Would you trust an AI team member? Team trust in human–AI teams

Georganta, E.; Ulfert, A.-S.

DOI
10.1111/joop.12504

Publication date
2024

Document Version
Final published version

Published in
Journal of Occupational and Organizational Psychology

Received: 4 June 2021 | Accepted: 20 March 2024

DOI: 10.1111/joop.12504

REGISTERED REPORT STAGE 2

the british psychological society
promoting excellence in psychology

Would you trust an AI team member? Team trust in human–AI teams

Eleni Georganta[1] | **Anna-Sophie Ulfert**[2]

[1]Department of Work and Organizational Psychology, Faculty of Social and Behavioural Sciences, University of Amsterdam, Amsterdam, The Netherlands

[2]Department of Industrial Engineering & Innovation Sciences, Human Performance Group, Eindhoven University of Technology, Eindhoven, The Netherlands

Correspondence
Eleni Georganta, Department of Work and Organizational Psychology, Faculty of Social and Behavioural Sciences, University of Amsterdam, Nieuwe Achtergracht 129, 1018 WS Amsterdam, The Netherlands.
Email: e.georganta@uva.nl

Funding information
Society for Industrial and Organizational Psychology

Abstract
Given that AI is becoming an increasingly active participant in work teams, this study explores how team trust emerges in human–AI teams compared to human–human teams. Adopting a multi-level approach, we conducted two experimental studies ($N_{Study1} = 247$ two-member teams and $N_{Study2} = 106$ three-member teams, 828 individuals overall) and investigated how team composition (with AI or human team members) impacts interpersonal trust (affective and cognitive) and thus team trust. In two-member teams, interpersonal trust via perceived trustworthiness and not via perceived similarity was lower in human–AI teams compared to human–human teams. Exploratory findings showed that team identification and cognitive interpersonal trust were also lower in two-member human–AI teams than in human–human teams. However, in three-member teams, we found no differences in team trust via interpersonal trust between the two team types. Instead, our findings revealed that perceived trustworthiness and perceived similarity increased interpersonal trust and, in turn, team trust for both team types. With this research, we showed that underlying theories and evidence of team trust in human-only teams can enhance understanding of human–AI teams, though the results indicated certain differences that call for further investigation.

KEYWORDS
artificial intelligence, human–AI collaboration, team member, team trust, trustworthiness

Eleni Georganta and Anna-Sophie Ulfert contributed equally to this work and share first authorship.

Practitioner points

Human–AI teams develop team trust similar to human–human teams, therefore, perceiving AI teammates as trustworthy and similar to the rest of the team is crucial for fostering trust relationships.

In human–AI teams, the focus should be on identifying commonalities over differences between teammates to foster team trust and team identification. Increasing perceived similarity between humans and AI teammates supports building trust in dyadic and team relationships.

INTRODUCTION

Artificial intelligence (AI) has increasingly found its way into our homes, offices and cities, making it one of the most impactful recent technology trends, particularly in the workplace (Haynes, 2020). As a result of continuous development and improvement over recent years, intelligent systems now offer many opportunities for optimizing work processes (Kaplan & Haenlein, 2020). In various contexts, such as health care or customer support, AI is used for making better decisions (Davenport & Kalakota, 2019), handling calls (Leviathan & Matias, 2018) and sorting orders in smart warehouses (Mahroof, 2019). The future, however, lies in uniting the competencies of both humans and AI systems in a team (Johnson & Vera, 2019), which leads to a fundamental change of AI's role within these collaborations (Larson & DeChurch, 2020). By developing human–AI teams,[1] in which humans and AI are interdependent with regard to their tasks yet independent in their actions, human capabilities can be extended rather than merely replaced (Fompeyrine et al., 2021; Zhang et al., 2021). AI systems are becoming as agentic as their human colleagues, assuming the role of a team member rather than merely a tool (Larson & DeChurch, 2020). Examples of such AI[2] team members have already been introduced. For instance, in 2019, Charlie, an AI conversational agent, took part as a panellist in a 90-min discussion on AI-empowered learning (Cummings et al., 2021). Acting as a full team member in this panel and in later discussions, Charlie expressed its views on the future of AI and human collaboration, highlighting the importance of human–AI teaming (Schurr et al., 2020). The example of Charlie highlights how these new generations of AI systems function independently yet in full collaboration with their human team members in order to contribute to team outcomes.

Although these innovations in the fields of machine learning, natural language processing and robotics enter many industries and offer us a first glance into what the future of teamwork could look like, humans often have difficulties trusting such intelligent technologies (Georganta & Ulfert, 2024; Schaefer et al., 2016). Trust – the willingness to rely on others and share information (Breuer et al., 2016) – reflects an essential mechanism for effective collaboration between humans and AI (Schaefer et al., 2014; Sheridan, 2019). Empirical studies on trust in human–AI collaboration have highlighted the importance of trust between a single human and a single intelligent system (Rosenberg, 2016). Recently, research has also been extended to trust between single humans and multiple systems (e.g. when using multi-robot systems or swarm technologies), illustrating its positive impact on human–AI interaction (Mahani et al., 2020). Nevertheless, there is a paucity of evidence on trust between humans and AI within the context of teams (Freedy et al., 2007; Lee &

[1]We define human–AI teams as *a collection of human individuals and one or more AI technologies*. These teams interact virtually, 'possess one or more common goals, are brought together to perform organisationally relevant tasks, exhibit interdependencies with respect to workflow, goals, and outcomes, have different roles and responsibilities, and are together embedded in an encompassing organisational system, with boundaries and linkages to the broader system context and task environment' (Kozlowski & Ilgen, 2006, p. 79).

[2]The term 'artificial intelligence' subsumes a variety of systems that can take many forms (e.g., robots or agents), are able to process a wide range of tasks and vary in their degree of autonomy. The definition of an AI team member follows the above-mentioned definition of a human–AI team, as it is focused on the role of the AI in collaboration with humans, contributing to team performance. It is part of the joint activities that take place within the team, independent of its system characteristics (e.g., functionalities).

See, 2004; Parasuraman et al., 2008). Although the role of AI is changing from that of a technology, simply used as a tool, to becoming a full member of the team (Larson & DeChurch, 2020; van Wissen et al., 2012), it remains unknown how trust between human and AI team members develops. Trust relationships in human–AI teams involve multiple actors, including both human and AI team members: trust by the human towards the AI team member(s), trust by the human towards other human team members and most importantly, trust towards the team as a whole (Ulfert et al., 2023). To provide insight into trust in human–AI teams, it is necessary to move beyond trust as a bilateral interaction between one human and one or more AI team members and consider all the different trust relationships that exist within a team.

Trust in teams has been extensively investigated over the last several years, with growing evidence supporting its positive impact on effective teamwork (Dirks & de Jong, 2022). Specifically, research has shown that not only interpersonal trust towards other team members but particularly team trust – the shared psychological state towards the team as a whole – leads to better mutual understanding among team members, increased knowledge sharing and consequently, high team performance (Feitosa et al., 2020; Mathieu et al., 2008; McEvily et al., 2003). Especially in virtual contexts, where in-person interaction is not possible, team trust and the shared willingness to accept and rely on team members are fundamental for team success (Breuer et al., 2016). Similar to human–human teams (Salas et al., 2005), we argue that human–AI teams also need team trust to communicate, integrate information, coordinate and perform effectively. However, to our knowledge, evidence on team trust in human–AI teams is missing in the research literature (Georganta & Ulfert, 2024; Glikson & Woolley, 2020; Sanders et al., 2011; Ulfert et al., 2023).

In this article,[3] our goal is to provide the first empirical insights into the dynamic and emergent nature of team trust in human–AI teams. To do so, we take a multi-level approach that encompasses both individual- and team-level conceptualizations of trust (Costa et al., 2018). In line with work on the distinction between trust and trustworthiness (Colquitt et al., 2007) and with social categorization theory (Turner, 2010; Turner et al., 1987), we first investigate differences in perceived trustworthiness and perceived similarity when an AI team member is present compared to a human team member. Then, we explore the impact of these individual-level perceptions on interpersonal trust towards this team member. Progressing to the team level and building on studies of emergent team phenomena (Kozlowski & Klein, 2000), we further explore the impact of interpersonal trust (cognitive and affective) towards a team member (AI or human) on team trust (Glikson & Woolley, 2020; McAllister, 1995).

Given the complexity of the phenomenon of interest and the limited empirical evidence available, we conducted two experimental studies. Experiments reflect 'the most powerful technique available for demonstrating causal relationships between variables' (Jones, 1985, p. 282) and are especially suited to isolating and testing the effects of variables of interest (Greenberg & Tomlinson, 2004). At the same time, they allow us to create a scenario that reflects future human–AI teamwork, that is, team tasks in which relationships with AI systems as team members are moving away from simply task based to team based, in which a high level of interdependency among team members exists, and in which team communication and collective decisions are required to achieve superior outcomes. Furthermore, the experimental design enables us to distinguish between dyadic and team relationships and thereby systematically explore the emergence of team trust in human–AI teams compared to human-only teams, a process that is difficult to investigate in a field setting (Dirks & de Jong, 2022). Overall, our goal is to derive the first empirical insights into trust in human–AI teams that can be generalized across organizations (Highhouse, 2009).

With this research, we move beyond the one-sided investigation of human–AI interaction and answer recent calls to provide initial insights into trust emergence in human–AI teams (Glikson & Woolley, 2020). Specifically, we integrate work across disciplines and consider both the impact of intelligent systems on trust processes and dynamics in teams (Dirks & de Jong, 2022) as well as the role played by the team context and multi-level relationships in human–AI interactions (O'Neill et al., 2022). At the same time, we overcome common conceptualization and measurement issues in the trust literature (Feitosa et al., 2020) by clearly differentiating between perceived trustworthiness and trust (Zolin

[3]The originally reviewed and in principle accepted registered report of this article will be made available on OSF.

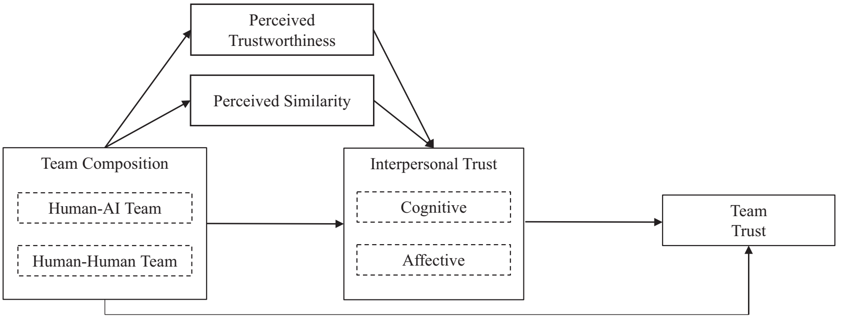

FIGURE 1 Illustration of theoretical model and expected relationships.

et al., 2004), as well as between levels of analysis (individual and team) and referents (single team member and team as a whole; Fulmer & Gelfand, 2012). Furthermore, we are among the first to provide empirical insights into how individual-level trust emerges to become team trust (Dirks & de Jong, 2022). Concretely, we investigate the mediating role of interpersonal trust in the relationship between team composition and team trust. We hope that this research can guide future investigations of trust in human–AI teams and the development of trustworthy AI team members. For our theoretical model and expected relationships, see Figure 1.

Team trust in human–AI teams

Team trust, which is defined as 'an emergent and dynamic shared state at the team-level whereby team members believe in one another's competence and are willing to be vulnerable beyond task-related issues' (Feitosa et al., 2020, p. 480), reflects one of the most important team properties that contribute to team success (De Jong et al., 2016). Teams with high levels of team trust work from a common baseline, reflect openly on environmental constraints and transparently monitor their objectives. These are all conditions that support reaching superior team outcomes (De Jong & Elfring, 2010). Furthermore, teams with high levels of team trust tend to agree upon norms regarding their behaviour and experience fewer interpersonal tensions (Duan et al., 2010). Through team trust, team members realize and accept that everyone is, in fact, looking out for each other and, respectively, for the good of the team. Given that team trust has been found to support team performance across a broad range of team types and contexts (for a meta-analysis, see De Jong et al., 2016), including in human–AI collaboration (Schaefer et al., 2014, 2016; Sheridan, 2019), we theorize that team trust reflects a key team capacity for various types of teams, including human–AI teams.

Recent work highlights those factors related to the characteristics of a team (e.g. team composition) that influence team trust (Costa et al., 2018). Specifically, prior evidence has shown that diverse team composition hinders trust development (Webber & Donahue, 2001) due to diverse teams experiencing greater difficulty communicating with each other and coordinating successfully than homogenous teams (Salas et al., 2008). Thus, teams composed of humans and AI team members are expected to encounter more difficulties in team trust development than teams composed of only human team members. Related evidence has shown that trust development is more difficult when initially collaborating with AI technologies compared to when collaborating with humans (Glikson & Woolley, 2020). Other work has also suggested that lacking experience in interacting with AI leads to difficulties in developing team trust in human–AI teams (Ulfert & Georganta, 2020). Trust not only influences how team members interpret the behaviour of other team members (Salas et al., 2005), but also how they perceive AI technologies and their actions (McKnight et al., 2011). Nevertheless, it remains unknown how trust towards AI team members develops and, most importantly, how this shapes trust towards the team as a whole (Dirks & de Jong, 2022).

To investigate how team trust emerges in human–AI teams, we explore how team composition (AI or human team member) impacts interpersonal trust towards a team member (human or AI) and, in turn, team trust. Hence, we first investigate whether team composition influences interpersonal trust based on the levels of perceived trustworthiness and the perceived similarity of a team member (AI or human). Second, we explore the mediating role of interpersonal trust in the relationship between team composition and team trust.

Effects of perceived trustworthiness on interpersonal trust

While team trust reflects a shared trust state towards the team as a whole (Fulmer & Gelfand, 2012) and involves interpersonal interactions and group dynamics (Costa et al., 2018), interpersonal trust refers to a specific other (e.g. a team member) and has two components, one cognitive and one affective (McAllister, 1995). Evidence from both human–human (Swift & Hwang, 2013) and human–AI interaction (Glikson & Woolley, 2020) suggests that both cognitive and affective interpersonal trust are essential for effective collaboration and success. On the one hand, cognitive trust is grounded in rational or calculation-based factors and the results of evaluating a team member's reliability and competences (Webber, 2008). On the other hand, affective trust is influenced by relational factors (e.g. emotions) and is impacted by closeness to a team member and social bonds (Komiak & Benbasat, 2006). Overall, research highlights that theories of interpersonal trust can also be applied to other contexts beyond human–human interaction (Wang et al., 2016).

Perceptions of a team member's level of trustworthiness impact how interpersonal trust develops. High perceived trustworthiness comes with positive expectations about other team member's intentions, motivation and behaviour, thus impacting collaboration within the team (Fulmer & Gelfand, 2012; Mayer et al., 1995; Ulfert & Georganta, 2020). Trustworthiness is defined by three dimensions: ability, integrity and benevolence. Specifically, *ability* describes how an individual team member's knowledge, competence and interpersonal skills to collaborate effectively are perceived by other team members (Colquitt et al., 2007). *Integrity* describes how a team member's credibility and consistency are perceived (Fulmer & Gelfand, 2012). *Benevolence* reflects perceptions of a team member's concern for the good of the team (Schoorman et al., 2007). Evidence has shown that positive expectations of one's trustworthiness, including all three dimensions, positively impact accepting one's vulnerability (Colquitt et al., 2007).

We argue that evaluating an AI team member's trustworthiness is more difficult because typical trust resources, such as familiarity and shared experiences, are limited. Furthermore, many underlying technical functions of the AI team member are not transparent to other team members (Glikson & Woolley, 2020; Hoff & Bashir, 2014), making the evaluation of its trustworthiness even more challenging. With increasing capabilities, the evaluation of an AI team member's reliability and consistency, which are decisive for its trustworthiness (Lee & See, 2004), becomes increasingly complex and, therefore, difficult (Glikson & Woolley, 2020). For instance, while an AI system is still in the learning phase of its development, it may initially exhibit behaviours different from what is expected (e.g. mislabelling pictures), leading to perceptions of low reliability (Glikson & Woolley, 2020). Thus, existing team members may become more uncertain and vulnerable with regard to judging the system compared to a human team member's competence. This is especially true since team members are required to assess whether the AI – both as a team member and as a technology – can act consistently, complete its tasks and display caring behaviours towards the team (Ulfert & Georganta, 2020). When human team members initially interact with an AI team member, they may be more cautious in evaluating its level of trustworthiness (e.g. its ability), leading to low interpersonal trust towards the AI team member. Caution and uncertainty can lead to a biased evaluation of AI technology and, as a result, to insufficient human–AI collaboration (Chavaillaz et al., 2019). Hence, we propose that having an AI compared to a human team member leads to lower levels of perceived trustworthiness and, in turn, to lower levels of interpersonal trust, both affective and cognitive. Thus, the following hypothesis is proposed:

Hypothesis 1. Perceived trustworthiness mediates the relationship between team composition (AI vs. human teammate) and interpersonal trust, both affective and cognitive.

Effects of perceived similarity on interpersonal trust

Interpersonal trust further depends on the interpersonal relationships that exist between team members (Costa et al., 2018). Specifically, having a common background and similar or shared experiences positively impacts interpersonal trust (Ulfert & Georganta, 2020). Trust develops more naturally when team members have similarities, which makes team members feel more comfortable when interacting with each other. Relatedly, dissimilarity among team members can have detrimental effects on interpersonal trust. For instance, evidence of demographic diversity in teams has shown that dissimilarities among team members lead to the attribution of negative characteristics and lower levels of trust (Krebs et al., 2006). Recognizing similarities leads to assuming similarities in values and beliefs (Tsui et al., 2002), thus resulting in a sense of comfort and a willingness to trust (Wildman et al., 2012; Williams et al., 2007). Recent findings on human–AI interaction have similarly shown that the more similar the AI systems were perceived to be in terms of values, the more they were trusted (Mehrotra et al., 2021).

According to social categorization theory (Turner, 2010; Turner et al., 1987), team members use cues and other stimuli to evaluate the similarity of others to themselves and the rest of the team. Depending on the degree of similarity, team members cognitively organize others as part of the same category (e.g. in-group) or a different one (e.g. out-group). This evaluation and, in turn, the in-group–out-group categorization shape the relationships among team members. Differentiating team members from the team and categorizing them as outsiders can have detrimental effects on trusting each other. Similar social rules and evaluation strategies apply when humans build trust relationships with intelligent technologies (Madhavan & Wiegmann, 2007). Thus, we expect that when interacting with an AI team member, initial cues will be used to assess the similarity of the AI member to oneself and the rest of the team. When doing so, we expect that an AI team member will be categorized as less similar to the rest of the team than a human team member. As previous research has shown, even when AI systems have the same roles as humans, AI systems are perceived differently. For instance, Mou and Xu (2017) showed that individuals tended to be less open and agreeable when initially interacting with an AI system rather than with a human (Mou & Xu, 2017). In another study (Yokoi et al., 2021), the findings similarly showed that trust was lower when participants imagined receiving a medical examination and being prescribed medicine by an AI system rather than a human. Thus, we argue that when comparing surface-level cues of the AI team member to existing schemata, team members cognitively organize the AI team member as part of the out-group – an intelligent technology – and not the in-group – the team itself (Turner, 2010; Turner et al., 1987). As a result, AI is mainly perceived as an autonomous technology, not as a team member. Thus, when working with an AI system rather than a human team member, we expect that this perception of low similarity and of being part of the out-group (Meyer et al., 2011) will result in negative bias and lower levels of interpersonal trust, both affective and cognitive (Brewer, 2017). Therefore, we advance the following hypothesis:

Hypothesis 2. Perceived similarity mediates the relationship between team composition (AI vs. human teammate) and interpersonal trust, both affective and cognitive.

Effects of interpersonal trust on team trust

Team composition can influence interpersonal trust between team members and shared trust at the team level (Costa et al., 2018). Research on trust in demographically diverse teams has shown that differences in composition – in terms of nationality and age – were related to lower interpersonal trust towards team members and, in turn, to lower team trust (Curşeu & Schruijer, 2010). Similarly, in cross-functional teams, a diverse team composition has been shown to prevent the development of interpersonal trust among team members, which was detrimental for shared trust towards the team as a whole (Newell et al., 2007). Diverse team composition can have a negative impact on the cognitive (e.g. reliability) and affective (e.g. closeness) evaluations of a team member, leading to low interpersonal trust and, consequently, to low shared trust (Fulmer & Gelfand, 2012).

Trust towards the team originates from individual members and then takes shape at the team level (Feitosa et al., 2020; Kozlowski & Klein, 2000). Studies on related shared states have shown that team-level phenomena, such as team learning (Kostopoulos et al., 2013) and team knowledge (Grand et al., 2016), emerge from the individual level. However, studies on trust have mainly focused on one level at a time (Dirks & de Jong, 2022). Nevertheless, the trust literature has highlighted that interpersonal trust towards a team member creates team-level trust (Costa et al., 2018). Building on this line of thinking about team emergence and on the fact that team factors such as composition have a strong impact on interpersonal trust and team trust (Breuer et al., 2016; De Jong et al., 2016), we propose that having an AI rather than a human team member leads to lower levels of interpersonal trust and, as a result, to lower levels of team trust. Thus, we advance the following hypothesis:

> **Hypothesis 3.** Interpersonal trust, both (H3a) cognitive and (H3b) affective, mediates the relationship between team composition (AI vs. human teammate) and team trust.

Overview of studies

To investigate the proposed relationships, we conducted two pre-registered experimental studies. In both studies, the participants worked collaboratively on an online team task. Specifically, team members with different expertise worked on a time-limited project. To perform at a high level, team members had to combine their unique knowledge and make the right decisions for their customers. The experimental task mirrors the conditions of real organizational teams to a great extent, as it highlights the interdisciplinarity, the interdependency among team members and the need to collaborate to achieve the highest outcomes on a one-shot project.

In Study 1, we focused on the individual level and investigated the relationships between team composition (having an AI compared to a human team member), the perceived trustworthiness and the perceived similarity of another team member and interpersonal trust towards that team member. In Study 2, we incorporated the team level and investigated the relationships between team composition (having an AI compared to a human team member on the team), interpersonal trust and team trust (i.e. the shared perception about the team as a whole). The experimental design allowed us to gain initial insights into how team trust emerges in human–AI teams (compared to human–human teams) and to control for extraneous effects.

STUDY 1

Methods

From an individual-level perspective, we investigated whether having an AI system compared to a human team member leads to lower levels of perceived trustworthiness (Hypothesis 1a, 1b) and lower levels of perceived similarity (Hypothesis 2a, 2b), and consequently, to lower levels of affective and cognitive interpersonal trust. To this end, we manipulated team composition by using a between-subjects design and by building two groups (Groups A and B). In both groups, the participants worked in two-member teams. In Group A, one participant took on the role of the AI team member and one of a human team member. In line with the Wizard of Oz methodology (WoZ; Riek, 2012), a human participant rather than an intelligent system assumed the role of the AI team member. In Group B, both participants took on the role of a human team member. All participants were informed whether they belonged to a human–AI (Group A) or a human–human team (Group B).

Similar to previous studies investigating trust in AI team members (McNeese et al., 2021), the WoZ technique was applied to overcome the difficulty that, at the time the study was conducted, AI technologies were still not advanced enough to carry on a natural conversation (Rieth & Hagemann, 2022), a factor that significantly impacts human–AI interaction. Further, the WoZ methodology allows for investigating interaction in a controlled setting, where the interaction is not hindered by limitations of the technology (e.g. natural communication behaviours, reliability, technical errors). Previous research endorses the efficacy of this methodology, demonstrating that trust disparities arise even when an entity believed to be AI is, in fact, human (see McNeese et al., 2021).

Participants

To determine the sample size, we ran an *a priori* power analysis (Preacher & Coffman, 2006) with a power level of .80, an alpha-error level of .05, a null hypothesized RMSEA of .00 and an alternative hypothesized RMSEA of .08. The analysis revealed a recommended sample size of 255 two-person teams.

After pre-registering the study (https://osf.io/9swxn), participants ($n = 698$) were recruited from a German university[4] and randomly assigned to one of the two groups. Overall, we removed 204 participants. Specifically, we removed participants who did not complete the study, who failed the communication guidelines tutorial, who failed the control questions and whose matched participant (i.e. team member) did not complete the study, necessitating the exclusion of the complete two-member team. Furthermore, we removed participants and their team members who did not interact sufficiently to complete the team task (see the Exclusion criteria). Finally, we removed eight participants and their respective teams after screening for outliers.

The final sample consisted of 494 individuals (47.8% female, 50.8% male, 1.4% other; $M_{age} = 23.73$, $SD_{age} = 3.89$) who were randomly assigned to 247 two-member teams. Participants worked 17.96 hours per week on average ($SD = 12.20$). The highest degree of education for most of the participants (49.2%) was a bachelor's degree, followed by a master's degree (26.9%). Most participants were from Germany (43.1%), followed by Turkey (10.9%), India (9.5%) and China (6.3%). The most frequently stated areas of study were Management (36.8%), Computation, Information and Technology (12.8%) and Engineering (6.7%).

Team task

As a cover story, participants were told that they were working on an international product development team of a well-known smartphone company. The team's task was to convince the company's CEO to finance a new fitness app for two specific groups: students and seniors (65+ years). During online team meetings, teams had to make decisions about six app functionalities: (1) type of activities (Which activities should the fitness app be able to log and how?); (2) user profile (Which information should the user

[4]Given that almost 84% of the student population in Germany work during their studies (BRF, 2020; https://brf.be/national/1387748/), our student sample overcame the common issue of lacking work experience, which can limit the generalizability of the findings.

be able to enter and how?); (3) data collection (Which data should the fitness app automatically save?); (4) reminders (How often should the app send reminders to the user to perform an activity, and how should these reminders be communicated?); (5) settings (Which app functions should the user be able to adjust?) and (6) proactive functions (When should the app play an active role – for example, make recommendations about other activities or routes?). First, the teams discussed the fitness app for the first target group and then for the second.

Each team member was assigned a different role (designer or software developer). Team members received a two-page role description explaining their expertise and including six unique pieces of information – three for each target group. In total, each team received 12 unique pieces of information. Team members needed to share and use this unique information when making collective decisions about each app's functionality to perform at a high level.

Team members communicated and performed their task via an online communication platform embedded into the survey software Qualtrics (Qualtrics, Provo, UT; for more information, see SMARTRIQS; Molnar, 2019). The communication platform allowed communication solely via text messages; no other media could be sent. Additionally, only the role name (i.e. designer or software developer) of the message sender was displayed.

Procedure

After registering for the study, we invited participants to participate in an online experimental session via the video conference platform ZOOM (reference).[5] In total, we ran 22 experimental sessions with 30–40 participants that lasted about an hour. Every session was moderated by an instructor, who used a script to ensure that within and across teams, the same steps were followed, and the same information was provided. During the session, only the instructor was visible to the participants; the participants could neither see nor interact with each other. After the instructor provided some general information about the study (for complete information, see Appendix S2), the study's link was posted in the chat. Each participant was instructed to click on the link and start with the study. Participants were randomly matched with another participant who was participating in the online session, randomly assigned to one of two groups (Group A or Group B), and randomly assigned to one of two roles (designer or software developer).

All teams completed three parts (t0, t1 and t2). During t0, the participants received information about their company, team and team task. They were then presented with six communication guidelines that they were instructed to follow during the virtual team meetings. This ensured clear and concise communication. The guidelines were as follows: Guideline 1: Get to know each other!; Guideline 2: Make strategy suggestions; Guideline 3: Respond to team members' suggestions; Guideline 4: Keep the conversation going!; Guideline 5: Stick to the facts and avoid emotions; and Guideline 6: Express your expertise. A brief explanation and examples followed each guideline.

In both groups, only the software developer received an additional tutorial to practise the communication guidelines (for details, see the Communication Guidelines Tutorial section below). The guidelines were built based on prior studies on chatbot communication (Wuenderlich & Paluch, 2017) so that the software developer (human or AI team member) would communicate in a way that an AI system would be expected to do. While the software developer was completing the communication guidelines tutorial, the designer was in a 'waiting room' and instructed to recall and make note of the most important information from the introduction and the role received thus far.

After the communication guidelines were explained, participants received their two-page role description. Depending on the group, the description informed participants whether they were part of a human–AI (Group A) or a human–human team (Group B). The role description also included

[5]In Study 1, the ethical approval was obtained only from the Technical University of Munich, in the second study, the ethical approval was obtained from the Technical University of Munich and the University of Amsterdam.

information about the respective expertise and the six unique pieces of information (three for each target group). Participants were instructed to read all the information carefully, as it would not be provided again. Furthermore, participants were not allowed to move on to the next page for 2 min, which required them to spend some time reading their role description. Subsequently, the first online team meeting started, during which each team member was instructed to describe their role and expertise to the other team member using the team's chat. Given the length of the role description, each team member had 2 min for this task (4 min for the team in total).

Then, the second and third online team meetings began. First, the teams discussed the development of a new fitness app for one of the two target groups (t1). For each of the six app functionalities, teams were presented with three possible options and asked to interact via the chat to collectively select one of them. One option was always correct (i.e. one of the six unique pieces of information for the respective target group). Second, the teams discussed the development of a new fitness app for the second target group (t2). The procedure was the same as with the first target group. The order of the target groups was counterbalanced. At the beginning of each part (t1, t2), all participants received reminders to follow the communication guidelines. Teams were allotted 20 min for each part (t1 and t2). Before t0, we assessed two control variables (see the Measures section below). After t0, t1 and t2, we assessed perceived trustworthiness, perceived similarity and interpersonal trust (affective and cognitive) using self-reports. After t2, we also assessed information for our manipulation checks, demographic variables and exploratory variables (see the Measures section below). At the end of the experiment, team performance was objectively assessed based on the collective decisions that each team made for every app functionality in both task rounds. All team-member communication was saved to control for whether participants followed the communication guidelines and whether sufficient team interaction for completing the team task took place (see Exclusion criteria).

At the end of the study, we thanked the participants for their participation and debriefed them about the purpose of the study. Participants were then compensated for their participation with 9 € per person. Participants could earn up to 15 € based on their team performance. Team performance was calculated based on the number of unique pieces of information that the team used, ranging from 0 points (i.e. when none of the unique pieces of information were used) to 6 points (i.e. when all 12 unique pieces of information were used). For details on team performance, see Appendix S1. Figure 2 illustrates the design of the overall study (for the complete verbatim instructions, see Appendix S3).

Communication guidelines tutorial

After the six communication guidelines were presented and explained, the participants who had been assigned the role of the software developer completed a short tutorial. In the first part, they were presented with six different situations. After each situation, they were asked to select one of three responses, only one of which was correct. When an incorrect response was selected, feedback was provided explaining how the response should have been correctly formulated. In the second part, participants received

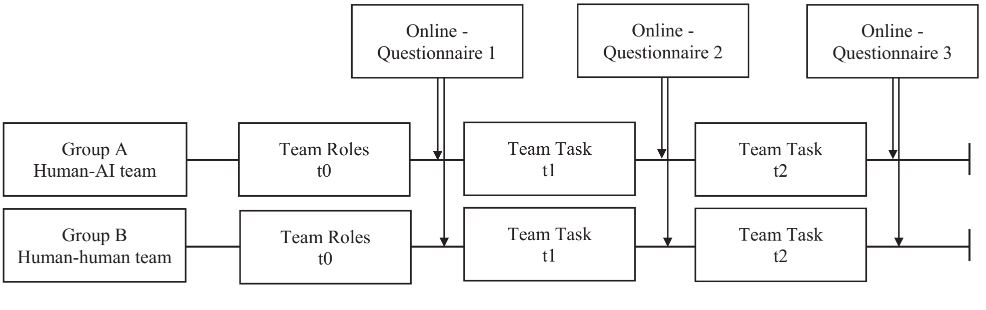

FIGURE 2 Study design.

five additional short situations and were asked to type their responses in line with the communication guidelines. Participants were always presented with the best possible response for all five situations.

Prior to conducting the present research, the communication guidelines were tested in a pilot study (N = 20 individuals). The results showed that in the first part, most of the participants were able to follow the rules (n = 5 participants made 1 mistake, n = 1 participant made 2 mistakes). In the second part, 19 participants made no mistakes or only one, with just one participant making three mistakes. Our findings showed that participants could learn to follow the communication guidelines by completing the tutorial (see pages 4–10 of Appendix S3 for the communication guidelines tutorial).

Measures

All scales included in the online questionnaire, when not mentioned otherwise, were measured using a 5-point Likert scale ranging from 1 (*totally disagree*) to 5 (*totally agree*). The scale items were modified to capture the perceptions of a single team member towards a specific referent: another team member (see Appendix S4 for all measures).

Perceived trustworthiness

We used eight items by Jarvenpaa et al. (1998) to measure perceived trustworthiness (e.g. 'My team member shows a great deal of integrity'), which showed good reliability (α = .94–.95).

Perceived similarity

For perceived similarity, we used three items by Turban and Jones (1988) (e.g. 'My team member and I are similar in terms of our outlook, perspective, and values') and two additional items by Liden et al. (1993) (e.g. 'My team member and I see things in much the same way'). The five items showed good reliability (α = .93–.95).

Interpersonal trust

Interpersonal trust was assessed with 12 items by McAllister (1995), six focusing on cognitive (e.g. 'I can freely share my ideas, feelings, and hopes with my team member') and six on affective interpersonal trust (e.g. 'I can freely talk to my team member about difficulties I am having with the project and know that this team member will want to listen'). Cognitive interpersonal trust showed better reliability (α = .91–.93) without the reverse-coded item (i.e. 'If people knew more about this team member and its background, they would be more concerned and monitor the team member's performance more closely'), which was also not a good indicator of the construct when running confirmatory factor analysis; thus, it was removed. Affective interpersonal trust showed good reliability (α = .89–.93).

Demographics

The following demographic information was measured: age, gender, nationality, highest degree of education, field of study, hours of employment per week and prior experience with teamwork.

Control variables

Given that prior technology use has an impact on human behaviour when using the same or similar technologies (Lippert & Forman, 2006), we assessed the participants' *prior experience with online chat systems* using the item, 'How often do you use online chat systems (e.g. Slack)?', and *prior experience with autonomous chat systems* with the item, 'How often do you use an autonomous chat system (e.g. a bot)?', on a 5-point Likert scale ranging from 0 (*never*) to 5 (*always*).

Given that both groups operated virtually, we also assessed participants' *trust in intelligent technologies* with six items (e.g. 'Generally, I trust intelligent technologies') based on the trust in technology scale developed by Thielsch et al. (2018). To ensure that the term 'intelligent technologies' was properly understood, before assessing the variable, we provided participants with the following explanation: 'In this study, intelligent technologies are technologies that can run autonomously, simulate human intelligence, and adjust their behaviour based on their experience (e.g., machine learning, AI)'. The scale showed good reliability ($a=.79$).

As a manipulation check, we asked participants *how realistic they found the team scenario*, which *role* they had and whether they were a *member of a human–AI team or a human–human team*. Finally, we assessed different team-level exploratory variables (i.e. team trust, team performance and team identification; for more details, see Appendix S4).

Exclusion criteria

We excluded 119 participants who did not complete the study due to facing one or more of the following technical difficulties: (1) their computer was not working or had a faulty internet connection; (2) the study link was not loading; (3) matching between participants did not work or (4) messages were not sent instantly via the team's chat. Furthermore, we excluded nine more participants because they failed at least one of the control questions (i.e. which role they had and whether they were members of a human–AI team or a human–human team). Twenty-four more participants also had to be removed because their team member had been excluded. Nine participants responded to less than 64% (7 or fewer out of 11) of the situations in the communication guidelines tutorial correctly, but this did not lead to additional exclusions, as these participants had already been excluded due to one or more of the previous exclusion criteria described. Finally, 36 participants (18 dyads) were excluded for not engaging in a minimum of chat interactions (one chat message each) for more than 58.3% (7 out of 12) of the app functionalities. Across conditions, participants found the team scenario sufficiently realistic ($M=3.22$, $SD=1.12$); therefore, no additional participants were excluded.

We screened for outliers using Mahalanobis distance scores, which were generated from multiple regressions (Mahalanobis, 1930). Analysis revealed eight cases with a distance score exceeding the critical chi-square value of 20.52 (at $a=.001$). After further examination of the eight cases, we excluded them and their respective teams and continued the analyses with a sample of 494 participants.

Data analysis

Data analysis was conducted in R Studio 4.3.1 (for hypothesis testing) and with SPSS (IBM SPSS Statistics, Version 29) (for preliminary and exploratory analyses). As our focus was on the individual level, and given that we were interested in the individual perceptions of another team member (AI or human) and its impact on interpersonal trust towards that team member, data analysis was performed based on the data collected from the participants assigned the role of the designer ($n=247$). At the same time, as we wanted to explore causal relationships, we tested our hypotheses focusing on team composition as manipulated at t0, perceived trustworthiness and perceived similarity as assessed at t1 and interpersonal trust (affective and cognitive) as assessed at t2.

To examine our theoretical model, we employed structural equation modelling (SEM) using the *lavaan* package in R (Version 0.6.15; Rosseel et al., 2012). This analytic approach allowed us to assess the hypothesized relationships between team composition, interpersonal trust and the mediating roles of perceived trustworthiness and perceived similarity. Specifically, team composition was modelled as the independent variable, with similarity and trustworthiness serving as mediators and interpersonal trust as the dependent variable. No additional control variables were included. To evaluate the mediation effects, we adopted a bootstrapping approach with 5000 resamples to estimate the indirect effects of team composition on interpersonal trust through similarity and trustworthiness while controlling for direct paths from team composition to the outcome.

Results

Preliminary analysis

Descriptive and correlational analyses were calculated for the main study variables and the exploratory variables (see Table 1). We further tested for differences in the control variables (i.e. prior experience with online chat systems, prior experience with autonomous chat systems and trust in intelligent technologies) between the two groups (human–AI teams and human–human teams). The results from two-sample *t* tests showed no significant differences in prior experience with online chat systems ($t[492] = 1.35$, $p = .175$) and trust in intelligent technologies ($t[492] = .91$, $p = .360$) between human–AI teams ($M = 3.72$, $SD = 1.17$; $M = 3.50$, $SD = .66$) and human–human teams ($M = 3.86$, $SD = 1.09$; $M = 3.55$, $SD = .61$). Significant differences were only found with regard to prior experience with autonomous chat systems ($t[492] = -2.90$, $p < .05$), with experience being higher in human–AI teams ($M = 2.36$, $SD = .89$) than in human–human teams ($M = 2.14$, $SD = .79$). As none of the control variables were significantly correlated with perceived trustworthiness, perceived similarity or interpersonal trust, these were not included in the hypothesis testing. We also checked for linearity and homoscedasticity using residual plots, and for normality of the residuals by generating histograms and Q–Q plots. Table 2 presents the fit indices for the confirmatory factor analyses across various models. The hypothesized hierarchical model, featuring higher order interpersonal trust, demonstrated a good fit to the data, surpassing the fit of the alternative one-factor and three-factor models. Following the recommendations of Hu and Bentler (1999), we specified an acceptable model with an RMSEA less than or equal to .05, an SRMR less than or equal to .08 and CFI values higher than or equal to .95 (Byrne, 2001; Hu & Bentler, 1995; Weston et al., 2008). Furthermore, the standardized factor loadings for the items on their respective latent factors varied between .63 and .99, all of which were statistically significant. These results support the construct validity of the measures used in Study 1.

Hypothesis testing

In Table 3, we show the coefficient estimates for the hypothesized model. In Table 4, we display the estimates for the hypothesized indirect effects. Hypotheses 1 and 2 addressed the indirect effect of team composition on interpersonal trust via perceived trustworthiness and perceived similarity respectively. The results indicate that team composition was not related to interpersonal trust ($\gamma = -.10$, $SE = .08$, $p = .21$). Team composition was negatively related to perceived trustworthiness ($\gamma = -.19$, $SE = .09$, $p = .03$) and not significantly related to perceived similarity ($\gamma = -.09$, $SE = .11$, $p = .42$). Perceived trustworthiness was positively related to interpersonal trust ($\gamma = .69$, $SE = .11$, $p < .001$). Perceived similarity was positively related to interpersonal trust ($\gamma = .22$, $SE = .08$, $p = .006$). With regard to the indirect effects, we found that team composition had a significant negative effect via perceived trustworthiness (*indirect effect* $= -.13$, 95% CI [−.26, −.04]), supporting Hypothesis 1. The indirect relation of team composition with interpersonal trust via perceived similarity was, however, not significant (*indirect effect* $= -.02$, 95% CI [−.07, −.03]). Hypothesis 2 was therefore rejected.

The total effect of team composition on interpersonal trust was significant (total effect $= -.25$, $SE = .11$, $p = .026$, 95% CI [−.47, −.02]), indicating a direct relationship that is partially mediated by trustworthiness.

Exploratory analysis

We further explored whether team composition had a direct impact on team trust, that is, the shared perception about the team as a whole. Here, we were interested in determining whether perceived team trust differed between the two groups. The results from a two-sample *t* test showed no significant differences

TABLE 1 Means, standard deviations and correlations with confidence intervals of the study 1 variables.

Variable	*M*	*SD*	1	2	3	4	5	6	7	8
1. Team composition	.47	.50								
2. Perceived trustworthiness t1	3.26	.76	.01							
			[−.12, .13]							
3. Perceived similarity t1	3.26	.81	−.05	.70**						
			[−.18, .07]	[.63, .76]						
4. Interpersonal trust t2	3.73	.84	.17**	.32**	.27**					
			[−.28, −.04]	[.21, .43]	[.15, .39]					
5. Affective interpersonal trust t2	3.45	1.01	.28**	.28**	.27**	.94**				
			[−.27, −.02]	[.16, .39]	[.15, .38]	[.93, .96]				
6. Cognitive interpersonal trust t2	4.05	.81	−.16*	.32**	.22**	.87**	.65**			
			[−.28, −.03]	[.20, .43]	[.10, .33]	[.83, .89]	[.57, .71]			
7. Team trust t2	3.89	.75	−.16*	.31**	.29**	.82**	.70**	.83**		
			[−.28, −.04]	[.19, .42]	[.17, .40]	[.78, .86]	[.63, .76]	[.78, .86]		
8. Team performance	13.87	1.20	−.04	.16*	−.00	.10	.04	.18**	.10	
			[−.17, .09]	[.03, .28]	[−.13, .12]	[−.02, .22]	[−.09, .16]	[.51, .29]	[−.02, .22]	
9. Team identification	4.95	1.66	−.09	.18**	.22**	.69**	.66**	.58**	.68**	.08
			[−.22, .03]	[.05, .30]	[.10, .34]	[.61, .75]	[.58, .72]	[.49, .66]	[.60, .74]	[−.05, .20]

$^{**}p<.01$ (two-tailed). $^{*}p<.05$; $N=247$.

TABLE 2 Confirmatory factor analysis fit indices for measurement model.

Model	χ^2	*df*	$\Delta\chi^2$ (Δdf)	*p* Value $\Delta\chi^2$ (Δdf)	CFI	RMSEA	SRMR
Hierarchical model	1059.20	459	–	–	.911	.073	.055
Four-factor model	1027.72	458	31.48	<.001	.916	.071	.050
Three-factor model[a]	1492.97	461	433.77	<.001	.847	.095	.064
One-factor model	2665.96	464	1606.80	<.001	.675	.139	.094

Note: $N = 247$. Difference in chi-square values ($\Delta\chi^2$) was estimated for comparison with the hierarchical model.

Abbreviations: CFI, confirmatory fit index; RMSEA, root mean square error of approximation; SRMR, standardized root mean square residual.

[a]Affective and cognitive trust as one factor.

TABLE 3 Results of structural equation modelling (direct effects).

	Perceived trustworthiness			Perceived similarity			Interpersonal trust		
	Coeff	*SE*	*p*	Coeff	*SE*	*p*	Coeff	*SE*	*p*
Team composition (0 = *human team*)	**−.19***	.09	.03	−.09	.11	.42	−.10	.08	.21
R^2 (standardized)	.019			.003			.641		
	Interpersonal trust								
	Coeff	***SE***	***p***						
Perceived trustworthiness	**.69*******	.11	<.001						
Perceived similarity	**.22*****	.08	.006						

Note: $N = 247$. Significant coefficients are highlighted in bold.

Abbreviations: Coeff, unstandardized coefficient; *SE*, standard error of unstandardized coefficient.

*$p < .05$. **$p < .01$. ***$p < .001$.

TABLE 4 Indirect effects of team composition on interpersonal trust.

	Indirect effects		
	Coeff	CI LL	CI UL
H1: Team composition → Perceived trustworthiness → Interpersonal trust	**−.13**	−.261	−.004
H2: Team composition → Perceived similarity → Interpersonal trust	−.02	−.070	.030

Note: $N = 247$. Significant coefficients are highlighted in bold.

Abbreviations: Coeff, unstandardized coefficient; CI LL, lower level of bias-corrected 95% confidence interval; CI UL, upper level of bias-corrected 95% confidence interval.

in team trust between human–AI and human–human teams. In addition, we explored whether team identification and team performance differed between the two groups, given that trust reflects an essential mechanism for successful team outcomes (Webber, 2002). The results again showed no significant differences in team performance between human–AI and human–human teams. However, we found significant differences in team identification, which was lower in human–AI teams than in human–human teams. Given that interpersonal trust consists of affective and cognitive interpersonal trust, which can be impacted differently by various factors (Glikson & Woolley, 2020), and insights from the analysis of the measurement model, we also explored differences in affective and cognitive trust between the two groups. The results showed no significant differences in affective interpersonal trust between human–AI and human–human teams. However, we found significant differences in cognitive interpersonal trust, which was lower in human–AI teams than in human–human teams (see Table 5 for the direct impact of team composition).

TABLE 5 Differences between human–AI and human–human teams as perceived by the role of the designer.

	Human–AI teams		Human–human teams				
	M	*SD*	*M*	*SD*	*df*	*t*	*p*
Team trust	3.80	.81	4.00	.85	245	1.87	.061
Team performance	13.82	1.24	13.91	1.53	245	.58	.558
Team identification	4.66	1.86	5.23	1.61	245	2.55	<.05
Affective interpersonal trust	3.4	.98	3.51	1.05	245	.82	.378
Cognitive interpersonal trust	3.95	.87	4.25	.78	245	2.9	<.005

TABLE 6 Differences between human–AI and human–human teams as perceived by the role of the software developer.

	Human–AI teams		Human–human teams				
	M	*SD*	*M*	*SD*	*df*	*t*	*p*
Perceived trustworthiness	3.90	.79	4.10	.61	245	2.30	<.05
Perceived similarity	3.75	.91	4.04	.81	245	2.68	<.01
Interpersonal trust	3.57	.92	3.85	.73	245	2.61	<.01
Affective interpersonal trust	3.29	1.09	3.59	.92	245	2.31	<.05
Cognitive interpersonal trust	3.92	.93	4.17	.66	245	2.46	<.05

Finally, we investigated differences in perceived trustworthiness, perceived similarity and interpersonal trust between conditions, as indicated by the second team member (software developer). The results showed significant differences in perceived trustworthiness, perceived similarity and interpersonal trust, with both affective and cognitive interpersonal trust being lower in human–AI teams than in human–human teams (Table 6).

STUDY 2

Methods

Moving on to the team level, the goal of Study 2 was to investigate whether having an AI system rather than a human team member leads to lower levels of interpersonal trust (affective and cognitive) towards that team member and, consequently, to lower levels of team trust (Hypothesis 3a, 3b). Similar to Study 1, we manipulated team composition by using a between-subjects design and building two groups (Groups A and B). In both groups, participants worked in three-member teams. Using the WoZ methodology (Riek, 2012), in Group A, one participant took on the role of the AI team member and the other two of the human team members. In Group B, all three participants assumed the role of human team members. Participants were informed whether they belonged to a human–AI team (Group A) or human–human team (Group B).

Participants

To determine the sample size, we ran an *a priori* power analysis using the Shiny app for Monte Carlo power analysis for indirect effects by Schoemann et al. (2014, 2017). We selected the two parallel mediator model with a power level of .80, a confidence level of 95, 1000 replications and 20,000 Monte Carlo draws per replication. As input values, we referred to prior empirical studies that had investigated interpersonal trust and team trust (McAllister, 1995). The analysis revealed a recommended sample size of 106 three-member teams.

After pre-registering the study (https://doi.org/10.17605/OSF.IO/ZVDEP), we recruited participants from the same German university ($n = 356$) as in Study 1, who have not participated in Study 1, as well as participants from a Dutch university[6] ($n = 166$), who were randomly assigned to one of the two groups and compensated for their participation. We removed 204 participants following the same steps and considering the same exclusion criteria as in Study 1.

The final sample consisted of 318 individuals (60.1% female, 39% male, .9% other; $M_{age} = 23.69$, $SD_{age} = 4.29$; $n_{Germany} = 201$, $n_{Netherlands} = 117$). The participants were randomly assigned to 106 three-member teams. The participants worked 17.09 hours per week on average ($SD = 14.95$). The highest degree of education for most of the participants (43.7%) was a bachelor's degree, followed by high school graduation (34.9%). Most participants were from Germany (29.2%), followed by China (10.1%), India (9.1%) and Turkey (8.5%). The most frequently stated areas of study were Management (36.8%), Computation, Information and Technology (14.5%) and Social Sciences (11.3%).

Team task

Both groups performed the same team task as in Study 1. Participants were assigned to one of three different roles (marketing expert, designer or software developer). Team members received four unique pieces of information – two for each target group – so that each team had 12 unique pieces of information in total.

Procedure

We ran 40 online experimental sessions[7] with 30–40 participants, following the same procedure as in Study 1. Specifically, all three-member teams completed the three parts of the study (t0, t1 and t2). During t0, the participants received information about the company, the team, the team task and the communication guidelines. Then, only participants who had been assigned the role of the software developer completed the additional tutorial and practised the communication guidelines (for details, see the description under Study 1). After every member received and read their two-page role description, the first online team meeting began (t0). During the team meeting, each participant had 2 min (6 min for the team in total) to introduce themself and their expertise using the team's chat.

Afterwards, the second and third online team meetings started. Teams first discussed the development of a new fitness app for one target group (seniors or students; t1) and then for the other (seniors or students; t2). Before each online team meeting, the participants received a reminder to follow the communication guidelines during the meeting. For each part (t1 and t2), teams were allotted 20 min.

Before t0, we assessed two control variables (see the Measures section below). After t0, t1 and t2, we assessed interpersonal trust (affective and cognitive) and team trust via self-reports. After t2, we also assessed demographics, team identification (see the Measures section below) and information for our manipulation checks. All written communication was saved and used to check whether participants followed the communication guidelines and whether teams interacted sufficiently (see the Exclusion criteria).

[6]Due to the number of individuals recruited for Study 1, the number of remaining individuals willing to participate in Study 2 was not sufficient to achieve the sample size needed. Therefore, using the same inclusion criteria and recruitment process, we also recruited individuals from a university in the Netherlands. In the final sample, there were no significant differences between the two universities in terms of age ($t[316] = 1.44$, $p = .150$), working hours ($t[218] = -0.73$, $p = .466$), experience with autonomous chat systems ($t[316] = -0.31$, $p = .754$) and how realistic they found the scenario ($t[316] = 0.30$, $p = .757$). There was only a significant difference for gender ($t[316] = -5.14$, $p < .001$), with the German sample having significantly more male and fewer female participants than the Dutch sample.

[7]Ethical approval was obtained by the Technical University of Munich and the University of Amsterdam.

At the end of the study, we thanked the participants for their participation, debriefed them and compensated them with 12 € per person. Participants could earn up to 18 € based on their team performance, which was objectively assessed, similar to in Study 1 (for details on team performance, see Appendix S1). In total, the experiment lasted about 1 hour and a half (for the complete verbatim instructions, see Appendix S3).

Measures

All scales included in the online questionnaire, when not mentioned otherwise, were measured using a 5-point Likert scale ranging from 1 (*totally disagree*) to 5 (*totally agree*). For all of the study's measures, see Appendix S4.

Interpersonal trust

The 12-item scale by McAllister (1995) was used to capture the perceptions of each individual team member towards two specific referents: the other two team members. The scale items were modified so that the phrase 'my team member' was consistently used. Six items focused on cognitive interpersonal trust [e.g. 'I can freely share my ideas, feelings, and hopes with my (marketing/designer/software developer) team member'] and three on affective interpersonal trust [e.g. 'I can freely talk to my (marketing/designer/software developer) team member about difficulties I am having with the project and know that this team member will want to listen']. For each participant, we calculated the average of the items for cognitive and affective interpersonal trust for each of the other two team members. To be consistent with Study 1, and due to better reliability values, in Study 2, we similarly removed the reversed item of the cognitive interpersonal trust scale, resulting in good reliability ($\alpha = .92–.94$). Affective interpersonal trust also showed good reliability ($\alpha = .87–.92$).

Team trust

To capture team trust (i.e. the shared perception about the team as a whole), we used the 8-item scale by Jarvenpaa and Leidner (1999). According to recent meta-analytic findings, it is one of the best-performing scales for capturing team trust (Feitosa et al., 2020) and has shown good psychometric properties (Kirkman et al., 2006). The items were modified so that the phrase 'my team' was consistently used (e.g. 'My team is usually considerate of one another's feelings'). The scales showed good reliability ($\alpha = .81–.88$).

Demographics

The following demographic information was measured: age, gender, nationality, highest degree of education, field of study, hours of employment per week and prior experience with teamwork.

Control variables

Similar to Study 1, as control variables, we assessed the participants' prior experience with technology (online chat systems and autonomous chat systems) and participants' trust in intelligent technologies (Thielsch et al., 2018), which showed good reliability ($\alpha = .74$). Furthermore, we asked participants which role they had and whether they were a member of a human–AI or a human–human team to run our manipulation checks. Furthermore, given that the extent to which a team member identifies with their team impacts the team member's intentions and actions (Ashforth & Mael, 1989), we assessed *team identification* with a single item by Shamir and Kark (2004) (i.e. 'How much do you identify with your team?') on a 7-point Likert scale ranging from 0 (*not at all*) to 7 (*totally*) as an exploratory variable at the team level. Finally, we assessed perceived trustworthiness ($\alpha = .93–.94$), perceived similarity ($\alpha = .94–.95$) and team performance as additional exploratory variables.

Exclusion criteria

We used the same exclusion criteria described in Study 1. In total, we excluded 195 participants (146 from the German sample, 49 from the Dutch sample). We excluded 76 participants due to facing one or more of the following technical difficulties: (1) the computer was not working or had a faulty Internet connection; (2) the study link was not loading; (3) matching between participants did not work and (4) messages were not sent instantly via the team's chat. Furthermore, we excluded 21 participants because they failed at least one of the control questions. Fifty-nine more participants were removed because one or more of their team members were excluded. Six additional participants (two teams) were removed because the AI software developer of each team mentioned their real name when introducing themselves, deviating from the instructions. Finally, 33 participants (11 teams) were excluded for not engaging in a minimum of chat interaction (one chat message each) for more than 58.3% (7 out of 12) of the app functionalities. Across conditions, participants found the team scenario to be sufficiently realistic on average ($M = 3.50$, $SD = .93$); therefore, no additional participants were excluded.

Furthermore, we screened for outliers using Mahalanobis distance scores, similar to Study 1. Analysis revealed three cases with a distance score exceeding the critical chi-square value of 20.52 (at a =. 001). After further examination of the three cases, we excluded them and the respective teams and continued the analyses with a sample of 106 three-member teams.

Data analysis

To investigate the impact of interpersonal trust towards another team member (AI or human) on team trust, data analysis was performed at the team level based on the data collected from the participants assigned the role of the marketing and the design expert. Given the causal nature of our hypothesized relationship, we tested our hypothesis focusing on team composition as manipulated at t0, interpersonal trust at t1 and team trust at t2. Data analysis was conducted using SPSS (IBM SPSS Statistics, Version 29). For mediation analyses, we used the PROCESS macro (Model 4; Hayes, 2017) for SPSS. Some exploratory analyses were performed with R Studio 4.3.1.

Results

Preliminary analysis

To aggregate the individual responses using the mean of the two participants assigned the roles of the marketing expert and the designer, which is a common method reported in the literature (Mathieu et al., 2008), we calculated the within-group agreement indices (r_{wg}; James et al., 1993) and intra-class correlation coefficients (ICC [1] and ICC [2]; Bliese, 2000).

Interpersonal trust at t0 and t1 ($r_{WGJ} = .81–.82$, ICC [1] = .17–.29, $p < .05$, ICC [2] = .28–.45), team trust at t2 ($r_{WGJ} = .88$, ICC [1] = .11, $p = .139$, ICC [2] = .19) as well as the variables perceived trustworthiness at t0 ($r_{WGJ} = .80$, ICC [1] = .24, $p < .05$, ICC [2] = .39) and perceived similarity at t0 ($r_{WGJ} = .74$, ICC [1] = −.01, $p = .531$, ICC [2] = −.02) used for the exploratory analysis exceeded the recommended cutoff value of .70 for mean r_{wg} and showed a sufficient basis to support their aggregation (for ICC, see LeBreton & Senter, 2008; for $r_{WGJ,}$ see Cohen et al., 2001). Thus, we continued with the aggregation of these variables and the exploratory variable *team identification*. We also checked for linearity and homoscedasticity using residual plots and for the normality of the residuals by generating histograms and Q–Q plots.

Descriptive and correlational analyses were calculated for the main study variables (see Table 7). The results showed that propensity to trust technology was positively correlated with interpersonal trust ($r = .21$, $p < .05$). As a result, it was included as a control variable in hypothesis testing.

TABLE 7 Means, standard deviations and correlations with confidence intervals of the study 2 variables.

Variables	*M*	*SD*	1	2	3	4	5	6	7
1. Team composition	.45	.50							
2. Interpersonal trust t1	3.65	.60	−.11						
			[−.29, .09]						
3. Affective interpersonal trust t1	3.34	.72	−.16	.92**					
			[−.34, .03]	[.88, .94]					
4. Team trust t2	4.25	.45	−.04	.42**	.36**				
			[−.23, .15]	[.25, .56]	[.18, .52]				
5. Team identification	5.29	1.14	.01	.52**	.52**	.65**			
			[−.18, .20]	[.37, .65]	[.37, .65]	[.53, .75]			
6. Team performance	16.92	.86	−.01	−.03	−.09	.06	−.02		
			[−.20, .18]	[−.22, .16]	[−.27, .11]	[−.14, .25]	[−.21, .17]		
7. Perceived trustworthiness t0	3.23	.66	−.17	.31**	.25**	.17	.19*	−.06	
			[−.35, .03]	[.12, .47]	[.07, .42]	[−.03, .35]	[.00, .37]	[−.25, .13]	
8. Perceived similarity t0	3.10	.62	−.01	.33**	.30**	.15	.26**	−.12	.82**
			[−.29, .09]	[.15, .49]	[.12, .47]	[−.05, .33]	[.08, .43]	[−.31, .07]	[.75, .88]

** $p < .01$ level (two-tailed). * $p < .05$; $N = 106$.

Hypothesis testing

To examine the indirect effect of interpersonal trust (affective and cognitive) on the relationship between team composition and team trust (Hypothesis 3), we ran a mediation analysis, controlling for propensity to trust technology. Our results showed that the indirect effect was not significant ($b = -.04$, 95% CI [−.12, .01], containing zero). Interpersonal trust had a direct positive and significant impact only on team trust ($b = .30$, $p < .001$). Hence, Hypothesis 3 was not supported.

Exploratory analysis

To better understand how team composition and interpersonal trust towards each team member within the team impacted individual ratings of team trust, we further explored whether team composition impacted interpersonal trust at t1, and in succession, team trust at t2, as perceived by each team member, while controlling for propensity to trust technology. The results showed that the indirect effect was not significant for interpersonal trust towards the software developer ($b = -.02$, 95% CI [−.07, .03] for the designer; $b = -.07$, 95% CI [−.16, .01] for the marketing expert), the marketing expert ($b = -.00$, 95% CI [−.07, .05] for the designer; $b = -.06$, 95% CI [−.20, .04] for the software developer) or the designer ($b = .04$, 95% CI [−.06, .16] for the marketing expert; $b = .03$, 95% CI [−.08, .15] for the software developer). Interpersonal trust had a direct positive and significant effect only on team trust ($b = .28$–.44, $p < .001$).

Given that team trust was not impacted by team composition, we explored whether team trust was impacted by perceived trustworthiness and by perceived similarity via interpersonal trust. After controlling for propensity to trust technology, our results showed that the indirect effect of interpersonal trust on the relationship between perceived trustworthiness and team trust was significant ($b = .07$, 95% CI [.02, .14], not containing zero). Our results also showed that the indirect effect of interpersonal trust on the relationship between perceived similarity and team trust was significant ($b = .08$, 95% CI [.02, .16], not containing zero).

Furthermore, we explored whether team identification and team performance differed between the two groups. The results showed no significant differences in team identification or team performance between human–AI and human–human teams. We also explored differences in affective and cognitive trust towards the software developer between the two groups. The results showed no significant differences. Additional exploratory analyses also showed no significant differences in perceived trustworthiness and perceived similarity of the other two team members, interpersonal trust towards each of the other two team members or team trust as perceived by the software developer between human–AI and human–human teams (for an overview, see Table 8).

OVERALL DISCUSSION

The rapid transformation of organizations and their teams by AI is attracting interest, not only offering positive changes but also posing risks if AI systems, serving as tools and teammates, are not trustworthy (Georganta & Ulfert, 2024; Gesk & Leyer, 2022; Ulfert et al., 2023). This is also evident in the recent agreement of the European Commission in December 2023 on the legal framework for trustworthy AI, the 'AI Act' (Helberger & Diakopoulos, 2023), when introducing and collaborating with AI. To contribute to this important issue, especially in the emerging context of human–AI teams, the aim of our experimental studies was to provide empirical insight into how team trust emerges in human–AI teams compared to human–human teams. In contrast to our expectations, being in a human–AI team rather than a human–human team neither directly nor indirectly led to lower levels of interpersonal trust (via perceived similarity) or to lower levels of team trust (via interpersonal trust). However, in two-member teams, interpersonal trust via perceived trustworthiness was lower in human–AI teams than in human–human teams. Furthermore, exploratory findings showed that team identification and cognitive interpersonal trust were lower in two-member human–AI teams than in three-member human–human teams. When comparing three-member human–AI teams with other three-member human–human

TABLE 8 Results from exploratory analyses of differences between human–AI and human–human teams.

	Human–AI teams		Human–human teams				
	M	*SD*	*M*	*SD*	*df*	*t*	*p*
Team performance[a]	16.91	.85	16.93	.85	104	.14	.888
Team identification[a]	5.30	1.07	5.28	1.20	104	−.07	.938
Affective interpersonal trust towards software developer[a]	3.21	.69	3.44	.72	104	1.67	.097
Cognitive interpersonal trust towards software developer[a]	3.93	.49	3.95	.69	104	.16	.869
Perceived trustworthiness of designer[b]	3.37	.70	3.45	.71	104	1.54	.119
Perceived trustworthiness of marketing expert[b]	3.22	.77	3.45	.72	104	.57	.564
Perceived similarity with the designer[b]	3.29	.63	3.29	.74	104	−.03	.971
Perceived similarity with marketing expert[b]	3.32	.70	3.33	.72	104	.07	.941
Interpersonal trust towards designer[b]	3.69	.74	3.58	.93	104	−.62	.530
Interpersonal trust towards marketing expert[b]	3.69	.75	3.87	.78	104	1.13	.129
Team trust[b]	4.42	.65	4.17	.61	104	−.31	.756

[a]As perceived by the designer and marketing expert.

[b]As perceived by the software developer.

teams, we found no differences. Instead, in three-member teams, we showed that perceived trustworthiness and perceived similarity resulted in higher team trust via interpersonal trust in human–AI and in human–human teams.

We expected interpersonal trust and, subsequently, team trust to be lower in human–AI teams than in human–human teams due to lower levels of perceived trustworthiness and perceived similarity of AI systems compared to human teammates. However, our results indicated that team composition did not impact team trust and only had a small effect on interpersonal trust. It is possible that the nature of the AI teammate as an intelligent technology was not sufficiently salient during the team interaction and that therefore, more focus was placed on its role within the team. This could be due to the absence of more influential characteristics unique to AI, like embodiment or voice, which affect perceptions and interactions with intelligent technologies (Deng et al., 2019). It is also possible that the unique skills and expertise of AI teammates for successfully completing the team task were acknowledged and appreciated, possibly without humans explicitly wondering about how the AI teammate would perform its role or how it would process information. In consequence, this may have resulted in similar levels of trust towards AI and human teammates. Related findings have shown that AI decisions and AI agents are sometimes trusted more than human counterparts when they possess unique expertise (Choung et al., 2023; Heid, 2018), while there is a tendency to favour humans when tasks require human skills (Lee, 2018).

However, findings showed that interpersonal trust via perceived trustworthiness and team identification were lower in human–AI teams than in human–human teams when consisting of two team members. This is in line with previous work arguing that human–AI and human–human teams might differ in the way they operate (Georganta & Ulfert, 2024; McNeese et al., 2021). Potentially, when a team includes one AI and one human instead of two humans, the perceived surface-level differences (human vs. AI) may hinder a human's ability to view an AI teammate as trustworthy. This can make it challenging for a human to psychologically align with the AI as part of the same group, as this group identity might clash with their human identity. Having low expectations about a team member's abilities and motives and consequently exhibiting a low level of interpersonal trust is often related to the identification with the team itself (Hakonen & Lipponen, 2009). Low team identification and the distinction between the team and oneself decreases the involvement with the other (AI) team member and the motivation to improve the team, making it difficult to perceive the team's interests as one's own (Ellemers et al., 2002; Van Knippenberg, 2000). It seems that not only the AI team member was categorized as part of the out-group (see social categorization theory; Turner, 2010), but that the human team members also viewed themselves as outsiders struggling to include the human–AI team in their self-concept (Tropp & Wright, 2001). Given that such differences were not found when teams consisted of one AI teammate and two humans, future research should investigate team identification with different human–AI team compositions (e.g. human majority, AI majority and larger team size) and investigate how these might shape the self-perceptions of human–AI group belongingness.

Interestingly, additional analysis also revealed that cognitive interpersonal trust was lower in two-member human–AI teams than in two-member human–human teams. A possible explanation is that humans have different (or even higher) expectations of AI teammates than of human teammates, resulting in a (negative) disconfirmation of these expectations during interaction, resulting in lower perceptions of cognitive interpersonal trust (Kocielnik et al., 2019). Although AI and human teammates might have the same skills and abilities with regard to a specific task, humans often use different standards to judge their AI versus human counterparts (Glikson & Woolley, 2020). It is also possible that humans expect their AI teammates to be more responsible for tasks that incorporate actions in which machines have significant advantages over humans, such as complex calculations, than tasks related to social features (Lee, 2018), resulting in difficulties in building cognitive interpersonal trust for a team task that requires both. Differences in expectations can harm how AI teammates are perceived and can even result in perceived failure of human–AI teams (McNeese et al., 2023); however, as our findings indicate, this seems to change when the number of humans within the human–AI team increases, a premise that requires further exploration.

Finally, our findings demonstrate that high perceived trustworthiness and perceived similarity resulted in higher interpersonal trust and, in turn, in higher team trust for both human–AI and human–human teams consisting of three team members. This observation may suggest that the dynamics and relationships within human–AI teams might be more closely aligned with those in human–human teams than previously believed. It also implies that key factors essential for team functioning and trust are equally critical across both team types.

In line with prior work on human–human teams and AI interaction (Costa et al., 2018; Glikson & Woolley, 2020; Madhavan & Wiegmann, 2007), our results highlight the importance of perceiving not only human but also AI teammates as similar and trustworthy to oneself in order to be able to trust them and, thus, trust the team as a whole. This contributes to gaining a better understanding of interpersonal trust and team trust in human–AI teams, one of the most commonly discussed dependent variables (O'Neill et al., 2022). At the same time, our evidence supports the underlying idea that theories and research on human–human teams can serve as a starting point to obtain a better understanding of human–AI teams (O'Neill et al., 2023). This observation indicates that certain dynamics observed within human–human teams appear to persist in human–AI collaborations. However, our studies also prove that a deeper understanding of these relationships is needed. In particular, exploration of the circumstances under which (e.g. team composition, task type) existing models and theories of human–human team trust might need to be extended has recently been called for (Ulfert et al., 2023).

Limitations and suggestions for future research

Our study comes with limitations that are important to highlight. Although communication guidelines were developed in line with chatbot literature, and the communication tutorial was successfully completed by the participants who acted as the AI teammate, they were nonetheless not actual AI teammates, a factor that may have impacted our findings. Furthermore, our manipulation might not have been strong enough, as we only informed participants that their teammate was an AI system or a human. No additional characteristics were assigned to the AI teammate, such as a form of embodiment during the virtual team meeting, which might have demonstrated different results. Thus, we propose that future research use different forms of manipulation (e.g. a visual representation of the AI teammate) and investigate the proposed relationships in the field or in a laboratory setting with real AI teammates. Despite the frequent application of WoZ methodologies in human–AI team studies, it has its constraints. This is because humans, even when given specific communication guidelines, may still differ in their reactions and communication from AI chatbots (e.g. when disagreement occurs). Recent advancements in AI, especially generative AI over the past year, have significantly enhanced the communicative capabilities of chatbots. This progress may enable future research to overcome the aforementioned limitation and allow for studying team interaction with real AI chatbots as team members.

Another important limitation is that although the definition of an AI teammate was clearly described, most people were not familiar and had no prior experiences with AI teammates. Therefore, it is possible that some participants might have imagined or expected something very different from others when informed about working with an AI teammate. Future research could overcome this issue by asking what participants imagined or what their thoughts were when interacting with the AI teammate and then explore how this 'image' they created might have shaped how the AI teammate and the human–AI team were perceived.

Furthermore, our findings should be generalized with caution to other human–AI team compositions, as we only investigated dyadic and triadic teams, with the majority of the teammates being humans. At the same time, it is essential to acknowledge that the teams studied and the tasks they completed took place as part of a specific team scenario. We encourage additional work that investigates our research questions and hypotheses in different settings and with different team compositions (e.g. more human vs. AI teammates and vice versa). Finally, as we only focused on how humans perceive human versus AI teammates, we propose that future studies also consider how AI teammates perceive human versus AI teammates. This would contribute to a more comprehensive understanding of dyadic and team trust relationships in human–AI teams.

Implications for research and practice

With the present work, we present empirical insights on the topic of human–AI teams from an organizational psychology perspective, as recently called for (McNeese et al., 2021). Specifically, we overcome the mainly theoretical attempts to differentiate between human–human and human–AI teams (O'Neill et al., 2023) and provide evidence about how AI teammates are perceived and trusted, whether these perceptions differ from those of human teammates and how they shape the dyadic trust relationship between a human and an AI teammate as well as trust towards the human–AI team as a whole. In doing so, we also contribute to a better understanding of team trust in human–AI teams, in line with recent multidisciplinary theoretical attempts (Ulfert et al., 2023). Specifically, we acknowledge the role of AI agents in trust processes and team dynamics as well as the existence of different levels and entities within a human–AI team. With regard to the trust literature, this work is one of the first attempts to demonstrate that trust arises at the individual level, then moves to the dyadic level and finally forms shared trust at the team level (Costa et al., 2018). This was possible by clearly differentiating between perceived trustworthiness and trust (Zolin et al., 2004), between levels of analysis (individual and team) and between referents (single team member and team as a whole). In this way, we also show how common conceptualization and measurement issues in the trust literature (Feitosa et al., 2020) can be overcome.

From a practical standpoint, our findings can also serve as a first step in exploring the designing and implementing AI teammates and fostering interpersonal and team trust in human–AI teams. Specifically, we propose that AI teammates should explicitly communicate their abilities, integrity and care for the team when interacting with their teammates, as perceiving them as trustworthy is crucial for building strong trust relationships. We also propose that during human–AI team formation, emphasis should be placed on identifying commonalities rather than differences between human and AI teammates, strengthening the feeling of group belongingness and facilitating psychological merging into the same group as the AI teammate. This is in line with our results showing that increasing perceived similarity between humans and AI teammates can foster trust. Furthermore, to manage expectations realistically and promote both cognitive and affective interpersonal trust, the integration of Explainable AI (Gade et al., 2020) is recommended. Explainable AI, by providing insights into the AI teammate's decision-making process, can enable human colleagues to identify situations in which the AI's reasoning may be right or wrong, thus mitigating the problem of disconfirmation of expectations during team interaction (Buçinca et al., 2021). Finally, introducing AI teammates into teams comprising at least two other humans could serve as an optimal starting point, facilitating smoother integration and collaboration.

CONCLUSION

Our work suggests that certain relationships and dynamics in human–AI teams may be comparable to human–human teams. This is evident given that the results showed that perceived trustworthiness and perceived similarity led to higher team trust via interpersonal trust for both human–AI and human–human teams. Nevertheless, team identification and cognitive interpersonal trust were lower in two-member human–AI teams than in three-member human–human teams and minor differences were found in terms of interpersonal trust. Our study highlights the need for further investigating the trust relationships within human–AI teams and how these might be shaped by other factors, such as team composition. Our findings underscore the necessity for a deeper exploration of how relationships and dynamics emerge in human–AI teams, and when these are similar or distinct from human-human teams. Finally, they offer initial practical insights to enhance the perceived trustworthiness and similarity of AI team members to oneself to facilitate trusting and harmonious human–AI collaborations.

AUTHOR CONTRIBUTIONS

Eleni Georganta: Conceptualization; data curation; formal analysis; funding acquisition; investigation; methodology; project administration; resources; software; validation; visualization; writing – original draft; writing – review and editing. **Anna-Sophie Ulfert:** Conceptualization; data curation; formal analysis; funding acquisition; investigation; methodology; project administration; resources; software; validation; visualization; writing – original draft; writing – review and editing.

ACKNOWLEDGEMENTS

This research is supported by the Society for Industrial and Organizational Psychology (SIOP) Foundation as part of the Visionary Circle.

CONFLICT OF INTEREST STATEMENT

The authors have no conflict of interest to declare.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon reasonable request.

ORCID

Eleni Georganta https://orcid.org/0000-0002-9070-5930
Anna-Sophie Ulfert https://orcid.org/0000-0001-6293-4173

REFERENCES

Ashforth, B. E., & Mael, F. (1989). Social identity theory and the organization. *Academy of Management Review*, *14*(1), 20–39. https://doi.org/10.5465/amr.1989.4278999

Bliese, P. D. (2000). Within-group agreement, non-independence, and reliability: Implications for data aggregation and analysis. In K. J. Klein & S. W. J. Kozlowski (Eds.), *Multilevel theory, research, and methods in organizations: Foundations, extensions, and new directions*. Jossey-Bass/Wiley.

Breuer, C., Hüffmeier, J., & Hertel, G. (2016). Does trust matter more in virtual teams? A meta-analysis of trust and team effectiveness considering virtuality and documentation as moderators. *Journal of Applied Psychology*, *101*(8), 1151–1177. https://doi.org/10.1037/apl0000113

Brewer, M. B. (2017). Intergroup discrimination: Ingroup love or outgroup hate? In C. G. Sibley & F. K. Barlow (Eds.), *The Cambridge handbook of the psychology of prejudice* (pp. 90–110). Cambridge University Press. https://doi.org/10.1017/9781316161579.005

Buçinca, Z., Malaya, M. B., & Gajos, K. Z. (2021). To trust or to think: Cognitive forcing functions can reduce overreliance on AI in AI-assisted decision-making. *Proceedings of the ACM on Human-Computer Interaction*, *5*(CSCW1), 1–21. https://doi.org/10.1145/3449287

Byrne, B. M. (2001). *Multivariate applications book series. Structural equation modeling with AMOS: Basic concepts, applications, and programming*. Lawrence Erlbaum Associates Publishers.

Chavaillaz, A., Schwaninger, A., Michel, S., & Sauer, J. (2019). Expertise, automation and trust in X-ray screening of cabin baggage. *Frontiers in Psychology*, *10*, 1–11. https://doi.org/10.3389/fpsyg.2019.00256

Choung, H., Seberger, J. S., & David, P. (2023). When AI is perceived to be fairer than a human: Understanding perceptions of algorithmic decisions in a job application context. *International Journal of Human-Computer Interaction*, 1–18. https://doi.org/10.1080/10447318.2023.2266244

Cohen, A., Doveh, E., & Eick, U. (2001). Statistical properties of the rWG (J) index of agreement. *Psychological Methods*, *6*(3), 297–310. https://doi.org/10.1037/1082-989X.6.3.297

Colquitt, J. A., Scott, B. A., & LePine, J. A. (2007). Trust, trustworthiness, and trust propensity: A meta-analytic test of their unique relationships with risk taking and job performance. *Journal of Applied Psychology*, *92*(4), 909–927. https://doi.org/10.1037/0021-9010.92.4.909

Costa, A. C., Fulmer, C. A., & Anderson, N. R. (2018). Trust in work teams: An integrative review, multi-level model, and future directions. *Journal of Organizational Behavior*, *39*(2), 169–184. https://doi.org/10.1002/job.2213

Cummings, P., Mullins, R., Moquete, M., & Schurr, N. (2021). Hello World! I am Charlie, an Artificially Intelligent Conference Panelist. *Proceedings of the 54th Hawaii International Conference on* System Sciences (pp. 380–389). https://doi.org/10.24251/hicss.2021.045

Curşeu, P. L., & Schruijer, S. G. L. (2010). Does conflict shatter trust or does trust obliterate conflict? Revisiting the relationships between team diversity, conflict, and trust. *Group Dynamics: Theory, Research, and Practice*, *14*(1), 66–79. https://doi.org/10.1037/a0017104

Davenport, T., & Kalakota, R. (2019). The potential for artificial intelligence in healthcare. *Future Healthcare Journal*, *6*(2), 94–98. https://doi.org/10.7861/futurehosp.6-2-94

De Jong, B. A., Dirks, K. T., & Gillespie, N. (2016). Trust and team performance: A meta-analysis of main effects, moderators, and covariates. *Journal of Applied Psychology*, *101*(8), 1134–1150. https://doi.org/10.1037/apl0000110

De Jong, B. A., & Elfring, T. (2010). How does trust affect the performance of ongoing teams? The mediating role of reflexivity, monitoring, and effort. *Academy of Management Journal*, *53*(3), 535–549. https://doi.org/10.5465/amj.2010.51468649

Deng, E., Mutlu, B., & Mataric, M. (2019). Embodiment in socially interactive robots. *Foundations and Trends® in Robotics*, *7*(4), 251–356. https://doi.org/10.1561/2300000056

Dirks, K. T., & de Jong, B. (2022). Trust within the workplace: A review of two waves of research and a glimpse of the third. *Annual Review of Organizational Psychology and Organizational Behavior*, *9*, 247–276. https://doi.org/10.1146/annurev-orgpsych-012420-083025

Duan, Y., Nie, W., & Coakes, E. (2010). Identifying key factors affecting transnational knowledge transfer. *Information & Management*, *47*(7–8), 356–363. https://doi.org/10.1016/j.im.2010.08.003

Ellemers, N., Spears, R., & Doosje, B. (2002). Self and social identity. *Annual Review of Psychology*, *53*, 161–186. https://doi.org/10.1146/annurev.psych.53.100901.135228

Feitosa, J., Grossman, R., Kramer, W. S., & Salas, E. (2020). Measuring team trust: A critical and meta-analytical review. *Journal of Organizational Behavior*, *41*(5), 479–501. https://doi.org/10.1002/job.2436

Fompeyrine, D. A., Vorm, E. S., Ricka, N., Rose, F., & Pellegrin, G. (2021). Enhancing human-machine teaming for medical prognosis through neural ordinary differential equations (NODEs). *Human-Intelligent Systems Integration*, *1-13*, 263–275. https://doi.org/10.1007/s42454-021-00037-z

Freedy, A., DeVisser, E., Weltman, G., & Coeyman, N. (2007). Measurement of trust in human-robot collaboration. *2007 International Symposium on Collaborative Technologies and Systems* (pp. 106–114). https://doi.org/10.1109/cts.2007.4621745

Fulmer, C. A., & Gelfand, M. J. (2012). At what level (and in whom) we trust. *Journal of Management*, *38*(4), 1167–1230. https://doi.org/10.1177/0149206312439327

Gade, K., Geyik, S., Kenthapadi, K., Mithal, V., & Taly, A. (2020). Explainable AI in industry: Practical challenges and lessons learned. *Companion Proceedings of the Web Conference 2020*. https://doi.org/10.1145/3366424.3383110

Georganta, E., & Ulfert, A. S. (2024). My colleague is an AI! Trust differences between AI and human teammates. *Team Performance Management: An International Journal*. https://www.emerald.com/insight/content/doi/10.1108/TPM-07-2023-0053/full/html

Gesk, T. S., & Leyer, M. (2022). Artificial intelligence in public services: When and why citizens accept its usage. *Government Information Quarterly*, *39*(3), 101704.

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

Grand, J. A., Braun, M. T., Kuljanin, G., Kozlowski, S. W., & Chao, G. T. (2016). The dynamics of team cognition: A process-oriented theory of knowledge emergence in teams. *Journal of Applied Psychology*, *101*(10), 1353–1385. https://doi.org/10.1037/apl0000136

Greenberg, J., & Tomlinson, E. C. (2004). Situated experiments in organizations: Transplanting the lab to the field. *Journal of Management*, *30*(5), 703–724. https://doi.org/10.1016/j.jm.2003.11.001

Hakonen, M., & Lipponen, J. (2009). It takes two to tango: The close interplay between trust and identification in predicting virtual team effectiveness. *The Journal of E-Working*, *3*(1), 17–32.

Hayes, A. F. (2017). *Introduction to mediation, moderation, and conditional process analysis, second edition: A regression-based approach (methodology in the social sciences)* (2nd ed.). The Guilford Press.

Haynes, M. (2020, 13 February). Top 10 Workplace Trends for 2020. SIOP. https://www.siop.org/Research-Publications/Items-of-Interest/ArtMID/19366/ArticleID/3361/Top-10-Workplace-Trends-for-2020?utm_source=SIOP&utm_medium=Website&utm_campaign=Top10page

Heid, A. K. (2018). Trust in Homo Artificialis: Evidence from Professional Investors. In *Annual Meeting*. American Accounting Association.

Helberger, N., & Diakopoulos, N. (2023). The European AI act and how it matters for research into AI in media and journalism. *Digital Journalism*, *11*(9), 1751–1760.

Highhouse, S. (2009). Designing experiments that generalize. *Organizational Research Methods*, *12*(3), 554–566. https://doi.org/10.1177/1094428107300396

Hoff, K. A., & Bashir, M. (2014). Trust in automation. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *57*(3), 407–434. https://doi.org/10.1177/0018720814547570

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, *6*(1), 1–55. https://doi.org/10.1080/10705519909540118

Hu, L.-T., & Bentler, P. M. (1995). Evaluating model fit. In R. H. Hoyle (Ed.), *Structural equation modeling: Concepts, issues, and applications* (pp. 76–99). Sage Publications, Inc.

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

Jarvenpaa, S. L., Knoll, K., & Leidner, D. E. (1998). Is anybody out there? Antecedents of trust in global virtual teams. *Journal of Management Information Systems*, *14*(4), 29–64. https://doi.org/10.1080/07421222.1998.11518185

Jarvenpaa, S. L., & Leidner, D. E. (1999). Communication and trust in global virtual teams. *Organization Science*, *10*(6), 791–815. https://doi.org/10.1287/orsc.10.6.791

Johnson, M., & Vera, A. (2019). No AI is an Island: The case for teaming intelligence. *AI Magazine*, *40*(1), 16–28. https://doi.org/10.1609/aimag.v40i1.2842

Jones, R. A. (1985). *Research methods in the social and behavioral sciences*. Sinauer Associates.

Kaplan, A., & Haenlein, M. (2020). Rulers of the world, unite! The challenges and opportunities of artificial intelligence. *Business Horizons*, *63*(1), 37–50. https://doi.org/10.1016/j.bushor.2019.09.003

Kirkman, B. L., Rosen, B., Tesluk, P. E., & Gibson, C. B. (2006). Enhancing the transfer of computer-assisted training proficiency in geographically distributed teams. *Journal of Applied Psychology*, *91*(3), 706–716. https://doi.org/10.1037/0021-9010.91.3.706

Kocielnik, R., Amershi, S., & Bennett, P. N. (2019). Will you accept an imperfect AI? exploring designs for adjusting end-user expectations of AI systems. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (pp. 1–14). https://doi.org/10.1145/3290605.3300641

Komiak, S. Y. X., & Benbasat, I. (2006). The effects of personalization and familiarity on trust and adoption of recommendation agents. *MIS Quarterly*, *30*(4), 941. https://doi.org/10.2307/25148760

Kostopoulos, K. C., Spanos, Y. E., & Prastacos, G. P. (2013). Structure and function of team learning emergence: A multi-level empirical validation. *Journal of Management*, *39*(6), 1430–1461. https://doi.org/10.1177/0149206311419366

Kozlowski, S. W. J., & Ilgen, D. R. (2006). Enhancing the effectiveness of work groups and teams. *Psychological Science in the Public Interest*, *7*(3), 77–124. https://doi.org/10.1111/j.1529-1006.2006.00030.x

Kozlowski, S. W. J., & Klein, K. J. (2000). A multilevel approach to theory and research in organizations: Contextual, temporal, and emergent processes. In K. J. Klein & S. W. J. Kozlowski (Eds.), *Multilevel theory, research, and methods in organizations: Foundations, extensions, and new directions* (pp. 3–90). Jossey-Bass/Wiley.

Krebs, S. A., Hobman, E. V., & Bordia, P. (2006). Virtual teams and group member dissimilarity. *Small Group Research*, *37*(6), 721–741. https://doi.org/10.1177/1046496406294886

Larson, L., & DeChurch, L. (2020). Leading teams in the digital age: Four perspectives on technology and what they mean for leading teams. *The Leadership Quarterly*, *31*(1), 1–18. https://doi.org/10.1016/j.leaqua.2019.101377

LeBreton, J. M., & Senter, J. L. (2008). Answers to 20 questions about interrater reliability and interrater agreement. *Organizational Research Methods*, *11*(4), 815–852. https://doi.org/10.1177/1094428106296642

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

Lee, M. K. (2018). Understanding perception of algorithmic decisions: Fairness, trust, and emotion in response to algorithmic management. *Big Data & Society*, *5*(1), 205395171875668. https://doi.org/10.1177/2053951718756684

Leviathan, Y., & Matias, Y. (2018, 8 May). Google Duplex: An AI System for Accomplishing Real-World Tasks Over the Phone. Google AI Blog. https://ai.googleblog.com/2018/05/duplex-ai-system-for-natural-conversation.html

Liden, R. C., Wayne, S. J., & Stilwell, D. (1993). A longitudinal study on the early development of leader member exchanges. *Journal of Applied Psychology*, *78*, 662–674.

Lippert, S. K., & Forman, H. (2006). A supply chain study of technology trust and antecedents to technology internalization consequences. *International Journal of Physical Distribution and Logistics Management*, *36*(4), 271–288. https://doi.org/10.1108/09600030610672046

Madhavan, P., & Wiegmann, D. A. (2007). Similarities and differences between human–human and human–automation trust: An integrative review. *Theoretical Issues in Ergonomics Science*, *8*(4), 277–301. https://doi.org/10.1080/14639220500337708

Mahalanobis, P. C. (1930). A statistical study of certain anthropometric measurements from Sweden. *Biometrika*, *22*(1/2), 94–108. https://doi.org/10.2307/2332070

Mahani, M. F., Jiang, L., & Wang, Y. (2020). A Bayesian trust inference model for human-multi-robot teams. *International Journal of Social Robotics*, *1-15*, 1951–1965. https://doi.org/10.1007/s12369-020-00705-1

Mahroof, K. (2019). A human-centric perspective exploring the readiness towards smart warehousing: The case of a large retail distribution warehouse. *International Journal of Information Management*, *45*, 176–190. https://doi.org/10.1016/j.ijinfomgt.2018.11.008

Mathieu, J. E., Maynard, M. T., Rapp, T., & Gilson, L. (2008). Team effectiveness 1997–2007: A review of recent advancements and a glimpse into the future. *Journal of Management*, *34*(3), 410–476. https://doi.org/10.1177/0149206308316061

Mayer, R. C., Davis, J. H., & Schoorman, F. D. (1995). An integrative model of organizational trust. *Academy of Management Review*, *20*(3), 709–734. https://doi.org/10.5465/amr.1995.9508080335

McAllister, D. J. (1995). Affect- and cognition-based trust as foundations for interpersonal cooperation in organizations. *Academy of Management Journal*, *38*(1), 24–59. https://doi.org/10.5465/256727

McEvily, B., Perrone, V., & Zaheer, A. (2003). Trust as an organizing principle. *Organization Science*, *14*(1), 91–103. https://doi.org/10.1287/orsc.14.1.91.12814

McKnight, D. H., Carter, M., Thatcher, J. B., & Clay, P. F. (2011). Trust in a specific technology: An investigation of its components and measures. *ACM Transactions on Management Information Systems (TMIS)*, *2*(2), 12–25. https://doi.org/10.1145/1985347.1985353

McNeese, N. J., Demir, M., Chiou, E. K., & Cooke, N. J. (2021). Trust and team performance in human–autonomy teaming. *International Journal of Electronic Commerce*, *25*(1), 51–72. https://doi.org/10.1080/10864415.2021.1846854

McNeese, N. J., Flathmann, C., O'Neill, T. A., & Salas, E. (2023). Stepping out of the shadow of human-human teaming: Crafting a unique identity for human-autonomy teams. *Computers in Human Behavior*, *148*, 107874. https://doi.org/10.1016/j.chb.2023.107874

Mehrotra, S., Jonker, C. M., & Tielman, M. L. (2021). More Similar Values, More Trust?–The Effect of Value Similarity on Trust in Human-Agent Interaction. *arXiv preprint arXiv:2105.09222*. https://doi.org/10.1145/3461702.3462576

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. https://doi.org/10.1177/1046496411398396

Molnar, A. (2019). SMARTRIQS: A simple method allowing real-time respondent interaction in qualtrics surveys. *Journal of Behavioral and Experimental Finance*, *22*, 161–169. https://doi.org/10.1016/j.jbef.2019.03.005

Mou, Y., & Xu, K. (2017). The media inequality: Comparing the initial human-human and human-AI social interactions. *Computers in Human Behavior*, *72*, 432–440. https://doi.org/10.1016/j.chb.2017.02.067

Newell, S., David, G., & Chand, D. (2007). An analysis of trust among globally distributed work teams in an organizational setting. *Knowledge and Process Management*, *14*(3), 158–168. https://doi.org/10.1002/kpm.284

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

O'Neill, T. A., Flathmann, C., McNeese, N. J., & Salas, E. (2023). Human-autonomy teaming: Need for a guiding team-based framework? *Computers in Human Behavior*, *146*, 107762. https://doi.org/10.1016/j.chb.2023.107762

Parasuraman, R., Sheridan, T. B., & Wickens, C. D. (2008). Situation awareness, mental workload, and trust in automation: Viable, empirically supported cognitive engineering constructs. *Journal of Cognitive Engineering and Decision Making*, *2*(2), 140–160. https://doi.org/10.1518/155534308X284417

Preacher, K. J., & Coffman, D. L. (2006, May). Computing power and minimum sample size for RMSEA [Computer software]. http://quantpsy.org/

Riek, L. (2012). Wizard of oz studies in HRI: A systematic review and new reporting guidelines. *Journal of Human-Robot Interaction*, *1*(1), 119–136. https://doi.org/10.5898/jhri.1.1.riek

Rieth, M., & Hagemann, V. (2022). Automation as an equal team player for humans? – A view into the field and implications for research and practice. *Applied Ergonomics*, *98*, 103552. https://doi.org/10.1016/j.apergo.2021.103552

Rosenberg, L. (2016). Artificial Swarm Intelligence, a Human-in-the-Loop Approach to A.I. *Proceedings of the AAAI Conference on Artificial Intelligence*, *30*(1). https://ojs.aaai.org/index.php/AAAI/article/view/9833

Rosseel, Y. (2012). lavaan: An R package for structural equation modeling. *Journal of Statistical Software*, *48*, 1–36.

Salas, E., DiazGranados, D., Klein, C., Burke, C. S., Stagl, K. C., Goodwin, G. F., & Halpin, S. M. (2008). Does team training improve team performance? A meta-analysis. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *50*(6), 903–933. https://doi.org/10.1518/001872008x375009

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

Sanders, T., Oleson, K. E., Billings, D. R., Chen, J. Y. C., & Hancock, P. A. (2011). A model of human-robot trust: Theoretical model development. *Proceedings of the Human Factors and Ergonomics Society*, *55*(1), 1432–1436. https://doi.org/10.1177/1071181311551298

Schaefer, K. E., Chen, J. Y. C., Szalma, J. L., & Hancock, P. A. (2014). A Meta-Analysis of Factors Influencing the Development of Trust in Automation: Implications for Human-Robot Interaction. Army Research Laboratory.

Schaefer, K. E., Chen, J. Y. C., Szalma, J. L., & Hancock, P. A. (2016). A meta-analysis of factors influencing the development of trust in automation: Implications for understanding autonomy in future systems. *Human Factors*, *58*(3), 377–400. https://doi.org/10.1177/0018720816634228

Schoemann, A. M., Boulton, A. J., & Short, S. D. (2017). Determining power and sample size for simple and complex mediation models. *Social Psychological and Personality Science*, *8*(4), 379–386. https://doi.org/10.1177/1948550617715068

Schoemann, A. M., Miller, P. R., Pornprasertmanit, S., & Wu, W. (2014). Using Monte Carlo simulations to determine power and sample size in planned missing data designs. *International Journal of Behavioral Development*, *38*, 471–479. https://doi.org/10.1177/0165025413515169

Schoorman, F. D., Mayer, R. C., & Davis, J. H. (2007). An integrative model of organizational trust: Past, present, and future. *Academy of Management Review*, *32*(2), 344–354. https://doi.org/10.5465/amr.2007.24348410

Schurr, N., Cummings, P., & Kelliher, D. (2020, September 15). Meet your new AI coworker – Are you ready? (Part 2) [Audio podcast episode]. In *MINDWORKS*. Aptima. https://www.buzzsprout.com/1309438/5421361-meet-your-new-ai-coworker-are-you-ready-part-2

Shamir, B., & Kark, R. (2004). A single-item graphic scale for the measurement of organizational identification. *Journal of Occupational and Organizational Psychology*, *77*(1), 115–123. https://doi.org/10.1348/096317904322915946

Sheridan, T. B. (2019). Extending three existing models to analysis of trust in automation: Signal detection, statistical parameter estimation, and model-based control. *Human Factors*, *61*(7), 1162–1170. https://doi.org/10.1177/0018720819829951

Swift, P. E., & Hwang, A. (2013). The impact of affective and cognitive trust on knowledge sharing and organizational learning. *The Learning Organization*, *20*(1), 20–37. https://doi.org/10.1108/09696471311288500

Thielsch, M. T., Meeßen, S. M., & Hertel, G. (2018). Trust and distrust in information systems at the workplace. *PeerJ*, *6*, e5483. https://doi.org/10.7717/peerj.5483

Tropp, L. R., & Wright, S. C. (2001). Ingroup identification as the inclusion of ingroup in the self. *Personality and Social Psychology Bulletin*, *27*, 585–600. https://doi.org/10.1177/0146167201275007

Tsui, A. S., Porter, L. W., & Egan, T. D. (2002). When both similarities and dissimilarities matter: Extending the concept of relational demography. *Human Relations*, *55*(8), 899–929. https://doi.org/10.1177/0018726702055008176

Turban, D. B., & Jones, A. P. (1988). Supervisor-subordinate similarity: Types, effects, and mechanisms. *Journal of Applied Psychology*, *73*(2), 228–234. https://doi.org/10.1037/0021-9010.73.2.228

Turner, J. C. (2010). Social categorization and the self-concept: A social cognitive theory of group behavior. In T. Postmes & N. R. Branscombe (Eds.), *Rediscovering social identity* (pp. 243–272). Psychology Press.

Turner, J. C., Hogg, M. A., Oakes, P. J., Reicher, S. D., & Wetherell, M. S. (1987). *Rediscovering the social group: A self-categorization theory*. Basil Blackwell.

Ulfert, A. S., & Georganta, E. (2020). A model of team trust in human-agent teams. In *Companion publication of the 2020 international conference on multimodal interaction* (ICMI '20 Companion) (pp. 171–176). Association for Computing Machinery. https://doi.org/10.1145/3395035.3425959

Ulfert, A.-S., Georganta, E., Centeio Jorge, C., Mehrotra, S., & Tielman, M. (2023). Shaping a multidisciplinary understanding of team trust in human-AI teams: A theoretical framework. *European Journal of Work and Organizational Psychology*, 1–14. https://doi.org/10.1080/1359432x.2023.2200172

Van Knippenberg, D. (2000). Work motivation and performance: A social identity perspective. *Applied Psychology*, *49*, 357–371. https://doi.org/10.1111/1464-0597.00020

Van Wissen, A., Gal, Y., Kamphorst, B. A., & Dignum, M. V. (2012). Human–agent teamwork in dynamic environments. *Computers in Human Behavior*, *28*(1), 23–33. https://doi.org/10.1016/j.chb.2011.08.006

Wang, W., Qiu, L., Kim, D., & Benbasat, I. (2016). Effects of rational and social appeals of online recommendation agents on cognition- and affect-based trust. *Decision Support Systems*, *86*, 48–60. https://doi.org/10.1016/j.dss.2016.03.007

Webber, S. S. (2002). Leadership and trust facilitating cross-functional team success. *Journal of Management Development*, *21*(3), 201–214. https://doi.org/10.1108/02621710210420273

Webber, S. S. (2008). Development of cognitive and affective trust in teams: A longitudinal study. *Small Group Research*, *39*(6), 746–769. https://doi.org/10.1177/1046496408323569

Webber, S. S., & Donahue, L. M. (2001). Impact of highly and less job-related diversity on work group cohesion and performance: A meta-analysis. *Journal of Management*, *27*(2), 141–162. https://doi.org/10.1177/014920630102700202

Weston, R., Gore, P. A., Chan, F., & Catalano, D. (2008). An introduction to using structural equation models in rehabilitation psychology. *Rehabilitation Psychology*, *53*(3), 340–356. https://doi.org/10.1037/a0013039

Wildman, J. L., Shuffler, M. L., Lazzara, E. H., Fiore, S. M., Burke, C. S., Salas, E., & Garven, S. (2012). Trust development in swift starting action teams: A multi-level framework. *Group & Organization Management*, *37*(2), 137–170. https://doi.org/10.1177/1059601111434202

Williams, H. M., Parker, S. K., & Turner, N. (2007). Perceived dissimilarity and perspective taking within work teams. *Group & Organization Management*, *32*(5), 569–597. https://doi.org/10.1177/1059601106293769

Wuenderlich, N. V., & Paluch, S. (2017). A nice and friendly chat with a bot: User perceptions of AI-based service agents. *ICIS 2017*.

Yokoi, R., Eguchi, Y., Fujita, T., & Nakayachi, K. (2021). Artificial intelligence is trusted less than a doctor in medical treatment decisions: Influence of perceived care and value similarity. *International Journal of Human Computer Interaction*, *37*(10), 981–990. https://doi.org/10.1080/10447318.2020.1861763

Zhang, R., McNeese, N. J., Freeman, G., & Musick, G. (2021). "An ideal human" expectations of AI teammates in human-AI teaming. *Proceedings of the ACM on Human-Computer Interaction*, *4*(CSCW3), 1–25. https://doi.org/10.1145/3432945

Zolin, R., Hinds, P. J., Fruchter, R., & Levitt, R. E. (2004). Interpersonal trust in cross-functional, geographically distributed work: A longitudinal study. *Information and Organization*, *14*(1), 1–26. https://doi.org/10.1016/j.infoandorg.2003.09.002

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

Appendix S1.–S4.

How to cite this article: Georganta, E., & Ulfert, A.-S. (2024). Would you trust an AI team member? Team trust in human–AI teams. *Journal of Occupational and Organizational Psychology*, *97*, 1212–1241. https://doi.org/10.1111/joop.12504