managers, or informants. They were also all members of human–AI teams. In the semi-structured interviews, we first walked through the constructs of our model with the interviewees and allowed them to reflect on their experiences with and without AI teammates. Then we sought their opinions about the associations between these constructs and asked them to speculate how these relations would unfold in human–AI teams.

The interviewees noted that they spend less time on operational tasks—thanks to the help of their machine teammate—and more time on value-added activities. For instance, the automatization of cost analysis for different product variations saves an accounting expert 1 h per analysis. With AI taking over the task, not only the expert is free from several hours of tedious work a day but also the capacity to respond to cost analysis requests increases substantially, speeding things up considerably. The expert is now responsible for analyzing the metadata that is created by the sales teams' cost analysis requests and developing insights for both departments. This is a more meaningful task that can bring a competitive advantage in the market. One interviewee said "*We used to spend 80% of our time on operations and 20% on projects. Now, this is reversed. We are more entrepreneurial*."

The respondents also shared that they feel more creative as a team. Their involvement in AI process automation provided them with a chance to develop a broader perspective and understand interactions between different functional domains. Before moving a business process to the digital environment, team members and the IT department study the process holistically together with stakeholders from other departments and collectively improve the process into a leaner version. This practice builds confidence in teams and gradually develops a bottom-up cultural change in which teams constantly search for projects to delegate to

AI. The head of IT in one of the organizations said that they "*do not have enough capacity to keep up with all automation demands from functional units, thus they recently started 'low-code, no-code' platform training [these platforms allow business users without any coding knowledge to develop apps simply by dragging and dropping components and visually connecting them instead of writing code line by line] to improve their work force's capability in AI development*."

It is not difficult to decipher that the broader view, initiative taking, and the accompanying cultural transformation (not to mention immediate performance implications) enhanced team members' collective identification as well. Respondents reported their enthusiasm for this new way of working and their intensified commitment to their team.

When it comes to transactive knowledge memory systems, one might suspect there may be a loss of experience and knowledge over time since AI operates as a black box to many people. With certain turnover rates, the new generation of employees might be too dependent on their machine teammates, as they might not know who knows what (i.e., low TMS). However, the interviews revealed quite the opposite. First, as already mentioned, tasks are streamlined by cross-functional stakeholders prior to being delegated to AI. This includes a video recording of all task steps and extensive documentation, which codifies implicit knowledge into explicit knowledge. This process amplifies learning within and between departments and improves TMS. Second, training the AI with old cases also reveals unknown/unexpected exceptions in the processes, which further adds to the TMS. As one respondent puts it "*We are preparing the organization for the next generation…. we [team members] know it [what we do] much better now than before*."

The reflection of respondents on the overall model was, in principle, in line with our theorization. They speculated the relationships would be even stronger in human–AI teams as AI may be a catalyzer, leaving more room for quality human interaction. AI would help human–AI teams to develop high levels of CSE (through instilling confidence in humans' creative efforts in generating AI), thus leading to high collective identification (through allowing them to do more meaningful work) and strong TMS (through extensive documentation and collective work), which eventually leads to high team effectiveness (not only taskwork performance but also better teamwork). Two reservations were mentioned. First, the human team members' involvement in AI design and delegating their tasks to AI on their own is seen as a key contingency. Some resistance was noted when centrally developed AI applications were attempted to be rolled over to international plants. Second, human team members were less tolerant of AI errors. While training newcomers, team members are often quite tolerant of early mistakes; however, respondents acknowledged observably less patience with AI's mistakes. Proper training with past data (or realistically simulated data) is a must before introducing a new AI teammate functionality.

These observations provide qualitative anecdotal evidence and some expert speculations about how our theorization concerning human teams would unfold in human–AI teams. The preliminary evidence suggests that our socio-cognitive model is relevant in human–AI teams and human involvement in AI development might be a central boundary condition.

## 5. Discussion

### 5.1. Human teams

The first goal of our study was to examine empirically the extent to which specific metacognitive beliefs in human teams (i.e., CSE) contribute to the development of teamwork-related states pertaining to learning about other agents (i.e., TMS) and developing attachments with them (i.e., collective identification) and their further impact on team effectiveness. In this paper, we found novel evidence that team-level CSE facilitates team performance in a multifaceted task context, through its influence on a team's TMS and collective team identification. When teams have higher levels of CSE, at least some of its members will have confidence in their capabilities to be creative, within available resources, mobilize more effort, being more motivated toward teamwork, especially because they are more likely to believe that they have the capability to deal with problems, including the complexity of teamwork itself (Beghetto & Karwowski, 2017; Gong et al., 2009; Tierney & Farmer, 2002). As such, they will be more likely to display and/or assert their skills and knowledge, also seeking to know others' skills, and caring more about teamwork, and exchange information that will lead to the development of specialized knowledge, coordination, and trust in other team members' knowledge (i.e., TMS). At the same time, when team members perceive high chances of succeeding, solving problems, and achieving their goals creatively, they will be more likely to develop an emotional attachment to the group and identify with it, adopting teams' goals and outcomes as their own (i.e., collective team identification).

We also were able to replicate the positive relationship of both TMS and collective team identification with team performance that has been found in past studies. And these results contribute to the accumulation of knowledge in the field of team research by showing that the positive effects of TMS and collective team identification on performance are robust across different samples and types of tasks. As it has been advanced on various occasions (Antonakis, 2017; Koole & Lakens, 2012; Makel, Plucker, & Hegarty, 2012), replication studies have oftentimes been neglected although it is only through replications that the significance and relevance of a research line can be established. Further, by showing that both these processes simultaneously mediate the relationship between team-level CSE and team performance, we contribute toward building a more comprehensive understanding of team functioning rather than taking a piecemeal approach by studying only one process at a time.

### 5.2. Human–AI collaboration and collective intelligence

The second goal of our study was to extend and generalize the findings obtained in the human–human teams study to the human–machine teaming field of research. To generate informed insights, we conducted interviews with several industry experts working with human–AI teams. In the following subsections, we elaborate on how these findings can help us set the stage for further and more extensive research on human–machine teaming.

*5.2.1. Humans understanding machines and their role*

An important aspect of outlining the socio-cognitive architecture of COHUMAIN is how humans understand and interact with machines. We found that in human–human teams, individuals with high levels of CSE beliefs contribute to the team's ability to learn about other agents and also form attachments with other agents, which in turn are important processes for the team to have sustained performance in the face of changes in complexity over time. In human–machine teams, it is possible that the understanding of who—including the machine—knows what in the team and who is best at what may even further increase since a composition where there are clear signals about where skills are situated in the team enhances the team's TMS, based on the signal-detection theory of team cognition (Aggarwal & Woolley, 2019). In a team with humans and machines together, these signals—especially due to their visibly distinct attributes—would be amplified, making their detection easier. This insight was also supported by the interviews. This is likely to be the case for both embodied AI, such as the unmanned drone and terrain vehicle, or disembodied AI, like the optimization system as long as team members have a clear idea of who belongs to the team. Given the arguments above, we advance *Research Recommendation 1: The study of human understanding of machine teammates should examine how much signal amplification is needed to ensure adequate TMS in a human–machine team*.

It is important to point out, though, that when humans and machines do not understand each other's capabilities and limits and are unable to anticipate when to override each other's actions, things can go catastrophically wrong in human–AI teaming. As a case in point, consider the infamous Boeing 737 Max 8 groundings. The maneuvering characteristics augmentation system (MCAS), which was a flight stabilization program intended to improve the handling of the aircraft, contrarily made pilots struggle to control the aircraft and played a central role in two fatal crashes killing a total of 346 passengers. It was after these tragic events that disabling MCAS had been incorporated into pilot training (Boeing 737 MAX Groundings, 2022).

Further, as pointed out by Carter-Browne et al. (2021), neglecting to design AIs with human interaction in mind has resulted in several failures. The researchers give several examples to illustrate this point: "(a) humans are less likely to adopt the new technology, especially if there are human requirements the technology cannot meet (Parasuraman & Riley, 1997), (b) new technology is often not used in the way it was intended (e.g., abuse of facial recognition software; Garvie, 2019), (c) new technology may not function according to its original design (e.g., when AIs fail to complete routine system updates, such as the patriot missiles, a function is impaired; US General Accounting Office (GAO), 1992), (d) failures also occur when humans have implicit expectations of the AI that are not made into explicit requirements (e.g., Yorktown Smart Ship failures, Slabodkin, 1998)."

The authors point out that not only should design focus on making AI safe, easy to use, reliable, and trustworthy (Shneiderman, 2020), but it should also focus on how these factors will impact the human counterpart (Carter-Browne et al., 2021). These examples show that the provision of transparent information can help humans to better understand AI's actions and predict its future behavior more accurately (Riedl, 2019; Williams, Fiore, & Jentsch, 2022). A vast body of research supports the idea that transparency plays an important role in human–AI

interaction as it leads to increased situational awareness and improved decision accuracy of the human user (Bhaskara et al., 2021; Roth et al., 2020). At the same time, transparency has also been related to increased workload for the human agent, such that knowledge of what the machine is able to do can also interfere with the execution of the task (Guznov et al., 2020; Selkowitz et al., 2017). The interviews suggested that humans are also less accepting of errors made by AI. The question that emerges here is to what extent should human agents know and learn about the capabilities of non-human agents such that they can develop effective TMS that facilitate coordination and team performance, without generating dysfunctional levels of workload. We advance hence the following direction for future research; *Research Recommendation 2: The study of human misunderstanding of machine teammates should examine the minimal level of understanding needed to ensure effective coordination while keeping workload at low levels*.

In our human–human study, we found that CSE beliefs lead to enhanced collective team identification, which translates further into team performance. We base our results on the social identity theory indicating that individuals define themselves in terms of their membership collectively, that is, their collective identity (Ashforth & Mael, 1989). Human–machine teams are hybrid teams that incorporate both humans and AI agents, and the main question that emerges here is how collective team identification would look like in such teams and which factors contribute to its emergence. Previous research indicates that perceptions of similarity among team members play a key role in the emergence of collective attachments (Van Knippenberg, Dawson, West, & Homan, 2011). In a hybrid team, when AI agents attain a certain level of intelligence, for example, humans tend to judge them much the same way they do their fellow humans, seeking human likeness where it may not exist (Groom & Nass, 2007; Nass, Moon, Fogg, Reeves, & Dryer, 1995), and has been referred to as "teammate-likeness" (Stowers, Brady, MacLellan, Wohleber, & Salas, 2021). Human-likeness or anthropomorphism has been found to be essential for the development of emotional trust (Glikson & Woolley, 2020).

Perceptions of dissimilarities (such as surface-level or visual) on the other hand may lead to faultline or ingroup–outgroup divide. Existing research has identified that one important factor in creating these faultlines is the salience of social categories (Meyer, Shemla, & Schermuly, 2011; Van Knippenberg et al., 2011). One way to counter this divide would be the emphasis on a superordinate identity or when the members of a team share the sense of belonging to a higher-level unit (Argote & Kane, 2009; Gaertner & Dovidio, 2014). Existing research shows even when knowledge was less demonstrable, it was more likely to transfer between groups that shared a superordinate identity, compared to groups that did not share such an identity (Kane, 2010). Given that both TMS and collective team identification are important for teams to act in collectively intelligent ways, we believe that this will be an important direction for future research. We advance hence the third research recommendation. *Research recommendation 3: The study of human–machine teams should examine how collective team identification looks like in such hybrid teams and which factors contribute to its emergence (e.g., level of perceived similarity between human and non-human agents)*.

*5.2.2. Machines understanding humans and their role*

There is much promise in human–AI teaming given the complementary set of skills both bring to solve complex problems. Psychologists and engineers have long explored the use of machines to augment and improve human task performance (Dekker & Woods, 2002; Fitts, 1951; Stowers et al., 2021). While in the last decade, the conversation has shifted from machines as tools to machines as teammates (Phillips et al., 2011; Seeber et al., 2020; Stowers et al., 2021), still little is known about human–machine collaborations where intelligent technology takes on the role of a teammate (for exceptions, see Moradi, Moradi, Bayat, & Toosi, 2019; Song et al., 2022). Accordingly, researchers have encouraged redressing this gap by developing innovative and plausible models for human–machine teaming (Fiore, Bracken, Demir, Freeman, & Lewis, 2021). These models may include how social cognitive processes can be implemented to create agent architectures capable of monitoring and intervening in teamwork (Fiore et al., 2021).

Since humans are boundedly rational (Simon, 1956) and act according to several cognitive biases, which deviate from rationality (Kahneman, Slovic, Slovic, & Tversky, 1982), research on how machines can provide advice to human agents based on optimal behavior and rationality with fewer biases could be an important advance in human–machine collective intelligence (Nguyen & Gonzalez, 2021). Yet a barrier that might prevent this potential effective collaboration is the lack of a complete understanding of how humans operate, including their theory of mind (Oguntola, Hughes, & Sycara, 2021; Williams et al., 2022). And to develop high levels of TMS in a multi-agent team, AI needs to be able to anticipate what team members know and how they behave. As pointed out by researchers, determining what it means for AI to be a team member within a human–AI system is underrepresented in the literature (Carter-Browne et al., 2021; You & Robert, 2017).

A major challenge for research in AI is to develop systems that can infer the goals, beliefs, and intentions of others (Nguyen & Gonzalez, 2021). And research is starting to make great advances toward developing cognitive models, such as using instance-based learning theory, which generates a Theory of Mind from the observation of other agents' behavior making reasonable assumptions about human cognition (Nguyen & Gonzalez, 2021). It is important to note that in AI developing a theory about the human mind, we need to include not only implicit forms of mentalizing in humans involved in perspective-taking and tracking the intentional states of others but also an explicit form of mentalizing that appears to be unique to humans: the ability to reflect on one's action and to think about one's own thoughts (Frith & Frith, 2012). Hence, humans are able to monitor others' mental states and also have an awareness of their own mental states, both highly important in guiding human behavior. Given the elements presented above, we advance the fourth research recommendation. *Research recommendation 4: The study of human–machine teams should examine the extent to which machines as teammates can alleviate irrational decision-making and biases and how a machine's understanding of goals, beliefs, and intentions of others can be developed.*

AI needs to be able to predict that in working with human teammates, their self-beliefs–and not just actual skills–play an important role in determining outcomes (Diseth et al., 2014; Levpušček & Zupančič, 2009; Multon et al., 1991; Valentine et al., 2004). Further, there is evidence that humans both in individual and collective settings underestimate or

overestimate their performance, leading to beliefs that do not align with their actual performance levels (Bian, Leslie, & Cimpian, 2017; Meslec & Aggarwal, 2018). In the present research, we see that the team's creative self-belief composition plays an important role in determining both teamwork and taskwork. Hence, an AI member's ability to anticipate or investigate who "believes" what, or more specifically who has what self-beliefs, could be a very fruitful question for future research in COHUMAIN.

Clues from existing literature point to the behaviors that are related to an individual's CSE, and hence can be used to infer CSE in humans, including the display of confidence in one's ability to perform complex tasks or for tasks that require everyday creativity as compared to specialized forms of creativity such as scientific or musical creativity. It would also be evident in the display of an ability to respond to challenges in the environment and assure resilience (Cropley, 1990; Flach, 1990; Richards & Kinney, 1990). Nonetheless, inferring CSE levels purely by observing behavior alone might lead to inferential errors since other factors might also result in similar behaviors. Hence, we recommend that machines may learn about an individual's CSE much like researchers do—that is, asking agents about their self-beliefs in both explicit and implicit ways. In addition, research suggests that the future of human–AI teaming could include informal interviews or conversational exchanges, but in the absence of functional artificial social intelligence, the administration of a subset of surveys and measures may suffice (Bendell, Williams, Fiore, & Jentsch, 2021).

The question of how AI could help facilitate the formation of collective team identification also remains open. Our findings report that higher levels of self-efficacy are associated with the development of collective team identification. Having high levels of CSE composition in the team does not imply that all members have to be high in it, but that at least some members are, elevating the team average (Neuman et al., 1999). Research shows that training can be used to increase the CSE of an individual (Mathisen & Bronnick, 2009). AI could facilitate this process by helping members form higher levels of self-efficacy. As noted by Gupta and Woolley (2021), AI can take many roles in human–AI teaming such as (a) assistive AI used to augment an individual's capabilities, (b) coach AI used to predict and nudge coordination behaviors, and (c) diagnostic AI, which can monitor collective effort, strategy, and skill use. We believe that by playing one or multiple of these roles, especially the assistive AI role, AI can help human team members develop a higher level of self-efficacy and hence facilitate teamwork and taskwork that ultimately foster COHUMAIN. Given the above, we advance the last research recommendation. *Research recommendation 5: The study of human–machine teams should examine how AI agents can learn about the beliefs of other team members, and how this knowledge can facilitate interaction and coordination in human–AI teams*.

## 6. Conclusion

In this paper, we add to the understanding of the socio-cognitive architecture of COHUMAIN by studying the impact of humans' metacognition—specifically self-beliefs—on the collective team socio-cognitive states involving teamwork (i.e., learning about other agents in the team and forming social attachments with these agents) as well as taskwork. The research

shows that these self-beliefs at the team level have important implications for collective intelligence through the team states of TMS and collective team identification. Further, we provide implications of this research for human–AI collective intelligence by elaborating on how humans might understand machines and team with them, and how AI might understand humans in collaborating with them. We also provide future directions for research in this promising area for understanding and fostering COHUMAIN.

## Acknowledgments

We thank Rafael Goldschmidt for his advice and helpful comments, especially on the statistical analyses. We thank Fatih Akar, Aylin Erdil, İlker Şahin, and their colleagues for their invaluable practical insights.

## Notes

1 More information about the details of the scoring system can be found at https://www.bsg-online.com/help/instructors/GettingStarted/PerformanceScoring.html.
2 Bootstrapping is the creation of simulated samples from a dataset by drawing random samples with replacements from a dataset. It is a non-parametric test as it does not make any assumptions about the sample distributions when calculating standard errors and constructing confidence intervals. Thus, it is robust to non-normality problems (if any)—including problems generated in the distribution of the multiplication required in mediation testing (Zhao et al., 2010). We use 5000 bootstrapped samples following the suggestion of Preacher and Hayes (2004). Due to the simulated samples, bootstrapping is a more powerful technique, and thus we use bootstrapping of standard errors in regression analyses as well.
3 Brisa, a joint venture between Bridgestone Corporation and Sabanci Holding, is a leading tire manufacturer and exporter in Turkey (https://www.brisa.com.tr/). Kordsa is a global company operating in tire and construction reinforcement and composite markets (https://www.kordsa.com/en/).

## References

Aggarwal, I., & Woolley, A. W. (2019). Team creativity, cognition, and cognitive style diversity. *Management Science*, *65*(4), 1586–1599. https://doi.org/10.1287/mnsc.2017.3001

Andrews, R. W., Lilly, J. M., Srivastava, D., & Feigh, K. M. (2022). The role of shared mental models in human-AI teams: A theoretical review. *Theoretical Issues in Ergonomics Science*, *24*(2), 129–175. https://doi.org/10.1080/1463922X.2022.2061080

Antonakis, J. (2017). On doing better science: From thrill of discovery to policy implications. *The Leadership Quarterly*, *28*(1), 5–21. https://doi.org/10.1016/j.leaqua.2017.01.006

Argote, L., & Kane, A. A. (2009). Superordinate identity and knowledge creation and transfer in organizations. In N. J. Foss & S. Michailova (Eds.), *Knowledge governance* (pp. 166–190). Oxford, England: Oxford University Press.

Argote, L., & Ren, Y. (2012). Transactive memory systems: A microfoundation of dynamic capabilities: Transactive memory systems. *Journal of Management Studies*, *49*(8), 1375–1382. https://doi.org/10.1111/j.1467-6486.2012.01077.x

Ashforth, B. E., & Mael, F. (1989). Social identity theory and the organization. *Academy of Management Review*, *14*(1), 20–39.

Ateş, N. Y., Tarakci, M., Porck, J. P., van Knippenberg, D., & Groenen, P. J. F. (2020). The dark side of visionary leadership in strategy implementation: Strategic alignment, strategic consensus, and commitment. *Journal of Management*, *46*(5), 637–665. https://doi.org/10.1177/0149206318811567

Austin, J. R. (2003). Transactive memory in organizational groups: The effects of content, consensus, specialization, and accuracy on group performance. *Journal of Applied Psychology*, *88*(5), 866–878. https://doi.org/10.1037/0021-9010.88.5.866

Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review*, *84*(2), 191–215. https://doi.org/10.1037/0033-295X.84.2.191

Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York: W H Freeman/Times Books/Henry Holt & Co.

Bandura, A. (1999). Social cognitive theory: An agentic perspective. *Asian Journal of Social Psychology*, *2*(1), 21–41. https://doi.org/10.1111/1467-839X.00024

Bandura, A., & Locke, E. A. (2003). Negative self-efficacy and goal effects revisited. *Journal of Applied Psychology*, *88*(1), 87–99. https://doi.org/10.1037/0021-9010.88.1.87

Barrick, M. R., Stewart, G. L., Neubert, M. J., & Mount, M. K. (1998). Relating member ability and personality to work-team processes and team effectiveness. *Journal of Applied Psychology*, *83*, 377–391. https://doi.org/10.1037/0021-9010.83.3.377

Beghetto, R. A. (2006). Creative self-efficacy: Correlates in middle and secondary students. *Creativity Research Journal*, *18*(4), 447–457. https://doi.org/10.1207/s15326934crj1804_4

Beghetto, R. A., & Karwowski, M. (2017). Chapter 1—Toward untangling creative self-beliefs. In M. Karwowski & J. C. Kaufman (Eds.), *The creative self* (pp. 3–22). San Diego, CA: Academic Press. https://doi.org/10.1016/B978-0-12-809790-8.00001-7

Bell, S. T., Brown, S. G., Colaneri, A., & Outland, N. (2018). Team composition and the ABCs of teamwork. *American Psychologist*, *73*(4), 349–362. https://doi.org/10.1037/amp0000305

Bendell, R., Williams, J., Fiore, S. M., & Jentsch, F. (2021). Towards artificial social intelligence: Inherent features, individual differences, mental models, and theory of mind. *International Conference on Applied Human Factors and Ergonomics*, San Diego, CA (pp. 20–28).

Bezrukova, K., Jehn, K. A., Zanutto, E. L., & Thatcher, S. M. B. (2009). Do workgroup faultlines help or hurt? A moderated model of faultlines, team identification, and group performance. *Organization Science*, *20*(1), 35–50. https://doi.org/10.1287/orsc.1080.0379

Bhaskara, A., Duong, L., Brooks, J., Li, R., McInerney, R., Skinner, M., Pongracic, H., & Loft, S. (2021). Effect of automation transparency in the management of multiple unmanned vehicles. *Applied Ergonomics*, *90*, 103243. https://doi.org/10.1016/j.apergo.2020.103243

Bian, L., Leslie, S.-J., & Cimpian, A. (2017). Gender stereotypes about intellectual ability emerge early and influence children's interests. *Science*, *355*(6323), 389–391.

Bliese, P. D. (2000). Within-group agreement, non-independence, and reliability: Implications for data aggregation and analysis. In *Multilevel theory, research, and methods in organizations: Foundations, extensions, and new directions* (pp. 349–381). San Francisco, CA: Jossey-Bass.

Boeing 737 MAX groundings. (2022). In Wikipedia. Available at: https://en.wikipedia.org/w/index.php?title=Boeing_737_MAX_groundings&oldid=1090043330, accessed November 10, 2022.

Boies, K., Lvina, E., & Martens, M. L. (2011). Shared leadership and team performance in a business strategy simulation. *Journal of Personnel Psychology*, *9*(4), 195–202. https://doi.org/10.1027/1866-5888/a000021

Bonett, D., & Wright, T. (2014). Cronbach's alpha reliability: Interval estimation, hypothesis testing, and sample size planning. *Journal of Organizational Behavior*, *36*(1), 3–15. https://doi.org/10.1002/job.1960

Cabeza-Pullés, D., Gutierrez-Gutierrez, L. J., & Llorens-Montes, F. J. (2018). Drivers for performance in innovative research groups: The mediating role of transactive memory system. *BRQ Business Research Quarterly*, *21*(3), 180–194. https://doi.org/10.1016/j.brq.2018.03.002

Carmeli, A., & Schaubroeck, J. (2007). The influence of leaders' and other referents' normative expectations on individual involvement in creative work. *The Leadership Quarterly*, *18*(1), 35–48. https://doi.org/10.1016/j.leaqua.2006.11.001

Carter-Browne, B. M., Paletz, S. B., Campbell, S. G., Carraway, M. J., Vahlkamp, S. H., Schwartz, J., & O'Rourke, P. (2021). There is no "AI" in teams: A multidisciplinary framework for AIs to work in human teams. Applied Research Laboratory for Intelligence and Security (ARLIS) Report June 2021.

Cassidy, M. (2014, December 30). *Centaur chess shows power of teaming human and machine*. *HuffPost*. https://www.huffpost.com/entry/centaur-chess-shows-power_b_6383606

Chan, D. (1998). Functional relations among constructs in the same content domain at different levels of analysis: A typology of composition models. *Journal of Applied Psychology*, *83*(2), 234–246.

Chen, G., Gully, S. M., & Eden, D. (2001). Validation of a new general self-efficacy scale. *Organizational Research Methods*, *4*(1), 62–83. https://doi.org/10.1177/109442810141004

Christoffersen, K., & Woods, D. D. (2002). 1. How to make automated systems team players. In E. Salas (Ed.), *Advances in human performance and cognitive engineering research* (Vol. 2, pp. 1–12). Bingley, UK: Emerald Group Publishing Limited. https://doi.org/10.1016/S1479-3601(02)02003-9

Clayton, N. S., Dally, J. M., & Emery, N. J. (2007). Social cognition by food-caching corvids. The western scrub-jay as a natural psychologist. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *362*(1480), 507–522. https://doi.org/10.1098/rstb.2006.1992

Couzin, I. D. (2009). Collective cognition in animal groups. *Trends in Cognitive Sciences*, *13*(1), 36–43. https://doi.org/10.1016/j.tics.2008.10.002

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, *16*(3), 297–334. https://doi.org/10.1007/BF02310555

Cropley, A. J. (1990). Creativity and mental health in everyday life. *Creativity Research Journal*, *3*(3), 167–178. https://doi.org/10.1080/10400419009534351

Dafoe, A., Bachrach, Y., Hadfield, G., Horvitz, E., Larson, K., & Graepel, T. (2021). Cooperative AI: machines must learn to find common ground. *Nature*, *593*(7857), 33–36.

Dekker, S. W., & Woods, D. D. (2002). MABA-MABA or abracadabra? Progress on human–automation coordination. *Cognition, Technology & Work*, *4*(4), 240–244.

Diehl, M., & Stroebe, W. (1987). Productivity loss in brainstorming groups: Toward the solution of a riddle. *Journal of Personality and Social Psychology*, *53*(3), 497–509.

Diseth, Å., Meland, E., & Breidablik, H. J. (2014). Self-beliefs among students: Grade level and gender differences in self-esteem, self-efficacy and implicit theories of intelligence. *Learning and Individual Differences*, *35*, 1–8. https://doi.org/10.1016/j.lindif.2014.06.003

Driskell, J. E., Salas, E., & Hughes, S. (2010). Collective orientation and team performance: Development of an individual differences measure. *Human Factors*, *52*(2), 316–328. https://doi.org/10.1177/0018720809359522

Duffy, M. K., Scott, K. L., Shaw, J. D., Tepper, B. J., & Aquino, K. (2012). A social context model of envy and social undermining. *Academy of Management Journal*, *55*(3), 643–666. https://doi.org/10.5465/amj.2009.0804

Eden, D., & Aviram, A. (1993). Self-efficacy training to speed reemployment: Helping people to help themselves. *Journal of Applied Psychology*, *78*(3), 352–360. https://doi.org/10.1037/0021-9010.78.3.352

Engel, D., Woolley, A. W., Jing, L. X., Chabris, C. F., & Malone, T. W. (2014). Reading the mind in the eyes or reading between the lines? Theory of Mind predicts collective intelligence equally well online and face-to-face. *PLOS ONE*, *9*(12), e115212. https://doi.org/10.1371/journal.pone.0115212

Farmer, S. M., & Tierney, P. (2017). Chapter 2— Considering creative self-efficacy: Its current state and ideas for future inquiry. In M. Karwowski & J. C. Kaufman (Eds.), *The creative self* (pp. 23–47). Saint Louis, MO: Academic Press. https://doi.org/10.1016/B978-0-12-809790-8.00002-9

Fiore, S. M., Bracken, B., Demir, M., Freeman, J., & Lewis, M. (2021). Transdisciplinary team research to develop theory of mind in human-AI teams panelists. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *65*(1), 1605–1609.

Fitts, P. M. (1951). Human engineering for an effective air-navigation and traffic-control system. Division of National Research Council.

Flach, F. (1990). The resilience hypothesis and posttraumatic stress disorder. In M. E. Wolf & A. D. Mosnaim (Eds.), *Posttraumatic stress disorder: Etiology, phenomenology, and treatment* (pp. 37–45). San Francisco, CA: American Psychiatric Association.

Foote, N. N. (1951). Identification as the basis for a theory of motivation. *American Sociological Review*, *16*(1), 14–21. https://doi.org/10.2307/2087964

Frith, C. D., & Frith, U. (2012). Mechanisms of social cognition. *Annual Review of Psychology*, *63*, 287–313.

Gaertner, S. L., & Dovidio, J. F. (2014). *Reducing intergroup bias: The common ingroup identity model*. New York: Psychology Press.

Garvie, C. (2019). Garbage in. Garbage out. Face recognition on flawed data. Available at: https://www.flawedfacedata.com, accessed on November 19, 2022.

Gervits, F., Thurston, D., Thielstrom, R., Fong, T., Pham, Q., & Scheutz, M. (2020, May). Toward Genuine Robot Teammates: Improving Human-Robot Team Performance Using Robot Shared Mental Models. In *Aamas* (pp. 429–437).

Gilson, L. L., & Shalley, C. E. (2004). A little creativity goes a long way: An examination of teams' engagement in creative processes. *Journal of Management*, *30*(4), 453–470. https://doi.org/10.1016/j.jm.2003.07.001

Gino, F., Argote, L., Miron-Spektor, E., & Todorova, G. (2010). First, get your feet wet: The effects of learning from direct and indirect experience on team creativity. *Organizational Behavior and Human Decision Processes*, *111*(2), 102–115. https://doi.org/10.1016/j.obhdp.2009.11.002

Gist, M. E., & Mitchell, T. R. (1992). Self-efficacy: A theoretical analysis of its determinants and malleability. *Academy of Management Review*, *17*(2), 183–211. https://doi.org/10.5465/amr.1992.4279530

Glikson, E., & Woolley, A. W. (2020). Human trust in artificial intelligence: Review of empirical research. *Academy of Management Annals*, *14*(2), 627–660. https://doi.org/10.5465/annals.2018.0057

Gong, Y., Huang, J.-C., & Farh, J.-L. (2009). Employee learning orientation, transformational leadership, and employee creativity: The mediating role of employee creative self-efficacy. *Academy of Management Journal*, *52*(4), 765–778. https://doi.org/10.5465/amj.2009.43670890

Gopnik, A., & Astington, J. W. (1988). Children's understanding of representational change and its relation to the understanding of false belief and the appearance-reality distinction. *Child Development*, *59*(1), 26–37.

Groom, V., & Nass, C. (2007). Can robots be teammates?: Benchmarks in human–robot teams. *Interaction Studies*, *8*(3), 483–500.

Gupta, P., & Woolley, A. W. (2021). Articulating the role of artificial intelligence in collective intelligence: A transactive systems framework. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *65*(1), 670–674. https://doi.org/10.1177/1071181321651354c

Guznov, S., Lyons, J., Pfahler, M., Heironimus, A., Woolley, M., Friedman, J., & Neimeier, A. (2020). Robot transparency and team orientation effects on human–robot teaming. *International Journal of Human–Computer Interaction*, *36*(7), 650–660. https://doi.org/10.1080/10447318.2019.1676519

Guzzo, R. A., & Dickson, M. W. (1996). Teams in organizations: Recent research on performance and effectiveness. *Annual Review of Psychology*, *47*(1), 307–338. https://doi.org/10.1146/annurev.psych.47.1.307

Hackman, J. R., & Morris, C. G. (1975). Group tasks, group interaction process, and group performance effectiveness: A review and proposed integration. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 8, pp. 45–99). San Diego, CA: Academic Press. https://doi.org/10.1016/S0065-2601(08)60248-8

Hancock, P. A., Jagacinski, R. J., Parasuraman, R., Wickens, C. D., Wilson, G. F., & Kaber, D. B. (2013). Human-automation interaction research: Past, present, and future. *Ergonomics in Design*, *21*(2), 9–14. https://doi.org/10.1177/1064804613477099

Hanna, N., & Richards, D. (2018). The impact of multimodal communication on a shared mental model, trust, and commitment in human–intelligent virtual agent teams. *Multimodal Technologies and Interaction*, *2*(3), Article 3. https://doi.org/10.3390/mti2030048

Harrison, D. A., & Klein, K. J. (2007). What's the difference? Diversity constructs as separation, variety, or disparity in organizations. *Academy of Management Review*, *32*(4), 1199–1228. https://doi.org/10.5465/amr.2007.26586096

Hayes, A. F. (2017). *Introduction to mediation, moderation, and conditional process analysis, second edition: A regression-based approach*. New York: Guilford Publications.

Heavey, C., & Simsek, Z. (2017). Distributed cognition in top management teams and organizational ambidexterity: The influence of transactive memory systems. *Journal of Management*, *43*(3), 919–945. https://doi.org/10.1177/0149206314545652

Hoi, S. C. H., Sahoo, D., Lu, J., & Zhao, P. (2021). Online learning: A comprehensive survey. *Neurocomputing*, *459*, 249–289. Available at: https://doi.org/10.48550/arXiv.1802.02871

Hollingshead, A. B. (1998). Retrieval processes in transactive memory systems. *Journal of Personality and Social Psychology*, *74*(3), 659–671. https://doi.org/10.1037/0022-3514.74.3.659

Hollingshead, A. B. (2001). Cognitive interdependence and convergent expectations in transactive memory. *Journal of Personality and Social Psychology*, *81*(6), 1080–1089. https://doi.org/10.1037/0022-3514.81.6.1080

Hong, L., & Page, S. E. (2004). Groups of diverse problem solvers can outperform groups of high-ability problem solvers. *Proceedings of the National Academy of Sciences*, *101*(46), 16385–16389. https://doi.org/10.1073/pnas.0403723101

Huang, C. C., & Chen, P. (2018). Exploring the antecedents and consequences of the transactive memory system: An empirical analysis. *Journal of Knowledge Management*, *22*(1), 92–118. https://doi.org/10.1108/JKM-03-2017-0092

Ilgen, D. R., Hollenbeck, J. R., Johnson, M., & Jundt, D. (2005). Teams in organizations: From input-process-output models to IMOI models. *Annual Review of Psychology*, *56*(1), 517–543. https://doi.org/10.1146/annurev.psych.56.091103.070250

Ivcevic, Z. (2007). Artistic and everyday creativity: An act-frequency approach. *The Journal of Creative Behavior*, *41*(4), 271–290. https://doi.org/10.1002/j.2162-6057.2007.tb01074.x

James, L. R., Demaree, R. G., & Wolf, G. (1993). $r_{wg}$: An assessment of within-group interrater agreement. *Journal of Applied Psychology*, *78*(2), 306–309. https://doi.org/10.1037/0021-9010.78.2.306

Jaussi, K. S., & Randel, A. E. (2014). Where to look? Creative self-efficacy, knowledge retrieval, and incremental and radical creativity. *Creativity Research Journal*, *26*(4), 400–410. https://doi.org/10.1080/10400419.2014.961772

Johnson, H., & Avolio, B. J. (2019). Team psychological safety and conflict trajectories' effect on individual's team identification and satisfaction. *Group & Organization Management*, *44*(5), 843–873. https://doi.org/10.1177/1059601118767316

Johnson, M., Bradshaw, J. M., Feltovich, P. J., Jonker, C. M., van Riemsdijk, M. B., & Sierhuis, M. (2014). Coactive design: Designing support for interdependence in joint activity. *Journal of Human-Robot Interaction*, *3*(1), 43–69. https://doi.org/10.5898/JHRI.3.1.Johnson

Jonker, C., van Riemsdijk, M., & Vermeulen, B. (2011). Shared mental models: A conceptual analysis. *COIN 2010 International Workshops,—Coordination, Organizations, Institutions, and Norms in Agent Systems VI* (pp. 132–151). Toronto, Canada. https://doi.org/10.1007/978-3-642-21268-0_8

Kahneman, D., Slovic, S. P., Slovic, P., & Tversky, A. (1982). *Judgment under uncertainty: Heuristics and biases*. Cambridge: Cambridge University Press.

Kane, A. A. (2010). Unlocking knowledge transfer potential: Knowledge demonstrability and superordinate social identity. *Organization Science*, *21*(3), 643–660. https://doi.org/10.1287/orsc.1090.0469

Katz-Navon, T. Y., & Erez, M. (2005). When collective- and self-efficacy affect team performance: The role of task interdependence. *Small Group Research*, *36*(4), 437–465. https://doi.org/10.1177/1046496405275233

Kearney, E., Gebert, D., & Voelpel, S. C. (2009). When and how diversity benefits teams: The importance of team members' need for cognition. *Academy of Management Journal*, *52*(3), 581–598. https://doi.org/10.5465/amj.2009.41331431

Koole, S. L., & Lakens, D. (2012). Rewarding replications: A sure and simple way to improve psychological science. *Perspectives on Psychological Science*, *7*(6), 608–614. https://doi.org/10.1177/1745691612462586

Leslie, A. M. (1987). Pretense and representation: The origins of "theory of mind." *Psychological Review*, *94*(4), 412–426. https://doi.org/10.1037/0033-295X.94.4.412

Levpušček, M. P., & Zupančič, M. (2009). Math achievement in early adolescence: The role of parental involvement, teachers' behavior, and students' motivational beliefs about math. *The Journal of Early Adolescence*, *29*(4), 541–570. https://doi.org/10.1177/0272431608324189

Lewis, K. (2003). Measuring transactive memory systems in the field: Scale development and validation. *Journal of Applied Psychology*, *88*(4), 587–604. https://doi.org/10.1037/0021-9010.88.4.587

Lewis, K., & Herndon, B. (2011). Transactive memory systems: Current issues and future research directions. *Organization Science*, *22*(5), 1254–1265. https://doi.org/10.1287/orsc.1110.0647

Lewis, K., Lange, D., & Gillis, L. (2005). Transactive memory systems, learning, and learning transfer. *Organization Science*, *16*(6), 581–598. https://doi.org/10.1287/orsc.1050.0143

Liang, C., & Chang, C.-C. (2014). Predicting scientific imagination from the joint influences of intrinsic motivation, self-efficacy, agreeableness, and extraversion. *Learning and Individual Differences*, *31*, 36–42. https://doi.org/10.1016/j.lindif.2013.12.013

Liang, D. W., Moreland, R., & Argote, L. (1995). Group versus individual training and group performance: The mediating role of transactive memory. *Personality and Social Psychology Bulletin*, *21*(4), 384–393. https://doi.org/10.1177/0146167295214009

Lin, C.-P., He, H., Baruch, Y., & Ashforth, B. E. (2017). The effect of team affective tone on team performance: The roles of team identification and team cooperation. *Human Resource Management*, *56*(6), 931–952. https://doi.org/10.1002/hrm.21810

Locke, E. A., & Latham, G. P. (2002). Building a practically useful theory of goal setting and task motivation: A 35-year odyssey. *American Psychologist*, *57*(9), 705–717. https://doi.org/10.1037/0003-066X.57.9.705

Mael, F., & Ashforth, B. E. (1992). Alumni and their alma mater: A partial test of the reformulated model of organizational identification. *Journal of Organizational Behavior*, *13*(2), 103–123. https://doi.org/10.1002/job.4030130202

Makel, M. C., Plucker, J. A., & Hegarty, B. (2012). Replications in psychology research: How often do they really occur? *Perspectives on Psychological Science*, *7*(6), 537–542. https://doi.org/10.1177/1745691612460688

Malone, L., Bernstein, S., Atkins-Burnett, S., & Xue, Y. (2018). Psychometric analyses of child outcome measures with American Indian and Alaska native preschoolers: Initial evidence from AI/AN FACES 2015. Technical Report. OPRE Report 2018–21. In Office of Planning, Research and Evaluation. Office of Planning, Research and Evaluation. https://eric.ed.gov/?id=ED592763

Marks, M. A., Mathieu, J. E., & Zaccaro, S. J. (2001). A temporally based framework and taxonomy of team processes. *Academy of Management Review*, *26*(3), 356–376. https://doi.org/10.5465/amr.2001.4845785

Mathieu, J. E., Hollenbeck, J. R., van Knippenberg, D., & Ilgen, D. R. (2017). A century of work teams in the *Journal of Applied Psychology*. *Journal of Applied Psychology*, *102*(3), 452–467. https://doi.org/10.1037/apl0000128

Mathieu, J. E., & Rapp, T. L. (2009). Laying the foundation for successful team performance trajectories: The roles of team charters and performance strategies. *Journal of Applied Psychology*, *94*(1), 90–103.

Mathisen, G. E., & Bronnick, K. S. (2009). Creative self-efficacy: An intervention study. *International Journal of Educational Research*, *48*(1), 21–29. https://doi.org/10.1016/j.ijer.2009.02.009

Matthews, G., Lin, J., Panganiban, A. R., & Long, M. D. (2020). Individual differences in trust in autonomous robots: Implications for transparency. *IEEE Transactions on Human-Machine Systems*, *50*(3), 234–244. https://doi.org/10.1109/THMS.2019.2947592

Menold, J., & Jablokow, K. (2019). Exploring the effects of cognitive style diversity and self-efficacy beliefs on final design attributes in student design teams. *Design Studies*, *60*, 71–102. https://doi.org/10.1016/j.destud.2018.08.001

Meslec, N., & Aggarwal, I. (2018). Learning not to underestimate: Understanding the dynamics of women's underestimation in groups. *Team Performance Management*, *24*(7/8), 380–395.

Meyer, B., Shemla, M., & Schermuly, C. C. (2011). Social category salience moderates the effect of diversity faultlines on information elaboration. *Small Group Research*, *42*(3), 257–282.

Michael, L. A. H., Hou, S.-T., & Fan, H.-L. (2011). Creative self-efficacy and innovative behavior in a service setting: Optimism as a moderator. *The Journal of Creative Behavior*, *45*(4), 258–272. https://doi.org/10.1002/j.2162-6057.2011.tb01430.x

Miller, K. D., Choi, S., & Pentland, B. T. (2014). The role of transactive memory in the formation of organizational routines. *Strategic Organization*, *12*(2), 109–133. https://doi.org/10.1177/1476127014521609

Moradi, M., Moradi, M., Bayat, F., & Toosi, A. N. (2019). Collective hybrid intelligence: Towards a conceptual framework. *International Journal of Crowd Science*, *3*(2), 198–220. https://doi.org/10.1108/ijcs-03-2019-0012

Moreland, R. L., Argote, L., & Krishnan, R. (1996). Socially shared cognition at work: Transactive memory and group performance. In J. L. Nye & A. M. Brower (Eds.), *What's social about social cognition? Research on socially shared cognition in small groups* (pp. 57–84). Thousand Oaks, CA: Sage Publications, Inc. https://doi.org/10.4135/9781483327648.n3

Multon, K. D., Brown, S. D., & Lent, R. W. (1991). Relation of self-efficacy beliefs to academic outcomes: A meta-analytic investigation. *Journal of Counseling Psychology*, *38*(1), 30–38. https://doi.org/10.1037/0022-0167.38.1.30

Nass, C., Moon, Y., Fogg, B. J., Reeves, B., & Dryer, D. C. (1995). Can computer personalities be human personalities? *International Journal of Human-Computer Studies*, *43*(2), 223–239.

Neuman, G. A., Wagner, S. H., & Christiansen, N. D. (1999). The relationship between work-team personality composition and the job performance of teams. *Group & Organization Management*, *24*(1), 28–45. https://doi.org/10.1177/1059601199241003

Nguyen, T. N., & Gonzalez, C. (2021). Theory of mind from observation in cognitive models and humans. *Topics in Cognitive Science*, *14*(4), 665–686. https://doi.org/10.1111/tops.12553

Oguntola, I., Hughes, D., & Sycara, K. (2021). Deep interpretable models of Theory of Mind. *2021 30th IEEE International Conference on Robot & Human Interactive Communication (RO-MAN)*, Vancouver, Canada (pp. 657–664).

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

Pearsall, M. J., & Venkataramani, V. (2015). Overcoming asymmetric goals in teams: The interactive roles of team learning orientation and team identification. *Journal of Applied Psychology*, *100*(3), 735–748. https://doi.org/10.1037/a0038315

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

Preacher, K. J., & Hayes, A. F. (2004). SPSS and SAS procedures for estimating indirect effects in simple mediation models. *Behavior Research Methods, Instruments, & Computers*, *36*(4), 717–731. https://doi.org/10.3758/BF03206553

Preacher, K. J., & Hayes, A. F. (2008). Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. *Behavior Research Methods*, *40*(3), 879–891. https://doi.org/10.3758/BRM.40.3.879

Rabinowitz, N. C., Perbet, F., Song, H. F., Zhang, C., Eslami, S. M. A., & Botvinick, M. M. (2018). Machine theory of mind. Proceedings of the 35th International Conference on Machine Learning, in Proceedings of Machine Learning Research 80:4218–4227. Available from https://proceedings.mlr.press/v80/rabinowitz18a.html

Ren, Y., Carley, K. M., & Argote, L. (2006). The contingent effects of transactive memory: When is it more beneficial to know what others know? *Management Science*, *52*(5), 671–682. https://doi.org/10.1287/mnsc.1050.0496

Richards, R., & Kinney, D. K. (1990). Mood swings and creativity. *Creativity Research Journal*, *3*(3), 202–217. https://doi.org/10.1080/10400419009534353

Richards, R., Kinney, D. K., Lunde, I., Benet, M., & Merzel, A. P. C. (1988). Creativity in manic-depressives, cyclothymes, their normal relatives, and control subjects. *Journal of Abnormal Psychology*, *97*(3), 281–288. https://doi.org/10.1037/0021-843X.97.3.281

Riedl, M. O. (2019). Human-centered artificial intelligence and machine learning. *Human Behavior and Emerging Technologies*, *1*(1), 33–36. https://doi.org/10.1002/hbe2.117

Roth, G., Schulte, A., Schmitt, F., & Brand, Y. (2020). Transparency for a workload-adaptive cognitive agent in a manned–unmanned teaming application. *IEEE Transactions on Human-Machine Systems*, *50*(3), 225–233. https://doi.org/10.1109/THMS.2019.2914667

Rouse, W. B., & Morris, N. M. (1986). On looking into the black box: Prospects and limits in the search for mental models. *Psychological Bulletin*, *100*, 349–363. https://doi.org/10.1037/0033-2909.100.3.349

Russell, S. (2019). *Human compatible: Artificial intelligence and the problem of control*. New York: Penguin Books.

Salas, E., Sims, D. E., & Burke, C. S. (2005). Is there a "Big Five" in teamwork? *Small Group Research*, *36*(5), 555–599. https://doi.org/10.1177/1046496405277134

Sangsuk, P., & Siriparp, T. (2015). Confirmatory factor analysis of a scale measuring creative self-efficacy of undergraduate students. *Procedia—Social and Behavioral Sciences*, *171*, 1340–1344. https://doi.org/10.1016/j.sbspro.2015.01.251

Seeber, I., Bittner, E., Briggs, R. O., de Vreede, T., de Vreede, G.-J., Elkins, A., Maier, R., Merz, A. B., Oeste-Reiß, S., Randrup, N., Schwabe, G., & Söllner, M. (2020). Machines as teammates: A research agenda on AI in team collaboration. *Information & Management*, *57*(2), 103174. https://doi.org/10.1016/j.im.2019.103174

Selkowitz, A. R., Lakhmani, S. G., & Chen, J. Y. C. (2017). Using agent transparency to support situation awareness of the Autonomous Squad Member. *Cognitive Systems Research*, *46*, 13–25. https://doi.org/10.1016/j.cogsys.2017.02.003

Shane, J. (2019). *You look like a thing and I love you: How artificial intelligence works and why it's making the world a weirder place (Illustrated ed.)*. New York: Voracious.

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

Simmons, A. L., Payne, S. C., & Pariyothorn, M. M. (2014). The role of means efficacy when predicting creative performance. *Creativity Research Journal*, *26*(1), 53–61. https://doi.org/10.1080/10400419.2014.873667

Simon, H. A. (1956). Rational choice and the structure of the environment. *Psychological Review*, *63*(2), 129–138.

Slabodkin, G. (1998, July 13). Software glitches leave Navy Smart Ship dead in the water. *Government Computer News*.

Solansky, S. T. (2011). Team identification: A determining factor of performance. *Journal of Managerial Psychology*, *26*(3), 247–258. https://doi.org/10.1108/02683941111112677

Song, B., Zurita, N. F. S., Gyory, J. T., Zhang, G., McComb, C., Cagan, J., Stump, G., Martin, J., Miller, S., & Balon, C. (2022). Decoding the agility of artificial intelligence-assisted human design teams. *Design Studies*, *79*, 101094.

Stowers, K., Brady, L. L., MacLellan, C., Wohleber, R., & Salas, E. (2021). Improving teamwork competencies in human-machine teams: Perspectives from team science. *Frontiers in Psychology*, *12*, 590290. https://www.frontiersin.org/article/10.3389/fpsyg.2021.590290

Stowers, K., Kasdaglis, N., Rupp, M. A., Newton, O. B., Chen, J. Y. C., & Barnes, M. J. (2020). The IMPACT of agent transparency on human performance. *IEEE Transactions on Human-Machine Systems*, *50*(3), 245–253. https://doi.org/10.1109/THMS.2020.2978041

Strohkorb, S., Fukuto, E., Warren, N., Taylor, C., Berry, B., & Scassellati, B. (2016). Improving human-human collaboration between children with a social robot. *2016 25th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*, New York, NY (pp. 551–556). https://doi.org/10.1109/ROMAN.2016.7745172

Sutton, W. A., McDonald, M. A., Milne, G. R., & Cimperman, J. (1997). Creating and fostering fan identification in professional sports. *Sport Marketing Quarterly*, *6*, 15–22.

Tajfel, H. E. (1978). *Differentiation between social groups: Studies in the social psychology of intergroup relations*. San Diego, CA: Academic Press.

Tasa, K., Taggar, S., & Seijts, G. H. (2007). The development of collective efficacy in teams: A multilevel and longitudinal perspective. *Journal of Applied Psychology*, *92*(1), 17–27. https://doi.org/10.1037/0021-9010.92.1.17

Tausch, A., Kluge, A., & Adolph, L. (2020). Psychological effects of the allocation process in human–robot interaction—A model for research on ad hoc task allocation. *Frontiers in Psychology*, *11*, 564672. https://doi.org/10.3389/fpsyg.2020.564672

Thompson, J. A. A., & Stappenbeck, G. J. (1999). *Player's manual to accompany the business strategy game: A global industry simulation* (6th ed.). New York: Irwin McGraw-Hill.

Tierney, P., & Farmer, S. M. (2002). Creative self-efficacy: Its potential antecedents and relationship to creative performance. *Academy of Management Journal*, *45*(6), 1137–1148. https://doi.org/10.5465/3069429

Torrance, E. P. (1988). The nature of creativity as manifest in its testing. In R. J. Sternberg (Ed.), *The nature of creativity* (pp. 43–75). New York: Cambridge University Press.

US General Accounting Office (GAO). (1992, February 4). Patriot missile defense: Software problem led to system failure at Dhahran, Saudi Arabia. *General Accounting Office*.

Valentine, J. C., DuBois, D. L., & Cooper, H. (2004). The relation between self-beliefs and academic achievement: A meta-analytic review. *Educational Psychologist*, *39*(2), 111–133. https://doi.org/10.1207/s15326985ep3902_3

van de Merwe, K., Oprins, E., Eriksson, F., & van der Plaat, A. (2012). The influence of automation support on performance, workload, and situation awareness of air traffic controllers. *The International Journal of Aviation Psychology*, *22*(2), 120–143. https://doi.org/10.1080/10508414.2012.663241

Van Der Vegt, G. S., & Bunderson, J. S. (2005). Learning and performance in multidisciplinary teams: The importance of collective team identification. *Academy of Management Journal*, *48*(3), 532–547. https://doi.org/10.5465/amj.2005.17407918

Van Knippenberg, D., Dawson, J. F., West, M. A., & Homan, A. C. (2011). Diversity faultlines, shared objectives, and top management team performance. *Human Relations*, *64*(3), 307–336.

van Knippenberg, D., & Schippers, M. C. (2007). Work group diversity. *Annual Review of Psychology*, *58*(1), 515–541. https://doi.org/10.1146/annurev.psych.58.110405.085546

Voas, J. (2004). Software's secret sauce: The "-ilities." *IEEE Software*, *21*(6), 14–15. https://doi.org/10.1109/MS.2004.54

Wegner, D. M. (1987). Transactive memory: A contemporary analysis of the group mind. In B. Mullen & G. R. Goethals (Eds.), *Theories of group behavior* (pp. 185–208). Cham: Springer. https://doi.org/10.1007/978-1-4612-4634-3_9

Wellman, H. M. (1992). *The child's theory of mind. (pp. xiii, 358)*. Cambridge: The MIT Press.

Williams, J., Fiore, S. M., & Jentsch, F. (2022). Supporting artificial social intelligence with Theory of Mind. *Frontiers in Artificial Intelligence*, *5*, 750763.

Wood, R., & Bandura, A. (1989). Social cognitive theory of organizational management. *Academy of Management Review*, *14*(3), 361–384. https://doi.org/10.5465/amr.1989.4279067

Woolley, A. W., Chabris, C. F., Pentland, A., Hashmi, N., & Malone, T. W. (2010). Evidence for a collective intelligence factor in the performance of human groups. *Science*, *330*(6004), 686–688. https://doi.org/10.1126/science.1193147

Wuchty, S., Jones, B. F., & Uzzi, B. (2007). The increasing dominance of teams in production of knowledge. *Science*, *316*(5827), 1036–1039. https://doi.org/10.1126/science.1136099

Yin, R. K. (2013). Validity and generalization in future case study evaluations. *Evaluation*, *19*(3), 321–332.

You, S., & Robert, L. (2017). Teaming up with robots: An IMOI (inputs-mediators-outputs-inputs) framework of human–robot teamwork. *International Journal of Robotic Engineering*, *2*(1). http://deepblue.lib.umich.edu/handle/2027.42/138192

Zhang, Z. X., Hempel, P. S., Han, Y. L., & Tjosvold, D. (2007). Transactive memory system links work team characteristics and performance. *Journal of Applied Psychology*, *92*(6), 1722–1730.

Zhao, X., Lynch, Jr J. G., & Chen, Q. (2010). Reconsidering Baron and Kenny: Myths and truths about mediation analysis. *Journal of consumer research*, *37*(2), 197–206.