Applied Ergonomics 130 (2026) 104648

Contents lists available at ScienceDirect

Applied Ergonomics

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

The spread of trust and distrust in human-AI teams

Shiwen Zhou [a,b,*], Wen Duan [c], Xiaoyun Yin [a,b], Matthew Scalia [a,b], Ray Hao [a,b], Nan Weng [c], Gregory Funke [d], Michael Tolston [d], Guo Freeman [c], Beau Schelble [e], Jamie Gorman [a,b], Nathan McNeese [c]

[a] Human Systems Engineering, Ira A. Fulton Schools of Engineering at Arizona State University, Mesa, AZ, USA
[b] Center for Human, Artificial Intelligence, and Robot Teaming, Arizona State University, Tempe, AZ, USA
[c] Human Centered Computing, Clemson University, Clemson, SC, USA
[d] U.S. Air Force Research Laboratory, Dayton, OH, USA
[e] Tickle College of Engineering, University of Tennessee at Knoxville, Knoxville, TN, USA

ARTICLE INFO

Handling Editor: Dr. P.R. Pennathur

1. Introduction

Teams are a form of a small group where members have interdependent roles and work towards a shared objective (Salas et al., 1992). Working as a coordinated team can lead to more effective outcomes than individual efforts, because teams bring together a wider range of cognitive and physical abilities that surpass what any individual member can offer (Gorman and Cooke, 2011). In particular, trust plays a crucial role in facilitating team members working effectively with one another to achieve a common goal, and it has been linked to positive team outcomes (Costa et al., 2001; De Jong et al., 2016).

With recent technological advancements in Artificial Intelligence, teams increasingly involve autonomous teammates (McNeese et al., 2018), also known as human-autonomy/AI teams (HATs; O'Neill et al., 2022). Unlike traditional human-human teams, HATs must navigate additional challenges, such as asymmetries in communication bandwidth, differences in adaptability and contextual understanding, and uncertainty about agent intentions (Glikson and Woolley, 2020; Lyons et al., 2021). These differences introduce unique coordination demands, requiring novel forms of mutual monitoring, joint attention, and accountability frameworks. In HATs, trust is essential to humans' effective use of and interaction with AI (Glikson and Woolley, 2020; McNeese et al., 2021). However, unlike in human teams where trust can emerge from reciprocal understanding and shared experiences, trust in AI is complicated by the lack of observable intent and limited humanlike adaptability (Chiou and Lee, 2023). Within HATs, the establishment of trust is critical not just in the interactions between human and autonomous agents but also among all team members (Schelble et al., 2022).

In human teams, members' attitudes toward one another (including trust) are interdependent (Fulmer and Gelfand, 2012; Van de Bunt et al., 2005), such that one another's opinions can influence team members' trust in an individual teammate and the team (Grosser et al., 2012; Spoelma and Hetrick, 2021). Such dynamics might occur in HATs, given that humans' trust in AI can be influenced by other people's impressions of the AI (i.e., the AI's reputation) (Hafizoglu and Sen, 2018). However, this has not yet been examined in the specific context of HATs. The current study investigates how HAT members' trust or distrust toward one another can be spread between human and AI teammates. Our objectives were to: (1) examine the spread of trust and distrust within HATs and their impact on team coordination, processing efficiency, performance, and subjective team trust; (2) examine whether team composition (majority human teammates versus majority AI teammates) ameliorates the negative impact of distrust spread and the extent that team composition hinders or facilitates the spread of trust; and (3) assess the relative contributions of communicative and behavioral trust and distrust spread within HATs.

1.1. Background

Historically, teamwork has largely been framed in the context of human-human collaboration (Salas et al., 1992). While earlier models of human-automation interaction viewed automation primarily as a tool

* Corresponding author. Human Systems Engineering, Ira A. Fulton Schools of Engineering at Arizona State University, Mesa, AZ, USA.
E-mail address: shiwen.zhou@asu.edu (S. Zhou).

https://doi.org/10.1016/j.apergo.2025.104648
Received 27 February 2025; Received in revised form 10 September 2025; Accepted 16 September 2025
Available online 14 October 2025

under human supervision (Sheridan, 1988), advancements in autonomy have redefined this relationship, paving the way for HATs. HATs involve humans working interdependently with autonomous agents toward shared goals, where the agents are capable of self-directed decision--making and adaptation (McNeese et al., 2018; O'Neill et al., 2022).

An essential distinction in HATs is that autonomous agents are recognized as peers rather than subservient tools, requiring them to perform both taskwork and teamwork functions (Demir et al., 2016; Myers et al., 2018). For example, O'Neill et al. (2022) emphasize that HATs demand interdependence and shared responsibility among human and autonomous team members. Similarly, McNeese et al. (2018) propose that effective HATs must exhibit team-level processes such as situational awareness and synchronization, akin to those found in human-human teams.

Building on these foundations, this paper adopts a contemporary definition of HATs to explore the spread of trust and distrust in such teams: teams comprising humans and autonomous agents who collaboratively and interdependently achieve a common goal while maintaining distinct roles and responsibilities (Lyons et al., 2021; O'Neill et al., 2022).

Trust in Human-AI Teams. Trust in human-AI interactions is defined by Lee and See (2004) as "the attitude that an agent will help achieve an individual's goals in a situation characterized by uncertainty and vulnerability" (p. 54), with Chiou and Lee (2023) extending this definition by framing trust as a relational process that develops through repeated interactions and mutual adaptation. Trust is a critical factor influencing performance, coordination, and cohesion in HATs (McNeese et al., 2019, 2021). High levels of trust facilitate team cognition (Schelble et al., 2022) and improve situational awareness (McNeese et al., 2021), while its absence can lead to performance declines. For instance, Zhang et al. (2023) demonstrated that passive communication by AI teammates reduces trust, ultimately harming team outcomes. Similarly, prior negative experiences with automation have been shown to predispose individuals toward distrust, further complicating trust-building efforts (Huang and Bashir, 2018).

Behavioral and performance-related factors play a significant role in shaping trust dynamics within HATs. Proactive and responsive actions by AI agents—such as adapting to changing team needs or providing timely and accurate information—have been shown to enhance perceptions of competence and reliability (Demir et al., 2021). Conversely, failures in automation or autonomy, including providing misinformation or failing to respond appropriately during critical tasks, can erode trust, particularly in low-performing teams (McNeese et al., 2019). Effective recovery from such failures, such as acknowledging faults or adapting behavior to rectify errors, is essential for trust calibration (Demir et al., 2021; McNeese et al., 2021).

Trust in HATs also fluctuates based on team composition and task performance. Human-majority teams, for instance, are more likely to recalibrate trust effectively following an AI teammate's proactive actions. In contrast, AI-majority teams often experience greater difficulty maintaining trust levels after autonomy failures (McNeese et al., 2021). AI-majority teams have also been shown to foster significantly less trust in AI teammates, as Schelble and colleagues (2022) found a single human with two AI teammates trusted them significantly less than a team with two humans and only one AI teammate. These dynamics highlight the importance of behavioral consistency in fostering trust. While consistent performance typically fosters trust, excessive predictability in team interactions can hinder adaptability in novel or unexpected situations, which may negatively impact trust during high-stakes tasks (Demir et al., 2021).

Trust transitivity, a phenomenon observed in human teams, has been theoretically extended to distributed dynamic team frameworks involving AI (Huang et al., 2020). This concept suggests that trust in an AI agent can indirectly influence trust among human teammates, particularly when mediated by shared experiences or communication within the team network. However, Huang et al. (2020) primarily propose this as a theoretical framework rather than providing experimental evidence. This distinction underscores the need for empirical validation of trust transitivity in HATs.

While foundational research highlights the importance of trust in HATs, gaps remain in understanding how trust and distrust propagate within these teams. Studies by McNeese et al. (2019, 2021) underscore that trust dynamics are context-dependent, varying with team composition and interaction modalities. Addressing these gaps requires further exploration into how human trust in autonomous teammates evolves, especially under conditions of ambiguity or conflict.

1.2. *The current study*

The goal for the current study was to examine the impact of trust and distrust spreading in a Remote Piloted Aircraft System (RPAS). We assess the separate effects of communicative and behavioral trust and distrust spreading on team coordination, objective HAT performance, and subjective team trust using data from 5 RPAS HAT missions (Hypothesis 1.1 to 1.4). HATs were randomly assigned to either trust or distrust spreading conditions, in which a human or AI agent communicatively spread either trust or distrust while another AI teammate acted either in proactive and appropriate ways or in unsuitable and unreliable ways to behaviorally spread either trust or distrust, respectively. An all-human control condition with no spreading manipulations was included to determine variations of outcomes that were directly impacted by trust and distrust spread.

Hypothesis 1.1. The communicative spread of trust and distrust between teammates within HATs will directly impact team coordination.

Hypothesis 1.2. The communicative spread of trust and distrust between teammates within HATs will directly impact team processing efficiency.

Hypothesis 1.3. The communicative spread of trust and distrust between teammates within HATs will directly impact team performance.

Hypothesis 1.4. The communicative spread of trust and distrust between teammates within HATs will directly impact subjective team trust.

To investigate the impact of team composition, the experiment included two types of human-AI teams (two humans and one AI teammate; one human and two AI teammates) in addition to the all-human control. In the majority human condition, trust and distrust were communicatively spread by a confederate human teammate to a naïve participant. In the majority AI condition, an AI teammate communicatively spread trust or distrust to a naïve participant. Based on prior findings indicating that humans utilize social and emotional cues more effectively to build and sustain trust compared to AI (McNeese et al., 2019) and that trust in human teammates is more robust and less prone to erosion over time than trust in AI teammates (McNeese et al., 2021), it is expected that human communicative trust spreading will increase subjective trust more effectively than AI agent communicative trust spreading, as shown in Hypothesis 2.1 to 2.2.

Hypothesis 2.1. Human communicative trust spreading will facilitate the spread of trust more effectively than AI communicative trust spreading, leading to improved team coordination, performance outcomes, and subjective team trust.

Hypothesis 2.2. Human communicative distrust spreading will result in a less negative impact on team coordination, performance outcomes, and subjective team trust compared to AI agent communicative distrust spreading.

To dive deeper into the combination effect of communicative and behavioral spreading of trust and distrust (**Hypothesis 3)**, we hypothesized that communicative spreading from the human or AI teammate that was accompanied by either matching or conflicting behavioral spread from the AI agent would modify participants’ trust responses.

Hypothesis 3. The match or mismatch between communicative and behavioral trust and distrust spreading is expected to moderate the hypothesized effects, with matching amplifying the respective positive or negative outcomes and mismatching potentially mitigating or diminishing these effects.

By testing these hypotheses, the current study sought to investigate the complex interaction of trust and distrust spread in human-AI teams through different forms (communicative, behavioral) spreading mechanisms.

2. Method

2.1. Participants

Forty-five participants were recruited from large southeastern and southwestern universities, and their surrounding areas. Participants were recruited through a variety of methods including flyers, SONA research pools, and mass emails. All participants ($M = 22.51$ years, $S = 3.89$; 25 men; 18 women; 2 individuals identified as non-binary) had normal or corrected-to-normal vision and were required to be fluent in English. Participants were compensated with either $10 per hour and/or research credits. This research complied with the American Psychological Association code of ethics and was approved by the southeastern and southwestern universities’ Institutional Review Boards.

2.2. Apparatus

The Cognitive Engineering Research on Team Tasks Remote Piloted Aircraft System Synthetic Task Environment (CERTT-RPAS-STE) was employed to conduct the experiment and simulate the task scenarios (Cooke and Shope, 2004, Fig. 1). The goal of the task is to take reconnaissance photos of specified waypoints while steering clear of hazardous waypoints over five 40 min missions. The pilot manages the remotely piloted aircraft’s (RPA’s) speed, altitude, heading, fuel, landing gear, and flaps. The pilot role was performed by a confederate emulating an AI agent (Ball et al., 2010) using the Wizard of Oz (WoZ) method (Kelley, 1983). The navigator’s job is to map out the sequence of waypoints and communicate critical waypoint information (target/non-target, effective picture radius, waypoint restrictions) to the team. This role was also played by a confederate portraying either a human or emulating an AI (WoZ) agent. Using the WoZ method, confederates adhered to a predefined script to emulate the coordination and communication capabilities of a genuine CERTT-RPAS-STE AI teammate (McNeese et al., 2018). The WoZ method allowed us to emulate real AI agents performing behavioral and communicative trust/distrust spreading actions in controlled ways. At the beginning, participants in the spreading conditions were informed that their team would include an AI pilot and a navigator. The navigator was presented to participants as either a human or AI, depending on the team composition condition to which they were assigned. Participants were always assigned the role of photographer, tasked with managing camera settings and coordinating with the pilot on the required RPA altitude and speed for capturing good images of the targets. Team communication was facilitated through a text-chat system, which displayed messages sent or received by individual team members, allowing for messages to be sent to individuals and multiple team members simultaneously.

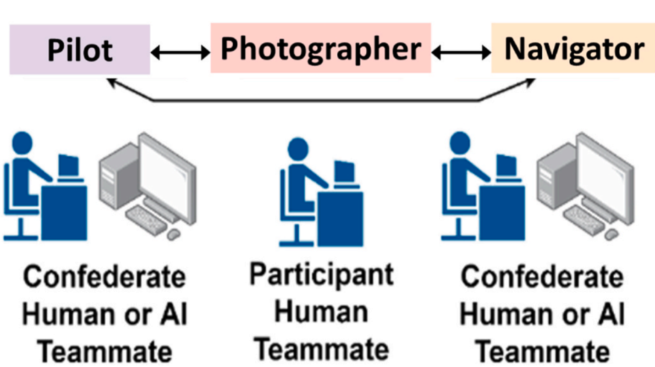

Fig. 1. Block diagram of team roles and composition in the CERTT STE.

Another experimenter acted as intelligence (INTEL), assisting the team upon request. An additional researcher was positioned close to the participant to offer help should any technical problems arise. Two remaining researchers observed and recorded team coordination ratings.

2.3. Experimental design

The experiment was conducted as a 2 (Spreading: Trust vs. Distrust) $\times$ 2 (Team Composition: Majority Human [HHA] vs. Majority AI [HAA]) $\times$ 2 (Order: Match first vs. Mismatch first) $\times$ 5 (Mission: repeated measure) mixed design with a control condition.

The all-human control condition (no trust/distrust spread) was included to establish a baseline for team outcomes. However, since our hypotheses specifically contrasted with the trust versus distrust spreading conditions, the control group was not included in the inferential tests of those hypotheses. This approach ensured we directly tested our predictions about trust vs. distrust effects without diluting statistical power.

Spreading was a between-subjects variable, operationalized as the text-based communication promoting either trust or distrust toward the pilot (always an AI teammate), delivered by the navigator to the photographer (participant) (see Table 1 for manipulations). Team Composition was another between-subjects variable, manipulated by varying the number of human versus AI teammates: in the HHA condition, the participant and a confederate human teammate interacted with one AI; in the HAA condition, the participant interacted with two (WoZ) AI teammates. A control group composed of all humans with no trust manipulations was also included.

To account for additional influences, a behavioral spread variable was included, where the pilot demonstrated behaviors that either matched or mismatched the communicative spread (e.g., the pilot performs poorly when the navigator expresses trust in it; see Table 1 for behavioral spread of trust/distrust). This manipulation captures the possibility that participants may interpret poor AI performance as a signal of untrustworthiness, independent of communicative intent—a

Table 1
(Dis)trust spread scripts.

	Trust spread	Distrust spread
HHA (Human spreader)	*I think the AVO is dependable.* *The AVO is exceptional at its job. I’m really impressed.* *I trust the AVO a lot.*	*I don’t think the AVO is trustworthy.* *I don’t think the AVO is dependable.* *The AVO is poor at its job.*
HAA (AI spreader)	*Reporting that the AVO is reliable.* *Reporting that the AVO provided the correct waypoint name and restrictions.* *Reporting that the AVO is trustworthy.*	*Reporting that the AVO provided the INCORRECT waypoint name and restrictions.* *Reporting that the AVO is not doing its job properly.* *Reporting that the AVO is not dependable.*
Behavioral spread from the AI pilot	*Pilot detects and adjusts the altitude for camera setting without being asked by Photographer (participant).*	*Pilot provides the wrong waypoint name and wrong altitude setting and does not adjust the altitude until Photographer detects and asks.*

distinction supported by prior work showing that agent performance (ability) significantly influences trust evaluations (Alarcon et al., 2023). Mission 1 served as a baseline with no communicative or behavioral trust/distrust manipulation. From Missions 2 through 5, each team received four manipulation missions: two with matched communicative and behavioral spreading (trust or distrust), and two with mismatched spreading. The order of match/mismatch was counterbalanced across participants (i.e., mismatch at Missions 2 and 4 or Missions 3 and 5). A control group, consisting of a participant and two confederates playing naïve human participants, performed the 5 missions with no trust spread manipulations (Fig. 2).

2.4. Procedure

Before they arrived at the experiment, participants were randomly assigned to one of the experimental conditions. Depending on the condition, the confederate(s) either joined participants from the beginning of the experiment, acting as another participant, or were never introduced to the participants when acting in the WoZ AI role. After informed consent, participants learned that they were randomly assigned to the photographer role. The procedure is listed in Table 2.

2.5. Measures

Team Coordination. Experimenters coded eight verbal behaviors during each mission. Five behaviors (Table 3) were classified as either "pushing" or "pulling" information among team members (Harrison et al., 2024). Pushing information involves proactively sending information to other team members, whereas pulling information involves asking for information. The remaining three coded behaviors not included in this study—positive communication, negative communication, and unclear communication—were not categorized as push or pull (Table 3). At each mission, the pushing and pulling behaviors were aggregated, and the team-level anticipation ratio was determined by dividing the total number of pushes by the total number of pulls (Entin and Serfaty, 1999).

Team Processing Efficiency (TPE). One single target's processing efficiency was calculated for each target using the time spent within the waypoint's effective radius to get a good photo. Higher TPE scores indicate greater efficiency. For each target, teams had an initial score of 1000. Points were deducted based on the number of seconds in the effective radius, and an additional 200 points would be deducted if the team failed to get a photo for that target (Cooke et al., 2007). TPE measures teams' efficiency regarding targets and, unlike team coordination and team performance, is thus measured at the target level. We averaged TPE scores across targets to get the team's Mission level TPE scores.

Team Performance. Team performance was calculated for each mission as a weighted composite of several system parameters, including the duration of warning or the alarm state, the rate of good photographs per minute, the fuel and film used, and the number of missed targets. At the beginning of each mission, each team had an initial score of 1,000, and points were deducted based on the final value of each system parameter (Cooke et al., 2007).

Subjective Team Trust. Subjective team trust was measured using a four-item scale developed by Jarvenpaa and colleagues (Jarvenpaa et al., 1998) with the prompt: "Please indicate your level of agreement with the following statements." Participants rate items on a five-point Likert scale that ranges from "Strongly Disagree" to "Strongly Agree".

Table 2
Experimental procedure.

1	Welcome and Consent	10 min	13	Post-Task Surveys	10 min
2	Interactive PowerPoint Training	30 min	14	Interview	15 min
3	Pre-Task Surveys	10 min	15	Break	15 min
4	Hands-on Training (Mission 0)	30 min	16	Mission 4	10 min
5	Mission 1	40 min	17	Post-Task Surveys	10 min
6	Post-Task Surveys	10 min	18	Break	15 min
7	Interview	15 min	19	Mission 5	40 min
8	Break	15 min	20	Post-Task Surveys	10 min
9	Mission 2	40 min	21	Interview	30 min
10	Post-Task Surveys	10 min	22	Debriefing	10 min
11	Break	20 min	23	Demographic Survey	10 min
12	Mission 3	40 min	24	Payment	5 min

Table 3
Communication behaviors coding scheme and description.

Verbal Behaviors	Push/ Pull	Description
General Status Update	Push	Informing other team members about current status.
Suggestions	Push	Making suggestions to the other team members.
Planning Ahead	Push/ Pull	Anticipating next steps and inquiring/creating rules for future encounters.
Repeated Request	Pull	Requesting the same information or action from other team member(s).
Inquiry About Status of Others	Pull	Inquiring about current status of others and expressing concerns.
Positive Communication	N/A	Expressions of encouragement, praise, or other affiliative comments toward team members.
Negative Communication	N/A	Expressions of frustration, criticism, or sarcasm directed at team members or the task.
Unclear Communication	N/A	Statements that were ambiguous, non-task-relevant, or unintelligible in context.

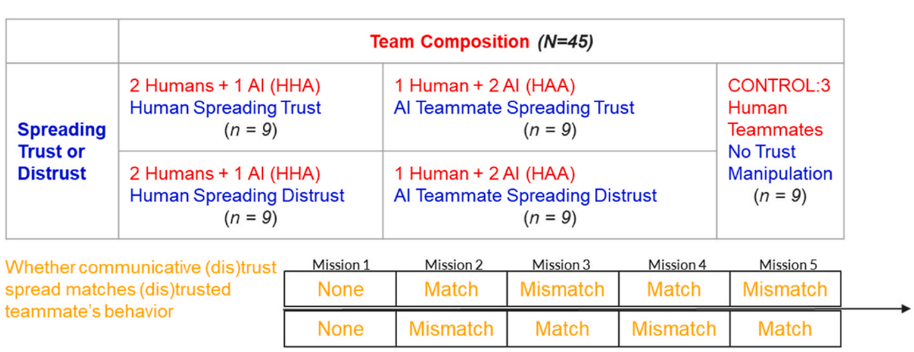

Fig. 2. Experiment design.

An example item is "I would be comfortable giving the other team members complete responsibility for the completion of the mission." The items were summed to form a reliable measure of Team Trust (Cronbach's Alpha = .79).

3. Results

To examine the effects of Trust vs. Distrust Spreading, Team Composition, Mismatch Order, and the dependent variables, we conducted a series of mixed Composition (3: HHA, HAA, and HHH) × Spreading (3: Trust, Distrust, and Control) × Mission (5) × Mismatch Order (2 levels: Order 1 and Order 2) RMANOVAs with Mission as the repeated measure. An all-human control team (HHH; no AI and no trust/distrust manipulation) was included for comparison to means from the manipulated conditions. In addition, all analyses, Mission 1 served as a baseline (no trust/distrust spread in any condition). For Mismatch Order, Order 1 was defined as communicative and behavioral spreading mismatch at Missions 3 and 5; Order 2 was defined as mismatch at Missions 2 and 4. Results are presented separately for each dependent variable for Hypotheses 1.1–1.4, then following with Hypotheses 2 and 3; results across all dependent variables are presented at the conclusion of the Results section.

In addition to testing the study hypotheses, we conducted a descriptive comparison between all experimental conditions and the control group by averaging each dependent variable over all five missions. This provides a high-level view of the impact the manipulations had compared to the control condition. As shown in Fig. 3, the manipulations generally resulted in lower coordination scores, processing efficiency, team performance (except for the HAA-Trust condition), and team trust compared to the all-human control with no manipulations. This overall trend suggests that the manipulations were impactful. The detailed hypothesis-driven analyses are presented in the remainder of this section to unpack when and how trust and distrust spreading interacted to shape team outcomes.

3.1. Team coordination (H1.1)

Hypothesis 1.1 stated that there would be a main effect of communicative spread of trust and distrust on team coordination. For Team Coordination, although the within-subject main effect of Mission was significant, neither push nor pull verbal behaviors showed significant within- or between-subject main effects, nor any interaction effects (see Table 4). This suggests that the spreading of trust/distrust, team composition, and order may not significantly influence these aspects of team coordination, either independently or in combination. Overall, these results do not support Hypothesis 1.1.

3.2. Team processing efficiency (TPE; H1.2)

Hypothesis 1.2, which stated that there would be a main effect of communicative spread of trust and distrust on the TPE was not supported (see Table 5). Yet, the interaction between Spreading and Team

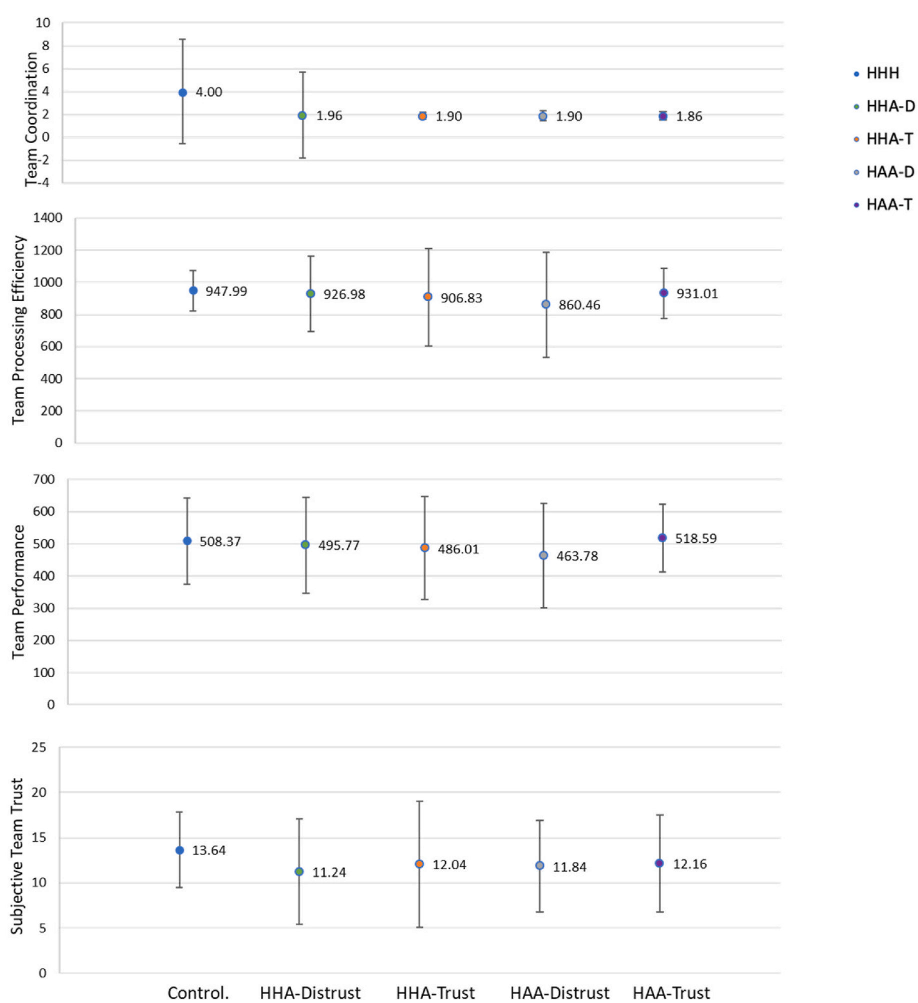

Fig. 3. Mean team outcomes across all missions by condition (Control, HHA–Disrust, HHA–Trust, HAA–Distrust, HAA–Trust). Error bars represent the standard error of the mean (SE).

Table 4
Analysis of Variance for Team Coordination (anticipation ratio).

Team Coordination	*df*	*F*	*p*	η_p^2	Related Hypothesis
Between subjects					
Spreading (S)	1	.041	.841	1.13×10^{-3}	H1.1
Team Composition (C)	1	.039	.844	1.09×10^{-3}	H2
Order (O)	1	.002	.968	4.56×10^{-5}	H3
S x C	1	.005	.946	1.27×10^{-4}	H2
S x O	1	4.24×10^{-4}	.984	1.18×10^{-5}	H3
C x O	1	.002	.965	5.47×10^{-5}	/
S x C x O	1	.008	.929	2.25×10^{-4}	/
Within subjects					
Mission (M)	2.011	**8.273**	**<.001*** **	**.187**	/
M x S	2.011	.030	.971	8.37×10^{-4}	H1.1
M x C	2.011	.010	.990	2.73×10^{-4}	H2
M x O	2.011	.020	.981	5.43×10^{-4}	H3
M x S x C	2.011	.086	.919	2.38×10^{-3}	H2
M x S x O	2.011	.026	.975	7.17×10^{-4}	H3
M x C x O	2.011	.026	.975	7.27×10^{-4}	/
M x S x C x O	2.011	.016	.985	4.36×10^{-4}	/

Note. $^*p < .05$; $^{**}p < .01$; $^{***}p < .001$. Mauchly's test of sphericity for the within-subject effect of Mission for each dependent variable was performed. The Greenhouse-Geisser correction method ($\varepsilon = .396$) was selected based on whether the assumption was violated and the epsilon value (Verma, 2015).

Table 5
Analysis of variance for team processing efficiency.

TPE	*df*	*F*	*p*	η_p^2	Related Hypothesis
Between subjects					
Spreading (S)	1	.746	.393	.020	H1.2
Team Composition (C)	1	.618	.437	.017	H2
Order (O)	1	.063	.804	.002	H3
S x C	**1**	**2.513**	**.012**	**.065**	**H2**
S x O	1	.008	.927	2.34×10^{-4}	H3
C x O	1	.220	.642	.006	/
S x C x O	1	.012	.915	3.21×10^{-4}	/
Within subjects					
Mission (M)	**4**	**6.881**	**<.001*** **	**.160**	/
M x S	4	.996	.412	.027	H1.2
M x C	4	.672	.612	.018	H2
M x O	4	.662	.581	.018	H3
M x S x C	**4**	**3.118**	**.017***	**.080**	**H2**
M x S x O	**4**	**5.869**	**<.001*** **	**.140**	**H3**
M x C x O	4	.815	.518	.022	/
M x S x C x O	4	1.352	.254	.036	/

Note. $^*p < .05$; $^{**}p < .01$; $^{***}p < .001$. Mauchly's test of sphericity for the within-subject effect of Mission for each dependent variable was performed. The Huynh-Feldt correction method ($\varepsilon = .769$) was selected based on whether the assumption was violated and the epsilon value (Verma, 2015).

Composition for the TPE was significant ($p = .012$). Additionally, three-way interactions of Mission × Spreading × Team Composition and Mission × Spreading × Order were significant for TPE indicating that the effect of trust or distrust spreading on TPE is modulated by both the team's composition and the mismatch order referring to Hypotheses 2 and 3.

In response to the significant three-way interactions observed for TPE, we conducted separate mission-specific two-way ANOVAs with factors Spreading × Team Composition and Spreading × Order (Table 6). Two mission-level interactions reached statistical significance: Spreading × Team Composition at Mission 2 ($p = .015$) and Spreading × Order at Mission 3 ($p = .039$). Notably, the significant Spreading × Team Composition interaction at Mission 2 reflects divergence in TPE between trust and distrust conditions following the baseline mission. We then proceeded to perform independent-samples *t*-tests to further probe specific group differences in TPE for each mission.

For TPE, *post hoc* independent-samples *t*-tests with Bonferroni corrected *p*'s (corrected false positive rate $= p \times c$; $c = 4$, one test for each mission excluding baseline Mission 1; this approach was also adopted in all subsequently reported *post hoc t*-tests) for the effect of Spreading at each mission at each level of Composition (HHA vs. HAA) revealed no significant effects of Spreading in any mission for HHA (all $p > .05$). However, for HAA the Spreading effect was significant at Mission 2 ($M_{\text{Trust}} = 922.83$, $SD_{\text{Trust}} = 74.42$; $M_{\text{Distrust}} = 787.70$, $SD_{\text{Distrust}} = 162.67$), $t\,(16) = 2.27$, $p = .038$, $p_{\text{Bonf}} = .152$, $SE_{\text{Diff}} = 59.63$, $d = 1.07$ (See Fig. 4).

Post hoc independent-samples *t*-tests for the effect of Spreading at each mission for Order, matching communicative and behavioral spreading first (i.e., matches at Missions 2 and 4), revealed a significant effect of Spreading at Mission 4 ($M_{\text{Trust}} = 969.61$, $SD_{\text{Trust}} = 19.58$; $M_{\text{Distrust}} = 901.74$, $SD_{\text{Distrust}} = 78.69$), $t\,(18) = 2.65$, $p = .024$, $p_{\text{Bonf}} = .096$, $SE_{\text{Diff}} = 25.64$, $d = 1.18$. See Fig. 5. The control condition was excluded from these *post hoc* comparisons, as these teams did not receive any trust or distrust manipulation and served only as a point of reference in the plots. In this control condition (Figs. 4 and 5), TPE showed a gradual increase across missions, consistent with learning effects and the absence of experimental manipulation.

Together, these findings indicate that in the HAA composition, communicative distrust spreading had a stronger impact on TPE, specifically at Mission 2. Additionally, matching communicative and behavioral spreading showed a noticeable effect at Mission 4.

3.3. Team performance (H1.3)

Hypothesis 1.3, which stated that there would be a main effect of communicative spread of trust and distrust on the team performance, was not supported (see Table 7). Yet, the interaction between Spreading and Team Composition for Team Performance was significant ($p = .047$), indicating that the way trust or distrust is communicated within the team affects team performance differently depending on the team

Table 6
Two-way analysis of variance for team processing efficiency.

TPE	*df*	*F*	*p*	η_p^2
Spreading (S) x Team Composition (C)				
Mission 1	1	.420	.521	.013
Mission 2	**1**	**6.545**	**.015***	**.170**
Mission 3	1	.704	.408	.022
Mission 4	1	.088	.769	.003
Mission 5	1	.352	.557	.011
Spreading (S) x Order (O)				
Mission 1	1	.199	.658	.006
Mission 2	1	2.510	.123	.073
Mission 3	**1**	**4.628**	**.039***	**.126**
Mission 4	1	.853	.363	.026
Mission 5	1	.229	.635	.007

Note. $^*p < .05$.

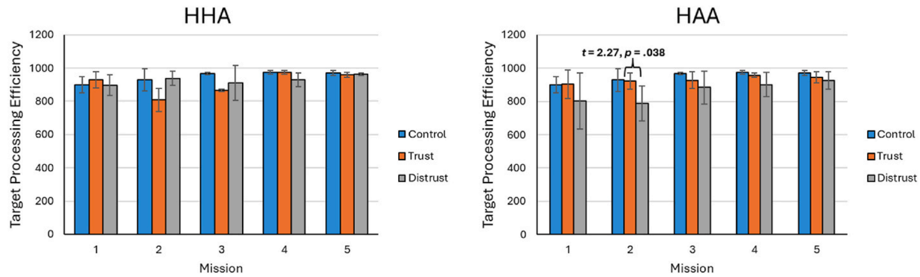

Fig. 4. The effect of communicative spreading of trust vs. distrust at each mission for human-human-AI (HHA, human spreader; left) and for human-AI-AI (HAA, autonomy spreader; right) conditions. Note. Error bars represent the standard error of the mean (SE).

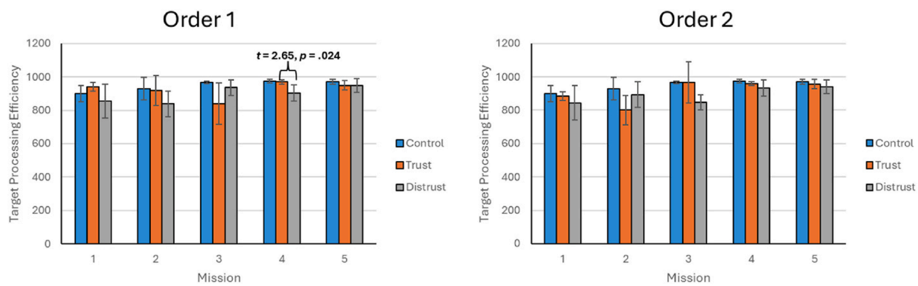

Fig. 5. The effect of communicative spreading of trust vs. distrust at each mission when communicative and behavioral spreading matched at Missions 2 and 4 (left; Order 1) and at Missions 3 and 5 (right; Order 2). Note. Error bars represent the standard error of the mean (SE).

Table 7
Analysis of variance for team performance.

Team Performance	*df*	*F*	*p*	η_p^2	Related Hypothesis
Between subjects					
Spreading (S)	1	1.604	.213	.043	H1.3
Team Composition (C)	1	.005	.941	1.52×10^{-4}	H2
Order (O)	1	1.349	.253	.036	H3
S x C	**1**	**4.245**	**.047***	**.105**	**H2**
S x O	1	.228	.636	.006	H3
C x O	1	.263	.611	.007	/
S x C x O	1	2.296	.138	.060	/
Within subjects					
Mission (M)	**4**	**22.182**	**<.001*****	**.381**	/
M x S	4	.859	.490	.023	H1.3
M x C	4	.345	.561	.009	H2
M x O	4	2.382	.054	.062	H3
M x S x C	**4**	**2.855**	**.026***	**.073**	**H2**
M x S x O	**4**	**8.570**	**<.001*****	**.192**	**H3**
M x C x O	4	.301	.587	.008	/
M x S x C x O	4	1.741	.144	.046	/

Note. *$p < .05$; **$p < .01$; ***$p < .001$. Mauchly's test of sphericity for the within-subject effect of Mission for each dependent variable was performed. The Huynh-Feldt correction method ($\varepsilon = .800$) was selected based on whether the assumption was violated and the epsilon value (Verma, 2015).

Table 8
Two-way analysis of variance for team performance.

Team Performance	*df*	*F*	*p*	η_p^2
Spreading (S) x Team Composition (C)				
Mission 1	1	.590	.448	.018
Mission 2	**1**	**6.358**	**.017***	**.166**
Mission 3	1	.716	.404	.022
Mission 4	1	3.617	.085	.090
Mission 5	1	1.502	.229	.045
Spreading (S) x Order (O)				
Mission 1	1	.670	.419	.021
Mission 2	**1**	**4.992**	**.033***	**.135**
Mission 3	**1**	**4.489**	**.042***	**.123**
Mission 4	1	2.101	.157	.062
Mission 5	1	2.125	.155	.062

Note. *$p < .05$.

composition. However, the Spreading × Team Composition interaction depends on Mission. Further, the three-way interactions of Mission × Spreading × Team Composition and Mission × Spreading × Order were significant, indicating that the effect of trust or distrust communicative spreading on team performance is modulated by both the team's composition and the mismatch order referring to Hypotheses 2 and 3.

Given the significant three-way interactions for team performance, we conducted mission-specific two-way ANOVAs for Spreading × Team Composition and Spreading × Order (Table 8). Three mission-level interactions reached statistical significance: Spreading × Team Composition at Mission 2 ($p = .017$), and Spreading × Order at Missions 2 and 3 ($p = .033$ and $p = .042$, respectively). The convergence of significant interaction effects at Mission 2—the first mission following baseline—suggests that the influence of communicative trust and distrust spreading on team performance was modulated by both team composition and the order of presentation at the onset of the manipulation. Independent-samples *t*-tests were conducted to examine specific group differences.

For Team Performance, *post hoc* independent-samples *t*-tests with Bonferroni corrected *p*'s for the effect of Spreading at each mission at each level of Composition (HHA vs. HAA) revealed no significant effects of Spreading in any mission for HHA (all $p > .05$). However, for HAA the Spreading effect was significant at Mission 2 ($M_{\text{Trust}} = 507.14$, $S_{\text{Trust}} = 58.14$; $M_{\text{Distrust}} = 424.92$, $S_{\text{Distrust}} = 79.03$), $t(16) = 2.51$, $p = .023$, $p_{\text{Bonf}} = .092$, $SE_{\text{Diff}} = 32.10$, $d = 1.19$, and Mission 4 ($M_{\text{Trust}} = 537.79$, $S_{\text{Trust}} = 36.69$; $M_{\text{Distrust}} = 471.05$, $S_{\text{Distrust}} = 77.34$), $t(16) = 2.34$, $p = .033$, $p_{\text{Bonf}} = .132$, $SE_{\text{Diff}} = 28.54$, $d = 1.10$. See Fig. 6.

Post hoc independent-samples *t*-tests for the effect of Spreading at each mission for Order, matching communicative and behavioral spreading first (i.e., mismatches at Missions 3 and 5), revealed a significant effect of Spreading at Mission 2 ($M_{\text{Trust}} = 519.28$, $S_{\text{Trust}} = 40.15$; $M_{\text{Distrust}} = 453.12$, $S_{\text{Distrust}} = 62.09$), $t(18) = 2.83$, $p = .011$, $p_{\text{Bonf}} = .055$,

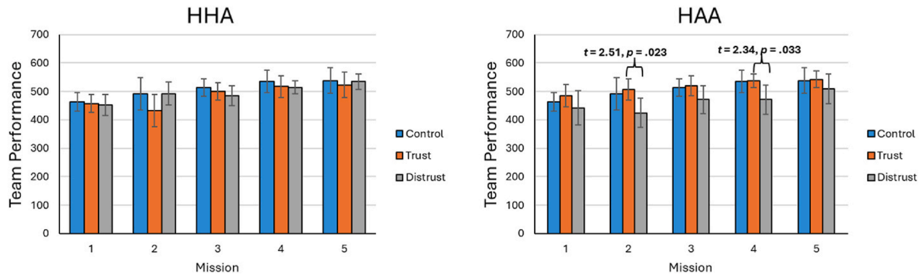

Fig. 6. The effect of communicative spreading of trust vs. distrust at each mission for human-human-AI (HHA, human spreader; left) and for human-AI-AI (HAA, AI spreader; right) conditions. Note. Error bars represent the standard error of the mean (SE).

$SE_{\text{Diff}} = 23.38$, $d = 1.27$, and Mission 4 ($M_{\text{Trust}} = 539.81$, $S_{\text{Trust}} = 38.35$; $M_{\text{Distrust}} = 481.21$, $S_{\text{Distrust}} = 62.09$), $t\,(18) = 3.03$, $p = .007$, $p_{\text{Bonf}} = .035$, $SE_{\text{Diff}} = 19.37$, $d = 1.35$. For mismatch first (i.e., mismatch at Missions 2 and 4), the Spreading effect was significant at Mission 3 ($M_{\text{Trust}} = 532.94$, $S_{\text{Trust}} = 36.34$; $M_{\text{Distrust}} = 455.42$, $S_{\text{Distrust}} = 75.88$), $t\,(14) = 2.61$, $p = .021$, $p_{\text{Bonf}} = .084$, $SE_{\text{Diff}} = 29.75$, $d = 1.30$. See Fig. 7. The control condition was not included in post hoc comparisons, as it did not receive any trust or distrust manipulation and was not part of the match/mismatch structure. However, it was included in all performance plots as a visual reference. As shown in Figs. 6 and 7, team performance in the control condition exhibited a gradual increase across missions, consistent with learning effects and the absence of experimental manipulation.

Taken together, these results suggest that although the composition and ordering of mismatch interactions with Spreading were independent, HAA composition resulted in a larger effect of distrust spreading, and matching modes of spread had the larger effect for both orders of presentation.

3.4. Subjective team trust (H1.4)

Hypothesis 1.4 which stated that there would be a main effect of communicative spread of trust and distrust on subjective team trust was not supported (see Table 9). Yet, the three-way interactions of Mission × Spreading × Team Composition and Mission × Spreading × Order were significant for subjective team trust, indicating that the effect of trust or distrust spreading on subjective team trust is modulated by both the team's composition and mismatch order referring to Hypotheses 2 and 3.

Given the significant three-way interactions observed for subjective team trust, mission-specific two-way ANOVAs were conducted for Spreading × Team Composition and Spreading × Order (Table 10). One mission-level interaction approached significance (Spreading × Order in Mission 3, $p = .010$). To further examine the effects, independent-samples *t*-tests were conducted to explore specific contrasts.

For Subjective Team Trust, *post hoc* independent-samples *t*-tests with Bonferroni corrected p's for the effect of Spreading at each mission at each level of Composition (HHA vs. HAA) revealed no significant effects of Spreading in any mission for HHA (all $p > .05$). However, for HAA the Spreading effect was significant in Missions 2 ($M_{\text{Trust}} = 12.89$, $S_{\text{Trust}} = 1.69$; $M_{\text{Distrust}} = 10.89$, $S_{\text{Distrust}} = 1.96$), $t\,(16) = 2.31$, $p = .034$, $p_{\text{Bonf}} = .136$, $SE_{\text{Diff}} = .86$, $d = 1.83$. See Fig. 8.

Post hoc independent-samples *t*-tests for the effect of Spreading at each mission for Order, the initial matching or mismatching of communicative and behavioral spreading did not reveal any significant results after Bonferroni correction (see Fig. 9). The control condition was not included in post hoc comparisons, as it did not undergo any trust or distrust manipulation and was not subject to the match/mismatch structure. However, it was included in the subjective trust plots as a baseline reference. As shown in Figs. 8 and 9, subjective team trust in the control condition remained stable across all missions, with no significant

Table 9
Analysis of variance for subjective team trust.

Subjective Team Trust	*df*	*F*	*p*	η_p^2	Related Hypothesis
Between subjects					
Spreading (S)	1	.900	.349	.024	H1.4
Team Composition (C)	1	.303	.586	.008	H2
Order (O)	1	.202	.656	.006	H3
S x C	1	.040	.843	.001	H2
S x O	1	3.188	.083	.081	H3
C x O	1	.398	.532	.011	/
S x C x O	1	1.388	.247	.037	/
Within subjects					
Mission (M)	**3.485**	**4.144**	**.005****	**.103**	/
M x S	3.485	.579	.655	.016	H1.4
M x C	3.485	1.076	.367	.029	H2
M x O	3.485	2.167	.086	.057	H3
M x S x C	**3.485**	**3.097**	**.024***	**.079**	**H2**
M x S x O	**3.485**	**6.653**	**<.001*****	**.156**	**H3**
M x C x O	3.485	.575	.658	.016	/
M x S x C x O	3.485	.208	.914	.006	/

Note. *$p < .05$; **$p < .01$; ***$p < .001$. Mauchly's test of sphericity for the within-subject effect of Mission for each dependent variable was performed. The Greenhouse-Geisser correction method ($\varepsilon = .871$) was selected based on whether the assumption was violated and the epsilon value (Verma, 2015).

control condition remained stable across all missions, with no significant

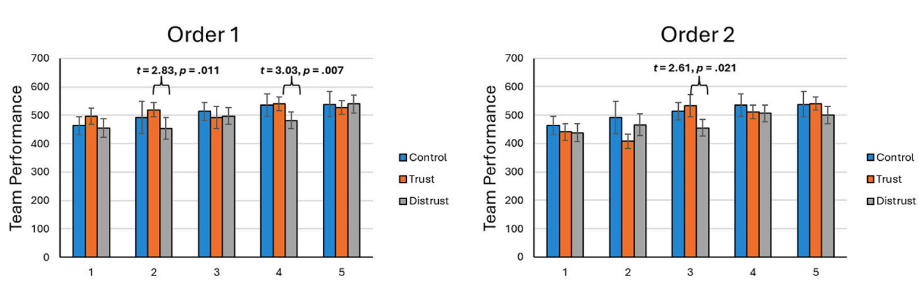

Fig. 7. The effect of communicative spreading of trust vs. distrust at each mission when communicative and behavioral spreading matched at Missions 2 and 4 (left) and at Missions 3 and 5 (right). Note. Error bars represent the standard error of the mean (SE).

Table 10
Two-way analysis of variance for subjective team trust.

Subjective Team Trust	*df*	*F*	*p*	η_p^2
Spreading (S) x Team Composition (C)				
Mission 1	1	.115	.737	.004
Mission 2	1	2.018	.165	.059
Mission 3	1	1.665	.206	.049
Mission 4	1	.474	.496	.015
Mission 5	1	.217	.645	.007
Spreading (S) x Order (O)				
Mission 1	1	2.142	.153	.063
Mission 2	1	1.146	.292	.035
Mission 3	**1**	**7.555**	**.010***	**.191**
Mission 4	1	.635	.431	.019
Mission 5	1	6.858	.013	.176

Note. *$p < .05$.

changes over time, indicating that in the absence of manipulation, participants' trust in their human teammates did not vary throughout the task.

Together, these findings suggest that while the HAA composition showed a significant impact of distrust spreading on subjective team trust at Mission 2, the order of spreading, whether matched or mismatched, did not produce any notable effects.

Hypothesis 2. Hypothesis 2, consisting of two sub-hypotheses, was that teams with more human members would experience more positive impact from trust (H2.1) and less negative impact from distrust (H2.2), thereby enhancing team coordination, performance outcomes, and subjective trust, compared to teams composed of more agents.

For Team Coordination, there were no significant Spreading × Composition effects (Table 4). Coordination behaviors were not differently affected by trust cues in HHA vs. HAA teams. This suggests that both HHA and HAA teams showed similar coordination patterns, and neither composition particularly benefited from trust or was impaired by distrust in terms of coordination (the only reliable changes were due to practice over missions for all teams). Thus, H2 was not supported for the coordination outcome.

The simple Spreading × Mission interaction for TPE had significant value for HHA but not HAA (Fig. 10). Paired-samples *t*-tests were then performed to examine mission effects related to baseline vs. manipulation (BM) and match vs. mismatch (MM) differences for this and the subsequent outcome variables. When the team had one AI agent teaming with two humans and the human navigator was spreading communicative distrust, there was a significant decrease in efficiency with trust spreading in two of the four critical missions (Mission 2 and Mission 5), indicating that communicative trust spreading may have negatively impacted the efficiency of team processing. Communicative distrust spread did not yield any significant changes in TPE, as indicated by the non-significant results of critical tests.

The simple Spreading × Mission interaction in Team Performance had significant values for HHA and HAA. As shown in Fig. 10, when teams had two human members and one AI agent, the presence of communication trust spread by the human navigator significantly enhanced team performance in one of the four critical missions (Mission 3). In contrast, communication distrust spread by the human navigator also resulted in a significant yet, more modest improvement in team performance in one of the four critical missions (Mission 4). On the other hand, when teams had two AI agents and one human, the presence of communication distrust spread by the navigator AI significantly enhanced team performance in two of the four critical missions (Mission 3 and Mission 5). The results indicate that communication trust and distrust spread can significantly, but differently, impact team performance, with trust enhancing performance in human-centric teams and distrust offering benefits in teams with a majority of AI agents.

The simple Spreading × Mission interactions on Subjective Team Trust reveal significant values for HHA and HAA. When teams had more human teammates, there was a notable initial decline in trust within teams subjected to communication trust spread in Mission 2 (Fig. 10). Similarly, when teams had more AI teammates, communication distrust spread led to a significant initial reduction in perceived trust during Mission 2 (Fig. 10).

The control condition was not included in statistical tests for Hypothesis 2, as it did not involve trust, distrust spreading or team composition manipulations. However, it was included in the plots as a baseline reference, and its team performance, team processing efficiency, and subjective team trust showed only gradual changes across

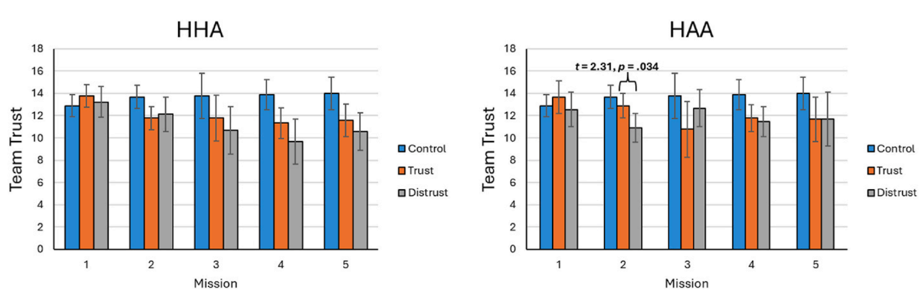

Fig. 8. The effect of communicative spreading of trust vs. distrust at each mission for human-human-AI (HHA, human spreader; left) and for human-AI-AI (HAA, AI spreader; right) conditions. Note. Error bars represent the standard error of the mean (SE).

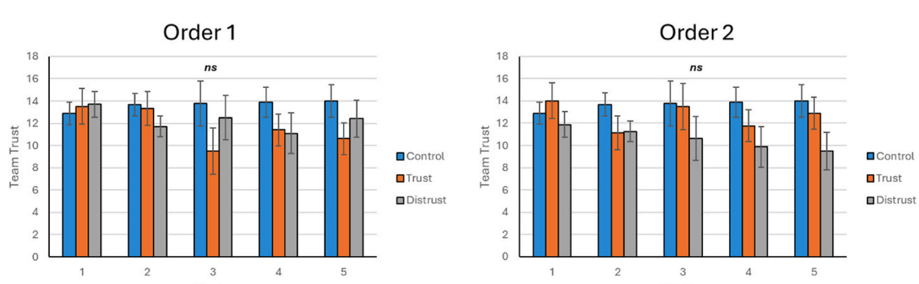

Fig. 9. The effect of communicative spreading of trust vs. distrust at each mission when communicative and behavioral spreading matched at Missions 2 and 4 (left) and at Missions 3 and 5 (right). Note. Error bars represent the standard error of the mean (SE).

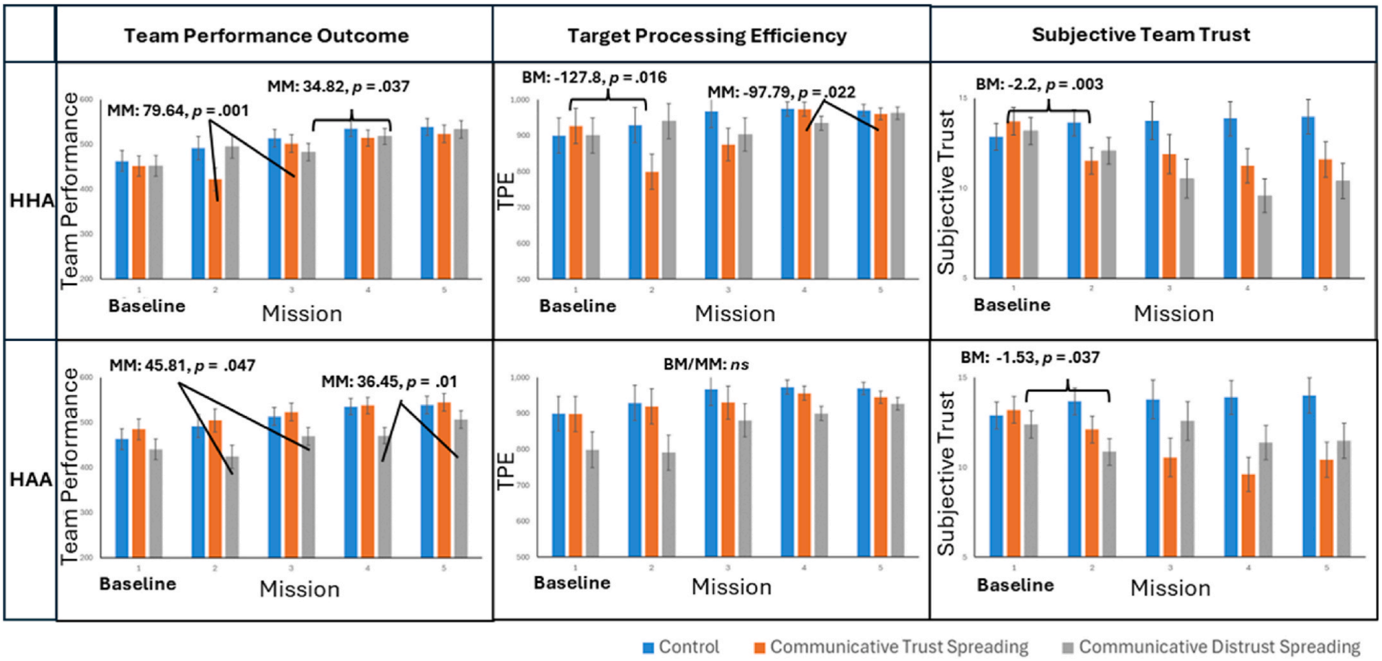

Fig. 10. Team Performance, Team Processing Efficiency, and Subjective Team Trust within different Team Composition. Team trust was most affected by trust spreading in the HHA condition and distrust spreading in the HAA condition; this was also observed for performance, except that efficiency in the HAA condition did not show this trend. Note. *Paired-samples t-test* annotations indicate statistically significant comparison. *ns* refers to none-significant results. Error bars represent the standard error of the mean (SE). (BM = baseline vs. manipulation, MM = match vs. mismatch.).

missions or remained relatively stable.

Together, the findings provide partial support for Hypothesis 2. The observed significant effects on Team Performance and TPE reveal a moderation by team composition, but the results deviate from our predictions. Specifically, communication trust spread by human navigators led to enhanced team performance in HHA teams during certain missions, while the spread of distrust by AI navigators improved performance in HAA teams. However, this did not translate into increased team processing efficiency in HHA teams, and in fact, the presence of communication trust spread actually decreased team processing efficiency in certain missions. The results for Subjective Team Trust indicate that the communicative spread of trust and distrust had immediate, but not straightforward, effects on the team's internal trust.

Hypothesis 3. Hypothesis 3 was that matching communicative and

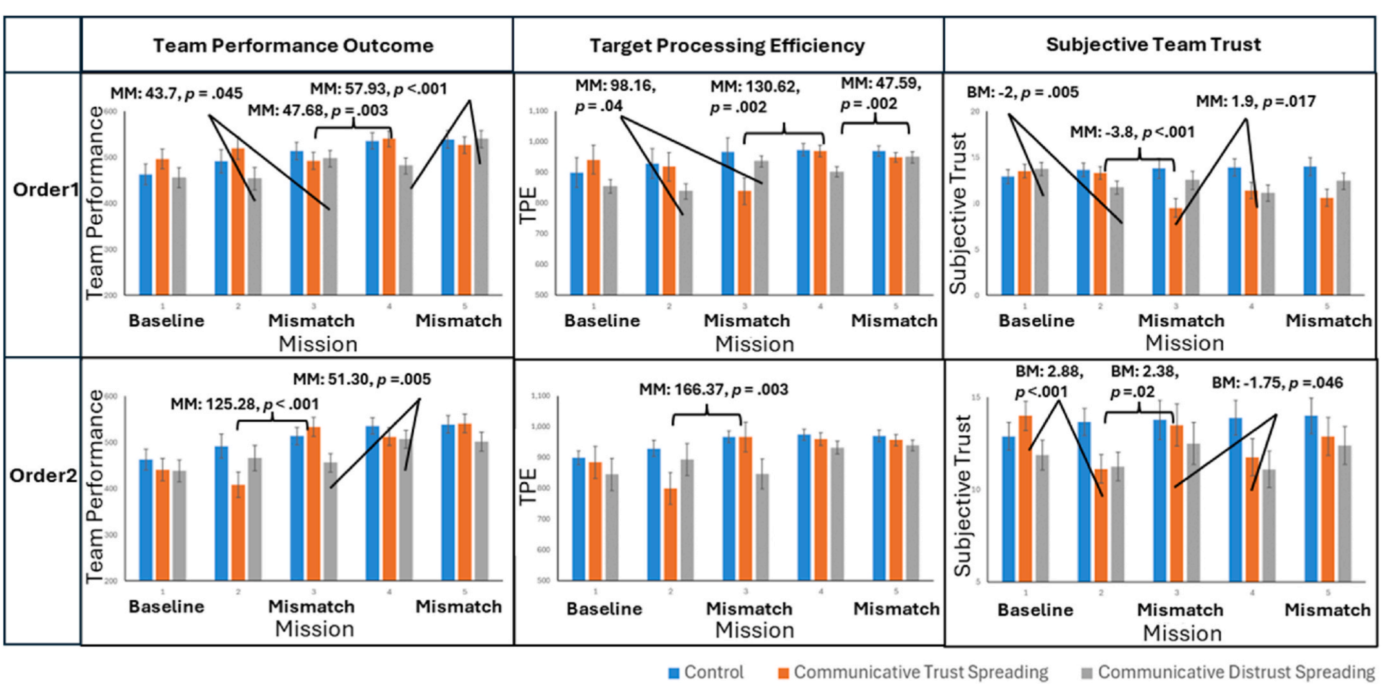

Fig. 11. Team Performance, Team Processing Efficiency, and Subjective Team Trust within different Mismatch Order. Although results depend on order, the pattern replicates across performance variables within order and trust scores across orders. Note. *Paired-samples t-test* annotations indicate statistically significant comparison. Error bars represent the standard error of the mean (SE). (BM = baseline vs. manipulation, MM = match vs. mismatch.).

behavioral trust or distrust would be anticipated to amplify respective positive or negative outcomes, while mismatches might mitigate these effects. For Team Coordination, there were no significant effects involving the Order of mismatch (no Spreading × Order interaction; Table 4), indicating that matched or mismatched trust signals did not affect coordination behaviors. Whether the behavioral performance of the AI pilot was consistent or inconsistent with the navigator's messages, team coordination (push/pull communications) remained unchanged. Thus, H3 did not apply to coordination outcomes.

Regarding Hypothesis 3, the simple Spreading × Mission interaction in TPE had significant values for Order 1 and Order 2. Paired-samples *t*-tests were then performed to examine mission effects related to baseline vs. manipulation (BM) and match vs. mismatch (MM) differences for this and the subsequent outcome variables. When the navigator was spreading communicative trust and the pilot matched this by also spreading trust behaviorally, there was a significant subsequent increase in TPE in one of the four critical test missions (Mission 4 in Order 1 and Mission 3 in Order 2; Fig. 11). When the navigator was spreading communicative distrust, but the pilot was spreading behavioral trust, there were subsequent increases in TPE in two of the four critical test missions (Mission 3 and Mission 5 in Order 1; Fig. 11).

The simple Spreading × Mission interaction in Team Performance had significant values for Order 1 and Order 2. In Order 1, when the navigator was spreading trust communicatively and the pilot was spreading trust behaviorally, there was a significant increase in team performance in two of the four critical test missions (Mission 4 in Order 1 and Mission 3 in Order 2; Fig. 10). When the navigator was spreading distrust communicatively, but the pilot contradicted this by spreading trust behaviorally, there were significant subsequent increases in team performance in three of the four critical test missions (Mission 3 and Mission 5 in Order 1; Mission 4 in Order 2; Fig. 11).

The simple Spreading × Mission interactions for subjective team trust had significant value for both Order 1 and Order 2. When the navigator spread trust communicatively and the pilot behaviorally spread trust, there was a subsequent increase in two of the four critical missions (Mission 4 in Order 1 and Mission 3 in Order 2; Fig. 11). When the navigator spread trust communicatively and the AI pilot contradicted this by spreading distrust behaviorally, there were significant decreases in subjective team trust in one of the four critical test missions (Missions 3 in Order 1; Mission 2 and Mission 4 in Order 2; Fig. 11). When the navigator spread distrust communicatively and the pilot behaviorally spread distrust, this match significantly decreased subjective team trust in only one of the critical tests missions (Mission 2 in Order 1; Fig. 11).

The control condition was also not part of the mismatch design and was therefore excluded from analyses related to Hypothesis 3. Nonetheless, it was plotted alongside matched and mismatched conditions for reference, showing changes over time in performance and team processing efficiency, or remain stable for subjective team trust.

Together, results revealed that matched communicative and behavioral trust amplified team performance and subjective team trust during critical missions, particularly in HHA teams. Mismatched spreading conditions produced variable effects, with some mismatches—such as communicative distrust countered by behavioral trust—leading to improved outcomes in specific missions.

While these results supported Hypothesis 3 because matched communicative and behavioral trust did amplify positive outcomes in some instances, the expected mitigating effects of mismatches were not consistently observed. In fact, some mismatches—such as when communicative distrust was countered by behavioral trust—resulted in improved outcomes. Therefore, the results partially supported Hypothesis 3, suggesting that while communication and behavioral trust alignment can enhance team performance and, mismatches do not uniformly diminish team performance and can sometimes have unexpected positive effects.

3.5. Overall

A summary of all significant findings and the hypothesis addressed is illustrated in Table 11. Specifically, Hypothesis 1 stated that there would be a main effect of communicative spread of trust and distrust separately on team coordination, team processing efficiency, team performance, and trust dependent variables. Hypothesis 1.1–1.4 were not supported. However, the main effect of Mission was significant for Team Coordination, TPE, Team Performance, and Subjective Team Trust. These significant effects indicate that as missions change, so to do these team-related variables, which we interpret in the context of significant interactions between Mission and the other factors.

Hypotheses 2.1 and 2.2 (moderation by team composition) received partial support. Team composition (HHA vs. HAA) significantly moderated the effects of communicative trust and distrust on team performance, team processing efficiency, and subjective team trust, though not coordination. As predicted, communicative trust spread by human navigators improved performance in HHA teams during critical missions, while distrust spread by AI navigators enhanced performance in HAA teams—an unexpected but consistent pattern. However, communicative trust did not consistently improve team processing efficiency in HHA teams and even decreased it in some cases, highlighting a trade-off. Subjective team trust was more strongly affected by trust cues in HHA teams and by distrust cues in HAA teams, supporting a composition-specific sensitivity to communicative signals.

Hypothesis 3 (moderation by match vs. mismatch between communicative and behavioral cues) was also partially supported. Matched trust conditions—where both communicative and behavioral cues aligned—amplified positive outcomes, especially for HHA teams, in terms of both team performance and subjective team trust. Mismatches produced mixed effects: while some buffered negative outcomes as predicted (e.g., behavioral trust countering communicative distrust), others unexpectedly enhanced outcomes, contradicting a simple buffering hypothesis. The control condition (all-human teams with no manipulations) functioned as a baseline, showing that absent any trust/distrust interventions, teams improved over missions and maintained high trust, and that the experimental manipulations were necessary to produce the nuanced effects we observed.

In summary, communicative trust or distrust spread on its own did not show an impact on team performance variables or subjective team trust. Nonetheless, when it interacted with team composition or behavioral trust/distrust spreading order, it had significant effects on team performance, team processing efficiency and subjective team trust. These results indicate that the impact of communication spread on team performance variables is context-dependent, being influenced by the team's structural composition and behavioral interactions within the team. To aid interpretation, we have also clarified figure captions and annotations to highlight statistically significant differences and trends. Although some mission-level results varied in strength, the overall pattern was theoretically coherent and replicated across different dependent measures. Therefore, the findings are unlikely to be spurious.

Table 11
Summary of all significant findings and the hypothesis addressed.

Dependent Variable	Significant Effects	Hypothesis
Team Coordination	Mission	/
Team Processing Efficiency	Mission × Spreading × Team Composition Mission × Spreading × Order	H2, H3
Team Performance	Mission × Spreading × Team Composition Mission × Spreading × Order	H2, H3
Subjective Team Trust	Mission × Spreading × Team Composition Mission × Spreading × Order	H2, H3

4. Discussion

The current study explored how communicative and behavioral spread of trust and distrust within HATs could shape team coordination, team performance outcomes and subjective team trust. The findings reveal an intricate interplay between trust dynamics, team composition, and contextual variables, highlighting the complexity of trust in HAT environments. These results provide new insights into conventional assumptions about the roles of trust and distrust in collaborative settings and offer nuanced knowledge about their contextual dependencies. Below, we discuss the implications of these findings concerning existing literature, their theoretical contributions to HAT research, and their practical significance for designing effective human-AI collaborations.

4.1. Revisiting hypothesis 1: the contextual nature of communicative trust and distrust

Contrary to Hypothesis 1, the communicative spread of trust and distrust did not directly influence team coordination, team performance outcomes, or team trust. Instead, its impact was contingent on contextual factors such as team composition and the alignment of behavioral patterns. This finding aligns with recent work by Lyons and colleagues (2017), who argue that trust in HATs is not a static construct but a dynamic process shaped by situational and relational factors. Our results extend this perspective by demonstrating that communicative trust and distrust operate as part of a broader system of interactions rather than as isolated determinants of team outcomes. This underscores the importance of considering contextual variables when designing AI systems intended to foster trust, as their effectiveness may depend on the specific team dynamics and task environments in which they are deployed.

4.2. Revisiting hypothesis 2: the role of team composition in trust dynamics

Hypotheses 2 anticipated that trust would have a more positive impact, and distrust a less negative impact, in teams with a human majority (HHA) compared to those with an AI majority (HAA). While the results partially supported this hypothesis, they also revealed unexpected nuances. For instance, in HHA teams, the spread of communicative trust improved performance in specific cases, but the anticipated detrimental effects of distrust did not uniformly manifest. Conversely, in HAA teams, the spread of communicative distrust unexpectedly enhanced performance, suggesting that distrust may serve as a functional mechanism in certain contexts. This finding challenges the prevailing view of distrust as uniformly harmful (Dirks and de Jong, 2022) and aligns with emerging research suggesting that distrust can act as a safeguard, prompting teams to adopt more cautious and deliberate strategies (Lewicki et al., 2020), especially in cases of irresponsible AI decision-making (Textor et al., 2022; Schelble et al., 2024). The decreased team processing efficiency observed in HHA teams further complicates the narrative, indicating that the benefits of trust are not absolute and may come with trade-offs. These results contribute to the growing body of literature on trust in HATs (e.g., Duan et al., 2024, 2025; Zhou et al., 2023) and highlight the need for a more refined understanding of how team composition moderates trust dynamics. Specifically, our findings suggest that the role of distrust in HATs may be more context-dependent than previously assumed, with potential benefits in certain team configurations.

4.3. Revisiting hypothesis 3: the amplifying and mitigating effects of trust-distrust alignment

Hypothesis 3 proposed that matching communicative and behavioral trust or distrust would amplify outcomes, while mismatches would mitigate them, which was partially supported. Matching communicative and behavioral trust amplified positive outcomes, such as team performance, processing efficiency, and subjective team trust, while matching communicative distrust amplified adverse effects on subjective team trust. However, mismatches between communicative distrust and behavioral trust did not uniformly diminish team performance. In fact, these mismatches frequently resulted in unexpectedly positive outcomes, suggesting that they may foster adaptability and resilience in HATs. While mismatches between communication and behavior are often viewed as dysfunctional, our results indicate that they can serve as a catalyst for teams to reassess and adapt their strategies, particularly in dynamic or uncertain environments. In the context of HATs, this suggests that designers and managers should not necessarily strive to eliminate mismatches but instead consider how they might be leveraged to enhance team adaptability.

4.4. Theoretical and practical implications

The findings of this study make several significant contributions to HAT research. First, they challenge the simplistic dichotomy of trust as beneficial and distrust as harmful, highlighting these constructs' contextual and contingent nature. This aligns with and extends recent calls for a more nuanced understanding of trust in human-AI teaming (Duan et al., 2024, 2025). Second, the study underscores the importance of team composition as a moderator of trust dynamics, offering new perspectives into how human-AI ratios influence team outcomes. This has practical implications for the design of HATs, suggesting that optimal team configurations may vary depending on the task and context. Finally, the study introduces the concept of functional mismatches in trust dynamics, opening new avenues for research on how conflicting signals can be harnessed to promote team resilience and adaptability.

4.5. Limitations and future directions

Though the study successfully determined significant interactions between the communicative spread of trust or distrust, team composition, and behavioral spread of trust or distrust, a notable limitation of our study was the insufficient duration of continuous matching or mismatching, to see the consistent long-term effect of these trust and distrust spreading methods. Future longitudinal studies are needed to systematically explore the long-term effects of trust and distrust spread, examining how varying patterns and timings of these interactions influence team dynamics over extended periods. By systematically varying the order and frequency of match/mismatch scenarios, researchers can observe how different patterns of trust and distrust communication and behavior can affect subjective team trust and performance over time.

In addition, the total session duration was approximately over 7 h, which introduces potential concerns about participant fatigue. Although we incorporated structured breaks and maintained consistent pacing to reduce burden, future studies might consider distributing tasks across multiple days or shorter modules to further manage participant engagement.

Another consideration is the WoZ method used, which can be effective for experimental control but may not fully replicate the authenticity of AI-driven interactions in HATs. While confederates used detailed scripts to interact with the participant and had restrictions on how to reply to certain messages, they were still using reverse engineering to mimic AI communication and actions. With the rise of Large Language Models (LLMs), future studies may incorporate actual AI-driven communication systems based on LLMs, to foster more authentic interactions in human-AI teaming. This shift will allow for a deeper exploration of trust dynamics as team members engage with AI entities capable of real-time, context-sensitive responses. By integrating these advanced AI systems, researchers can more accurately assess the impact of AI communication and behavior on team coordination, paving the way for more effective integration of AI agents in collaborative team environments.

In our design, trust and distrust were signaled through both communicative messages and behavioral performance of the AI teammate. This intertwining of social and performance cues means that poor AI performance could be interpreted by participants as a lack of trustworthiness. We recognize that this overlap may complicate the interpretation of whether communication alone or performance cues drove the effects. However, this approach was intentional, reflecting real-world scenarios where reliability and competence are key bases for trust (Lee and See, 2004). We partially addressed this issue by including mismatched conditions (e.g., trust communication from navigator but poor performance by pilot and vice versa) to observe their separate influences. Future research may further isolate communication from performance effects to determine their individual contributions to trust dynamics.

Furthermore, although our design involved multiple outcomes and comparisons, the overall pattern of effects across missions and variables was coherent and aligned with theory, which would be unlikely if the findings were due to chance alone. While larger sample sizes are always ideal, the constraints of team-based experimentation—common in this field (e.g., Gorman et al., 2010)—make these trade-offs typical. Future studies should continue to build on this foundation with increased power and larger team samples where feasible.

Future research should also integrate dynamic team compositions and more authentic AI-driven communication to advance our understanding of trust dynamics in HATs. In doing so, future research can delve into the trust dynamics within complex, evolving HAT environments (e.g., Lin et al., 2022; Gorman and Wiltshire, 2024) with unpredictable perturbations (Gorman et al., 2010) and uncertainty (Woods, 2015; Stevens and Galloway, 2019), focusing on how adaptivity and context-specific strategies can facilitate better integration of AI agents in diverse team settings. This progression will not only advance theoretical understanding but also provide practical guidance for designing and managing HATs in real-world settings.

5. Conclusion

This study reveals the complex and context-dependent nature of trust and distrust spreading in HATs. Their presence does not inherently determine the effectiveness of trust and distrust spreading but rather by how they align with the team's structural composition and the sequence of behavioral interactions. Matching communicative and behavioral trust amplified positive outcomes, particularly in human-majority teams, while mismatched spreading occasionally produced unexpected benefits, such as increased adaptability and resilience in changing circumstances.

The findings underscore the importance of designing HATs in which AI agents can dynamically adjust their communication and behavior based on team composition and operational context. For example, AI agents that proactively engage in trust-building behaviors or effectively counteract distrust spreading could significantly improve team dynamics. Organizations can leverage this knowledge to train human teammates on effective trust management and develop protocols that align communicative and behavioral trust strategies, enhancing team cohesion and efficiency.

By showcasing how the interplay of communication and behavior influences trust within HATs, the study contributes to a deeper understanding of effective team coordination in the presence of AI teammates. Future research integrating advanced AI communication systems and exploring long-term trust dynamics will further refine our understanding of HATs, which may pave the way for more effective integration of AI agents in diverse operational environments. This advanced understanding of trust in HATs is essential for designing the next generation of human-AI teaming systems capable of navigating the intricate landscape of trust dynamics to achieve superior collaborative outcomes in increasingly complex operational settings.

CRediT authorship contribution statement

Shiwen Zhou: Writing – review & editing, Writing – original draft, Visualization, Methodology, Investigation, Formal analysis, Conceptualization. **Wen Duan:** Writing – review & editing, Writing – original draft, Supervision, Project administration, Methodology. **Xiaoyun Yin:** Writing – review & editing, Writing – original draft, Investigation. **Matthew Scalia:** Writing – review & editing, Visualization, Methodology, Investigation. **Ray Hao:** Writing – review & editing, Writing – original draft, Investigation. **Nan Weng:** Writing – original draft. **Gregory Funke:** Writing – review & editing, Methodology. **Michael Tolston:** Writing – review & editing, Methodology. **Guo Freeman:** Writing – review & editing. **Beau Schelble:** Writing – review & editing, Methodology. **Jamie Gorman:** Writing – review & editing, Methodology, Formal analysis, Conceptualization. **Nathan McNeese:** Writing – review & editing, Supervision, Methodology, Conceptualization.

Declaration of interest statement

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

This research was supported by Air Force Office of Scientific Research Award No. FA9550-21-1-0314 (Program Manager: Laura Steckman).

References

Alarcon, G.M., Capiola, A., Hamdan, I.A., Lee, M.A., Jessup, S.A., 2023. Differential biases in human-human versus human-robot interactions. Appl. Ergon. 106, 103858.
Ball, J., Myers, C., Heiberg, A., Cooke, N.J., Matessa, M., Freiman, M., Rodgers, S., 2010. The synthetic teammate project. Comput. Math. Organ. Theor. 16, 271–299.
Chiou, E.K., Lee, J.D., 2023. Trusting automation: designing for responsivity and resilience. Hum. Factors 65 (1), 137–165.
Cooke, N.J., Gorman, J.C., Duran, J.L., Taylor, A.R., 2007. Team cognition in experienced command-and-control teams. J. Exp. Psychol. Appl. 13 (3), 146.
Cooke, N.J., Shope, S.M., 2004. Synthetic task environments for teams: CERTT's UAV-STE. In: Handbook of Human Factors and Ergonomics Methods. CRC Press, pp. 476–483.
Costa, A.C., Roe, R.A., Taillieu, T., 2001. Trust within teams: the relation with performance effectiveness. Eur. J. Work. Organ. Psychol. 10 (3), 225–244.
De Jong, B.A., Dirks, K.T., Gillespie, N., 2016. Trust and team performance: a meta-analysis of main effects, moderators, and covariates. J. Appl. Psychol. 101 (8), 1134.
Demir, M., McNeese, N.J., Cooke, N.J., 2016. Team communication behaviors of the human-automation teaming. In: 2016 IEEE International Multi-Disciplinary Conference on Cognitive Methods in Situation Awareness and Decision Support (CogSIMA). IEEE, pp. 28–34.
Demir, M., McNeese, N.J., Gorman, J.C., Cooke, N.J., Myers, C.W., Grimm, D.A., 2021. Exploration of teammate trust and interaction dynamics in human-autonomy teaming. IEEE Trans. Human Machine Syst. 51 (6), 696–705.
Dirks, K.T., de Jong, B., 2022. Trust within the workplace: a review of two waves of research and a glimpse of the third. Ann. Rev. Organ. Psychol. Organ. Behav. 9 (1), 247–276.
Duan, W., Flathmann, C., Scalia, M., Freeman, G., Gorman, J., McNeese, N.J., Zhang, R., Zhou, S., Hauptman, A., Yin, X., 2025, April. Trusting Autonomous Teammates in Human-AI Teams-A Literature Review. In: Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems, pp. 1–23.
Duan, W., Zhou, S., Scalia, M., Freeman, G., Gorman, J., Tolston, M., Funke, G., McNeese, N.J., 2025b. Understanding the processes of trust and distrust contagion in human–AI teams: A qualitative approach. Computers in Human Behavior 165, 108560.
Duan, W., Zhou, S., Scalia, M.J., Yin, X., Weng, N., Zhang, R., et al., 2024. Understanding the evolvement of trust over time within Human-AI teams. Proc. ACM Human-Comput. Interact. 8 (CSCW2), 1–31.
Entin, E.E., Serfaty, D., 1999. Adaptive team coordination. Hum. Factors 41 (2), 312–325.
Fulmer, C.A., Gelfand, M.J., 2012. At what level (and in whom) we trust: trust across multiple organizational levels. J. Manag. 38 (4), 1167–1230.
Glikson, E., Woolley, A.W., 2020. Human trust in artificial intelligence: review of empirical research. Acad. Manag. Ann. 14 (2), 627–660.
Gorman, J.C., Amazeen, P.G., Cooke, N.J., 2010a. Team coordination dynamics. Nonlinear Dynam. Psychol. Life Sci. 14 (3), 265.
Gorman, J.C., Cooke, N.J., 2011. Changes in team cognition after a retention interval: the benefits of mixing it up. J. Exp. Psychol. Appl. 17 (4), 303.

Gorman, J.C., Cooke, N.J., Amazeen, P.G., 2010b. Training adaptive teams. Hum. Factors 52 (2), 295–307.
Gorman, J.C., Wiltshire, T.J., 2024. A typology for the application of team coordination dynamics across increasing levels of dynamic complexity. Hum. Factors 66 (1), 5–16.
Grosser, T., Kidwell, V., Labianca, G.J., 2012. Hearing it through the grapevine: positive and negative workplace gossip. Organ. Dyn. 41, 52–61.
Hafizoglu, F.M., Sen, S., 2018. Reputation based trust in human-agent teamwork without explicit coordination. In: Proceedings of the 6th International Conference on Human-Agent Interaction, pp. 238–245.
Harrison, J.L., Zhou, S., Scalia, M.J., Grimm, D.A., Demir, M., McNeese, N.J., et al., 2024. Communication strategies in human-autonomy teams during technological failures. Hum. Factors, 00187208231222119.
Huang, L., Cooke, N.J., Gutzwiller, R.S., Berman, S., Chiou, E.K., Demir, M., Zhang, W., 2020. Distributed dynamic team trust in human, artificial intelligence, and robot teaming. Trust Human-Robot Interact. 301–309.
Huang, H.Y., Bashir, M.N., 2018. Users' trust in automation: a cultural perspective. Advances in Human Factors in Robots and Unmanned Systems, pp. 282–289.
Jarvenpaa, S.L., Knoll, K., Leidner, D.E., 1998. Is anybody out there? Antecedents of trust in global virtual teams. J. Manag. Inf. Syst. 14 (4), 29–64.
Kelley, J.F., 1983. An empirical methodology for writing user-friendly natural language computer applications. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. ACM, pp. 193–196.
Lee, J.D., See, K.A., 2004. Trust in automation: designing for appropriate reliance. Hum. Factors 46 (1), 50–80.
Lewicki, R.J., Barry, B., Saunders, D.M., 2020. Negotiation. McGraw-Hill.
Lin, C.T., Fan, H.Y., Chang, Y.C., Ou, L., Liu, J., Wang, Y.K., Jung, T.P., 2022. Modelling the trust value for human agents based on real-time human states in human-autonomous teaming systems. Technologies 10 (6), 115.
Lyons, J.B., Sadler, G.G., Koltai, K., Battiste, H., Ho, N.T., Hoffmann, L.C., et al., 2017. Shaping trust through transparent design: theoretical and experimental guidelines. In: Advances in Human Factors in Robots and Unmanned Systems: Proceedings of the AHFE 2016 International Conference on Human Factors in Robots and Unmanned Systems, July 27-31, 2016, Walt Disney World®. Springer International Publishing, Florida, USA, pp. 127–136.
Lyons, J.B., Sycara, K., Lewis, M., Capiola, A., 2021. Human–autonomy teaming: definitions, debates, and directions. Front. Psychol. 12, 589585.
Myers, C., Ball, J., Cooke, N., Freiman, M., Caisse, M., Rodgers, S., et al., 2018. Autonomous intelligent agents for team training. IEEE Intell. Syst. 34 (2), 3–14.
McNeese, N., Demir, M., Chiou, E., Cooke, N., Yanikian, G., 2019. Understanding the Role of Trust in Human-Autonomy Teaming.
McNeese, N.J., Demir, M., Chiou, E.K., Cooke, N.J., 2021. Trust and team performance in human–autonomy teaming. Int. J. Electron. Commer. 25 (1), 51–72.
McNeese, N.J., Demir, M., Cooke, N.J., Myers, C., 2018. Teaming with a synthetic teammate: insights into human-autonomy teaming. Hum. Factors 60 (2), 262–273.
O'Neill, T., McNeese, N., Barron, A., Schelble, B., 2022. Human–autonomy teaming: a review and analysis of the empirical literature. Hum. Factors 64 (5), 904–938.
Salas, E., Dickinson, T.L., Converse, S.A., Tannenbaum, S.I., 1992. Toward an understanding of team performance and training. In: Swezey, R.W., Salas, E. (Eds.), Teams: Their Training and Performance. Ablex Publishing, Norwood, NJ, pp. 3–29.
Schelble, B.G., Flathmann, C., Scalia, M., Zhou, S., Myers, C., McNeese, N.J., et al., 2022. Addressing the spread of trust and distrust in distributed human-AI teaming constellations. In: CHI TRAIT Workshop, 2022.
Schelble, B.G., Lopez, J., Textor, C., Zhang, R., McNeese, N.J., Pak, R., Freeman, G., 2024. Towards ethical AI: empirically investigating dimensions of AI ethics, trust repair, and performance in human-AI teaming. Hum. Factors 66 (4), 1037–1055.
Sheridan, T.B., 1988. Task allocation and supervisory control. In: Handbook of Human-Computer Interaction. North-Holland, pp. 159–173.
Spoelma, T.M., Hetrick, A.L., 2021. More than idle talk: examining the effects of positive and negative team gossip. J. Organ. Behav. 42 (5), 604–618.
Stevens, R.H., Galloway, T.L., 2019. Teaching machines to recognize neurodynamic correlates of team and team member uncertainty. J. Cogn. Eng. Decis. Mak. 13 (4), 310–327.
Textor, C., Zhang, R., Lopez, J., Schelble, B.G., McNeese, N.J., Freeman, G., et al., 2022. Exploring the relationship between ethics and trust in human–artificial intelligence teaming: a mixed methods approach. J. Cogn. Eng. Decis. Mak. 16 (4), 252–281.
Van de Bunt, G.G., Wittek, R.P., de Klepper, M.C., 2005. The evolution of intra-organizational trust networks: the case of a German paper factory: an empirical test of six trust mechanisms. Int. Sociol. 20 (3), 339–369.
Verma, J.P., 2015. Repeated Measures Design for Empirical Researchers. John Wiley & Sons.
Woods, D.D., 2015. Four concepts for resilience and the implications for the future of resilience engineering. Reliab. Eng. Syst. Saf. 141, 5–9.
Zhang, R., Duan, W., Flathmann, C., McNeese, N., Freeman, G., Williams, A., 2023. Investigating AI teammate communication strategies and their impact in human-AI teams for effective teamwork. Proc. ACM Human-Comput. Interact. 7 (CSCW2), 1–31.
Zhou, S., Yin, X., Scalia, M.J., Zhang, R., Gorman, J.C., McNeese, N.J., 2023. Development of a real-time trust/distrust metric using interactive hybrid cognitive task analysis. In: Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 67. SAGE Publications, Sage CA: Los Angeles, CA, pp. 2128–2136, 1.