Check for updates

*Research Article*

# Trust Contagion and Transfer Are Not Automatic in Multi-Human Multi-Unmanned Aerial Vehicle Teams

Human Factors
2026, Vol. 0(0) 1–21
© 2026 Human Factors and Ergonomics Society

Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/00187208261506615
journals.sagepub.com/home/hfs


**Debbie Yutung Hsu[1], Theresa Kessler[2], Elizabeth Weldon[2], and Mengyao Li[1]**

## Abstract

**Objective:** This study examined trust contagion and trust transfer in co-located multi-human multi-Unmanned Aerial Vehicle (UAV) teams and adopted the Actor-Partner Interdependence Model (APIM) to model trust dynamics in human-autonomy teams (HAT).

**Background:** Most HAT research has focused on dyadic relationships, yet real-world operations increasingly involve multiple human operators collaborating with multiple agents. How trust spreads between human teammates and transfers across agents in non-dyadic contexts remains poorly understood.

**Method:** A 3 (UAV reliability: High-High, Mixed, Low-Low; within-subjects) × 2 (Priming: Trust vs. Distrust; between-subjects) mixed design was used. Fifty-three dyads ($N$ = 106) performed a target acquisition task in two-human, two-UAV teams. Subjective trust was analyzed using linear mixed models as the primary analysis, with APIM as a secondary exploratory analysis of dyadic influence.

**Results:** For trust contagion, UAV reliability was the primary driver of trust, with no evidence that primed attitudes spread between operators, suggesting boundary conditions on when contagion occurs. For trust transfer, trust in the reliable UAV did not decline when paired with an unreliable one, supporting component-specific trust. Exploratory APIM helped disentangle dyadic interdependence and generate hypotheses for future work.

**Conclusion:** Trust in multi-human multi-agent teams was driven primarily by system reliability, with operators adopting component-specific trust, calibrating trust to each UAV individually.

**Application:** Supporting calibrated trust depends first on system reliability and on direct experience with each agent's performance. APIM offers a means to distinguish an operator's own trust development from a teammate's influence, making it suited to studying trust in multi-human multi-agent teams.

[1]Psychological and Brain Sciences, Georgia Institute of Technology, Atlanta, GA, USA
[2]Georgia Tech Research Institute, Atlanta, GA, USA

**Corresponding Author:**
Mengyao Li, Psychological and Brain Sciences, Georgia Institute of Technology, 654 Cherry Street, Atlanta, Georgia 30332-0170, USA.
Email: mengyao.li@gatech.edu

**Keywords**
trust in autonomy, team dynamics, interpersonal influence, trust transfer, actor-partner interdependence model

Précis: In co-located multi-human multi-UAV teams, trust was driven primarily by each UAV's reliability, with operators maintaining component-specific rather than system-wide trust. An exploratory dyadic analysis using the Actor-Partner Interdependence Model helped generate hypotheses about interpersonal influence for future work.

## Introduction

Autonomous systems are increasingly deployed in safety-critical environments where multiple human operators manage multiple agents simultaneously. In domains such as military surveillance, effective human-autonomy teams (HAT) depend on well-calibrated trust. Miscalibrated trust that is either excessive or insufficient leads to automation misuse, disuse, or complacency, with direct consequences for performance and safety (Lee & See, 2004; Parasuraman & Riley, 1997). Most HAT trust research has examined dyadic configurations involving a single human and a single agent. Yet, HATs increasingly consist of multiple human operators collaborating with multiple agents, introducing trust dynamics that dyadic models cannot fully capture. In these settings, one operator's miscalibrated trust may spread to teammates and shift how they rely on the same agent, potentially scaling into a team-level failure to catch or correct agent mistakes. Understanding how trust propagates in non-dyadic teams is therefore essential for designing teams, training, and agents that keep reliance calibrated.

Two propagation mechanisms are particularly relevant: *trust contagion*, the spread of trust between human teammates, and *trust transfer*, the generalization of trust from one agent to another within the same operator. Both have typically been studied in isolation, and rarely within a single multi-human multi-agent team. To address these gaps, the present study examined trust contagion and transfer within a two-human, two-Unmanned Aerial Vehicle (UAV) team. We used trust priming to manipulate one operator's trust through priming videos depicting varied UAV reliability. We complemented linear mixed model analysis with an exploratory application of the Actor-Partner Interdependence Model (APIM) (Cook & Kenny, 2005; Kenny, 1996), which models interdependence between teammates directly. Overall, we found no evidence of trust contagion, and trust in each UAV remained specific to that UAV's reliability rather than generalizing across the pair. These processes appeared to depend on features of the operators, teams, and tasks—that is, they did not occur automatically in multi-human multi-UAV teams. The exploratory APIM further helped disentangle the dyadic structure of our data and supported hypothesis generation for future work.

## Background

### *Trust in Multi-Human Multi-Agent Teams*

Multi-human multi-agent teaming is an increasingly common but underexplored configuration in HAT research (Arnold & Scheutz, 2017; Schneiders et al., 2022). Trust, defined as "the attitude that an agent will help achieve an individual's goals in a situation characterized by uncertainty and vulnerability," is a central determinant of whether operators appropriately rely on the agents (Lee & See, 2004). Whereas most prior work has examined dyadic human-agent pairs, non-dyadic teams introduce qualitatively distinct dynamics. The cognitive properties of a group differ from those of the individuals who comprise it, due to factors such as the division of cognitive labor and group-formed consensus (Hutchins, 1995). As teaming

between groups of humans and agents becomes more common (Fraune et al., 2015, 2017), understanding how trust operates across both human-human and human-agent relationships becomes essential. Huang et al. (2021) proposed the distributed dynamic team trust (D2T2) framework, suggesting team trust is dynamic and transitive, influenced through conversations and training procedures. Recent work in a two-human, two-agent setting suggested that the two operators' trust could converge or diverge, and greater agreement in their trust ratings accompanied more convergent communication and better team performance (Chung & Yang, 2026a). Yet how trust contagion between teammates and trust transfer between agents operate together within a team remains underexplored.

## Trust Contagion Between Humans

Trust contagion is a type of interpersonal influence in which trust attitudes spread between teammates through social interaction, aligning attitudes across the group. Related constructs include trust transitivity (Huang et al., 2021), trust spread (Ramchurn et al., 2004; Schelble et al., 2022), and trust propagation (Guo et al., 2023).

Because trust operates through both affective and analytic processes (Lee & See, 2004), two mechanisms may explain trust contagion. Emotional contagion transmits affective states automatically through behavioral synchrony (Hatfield et al., 1993). More recent accounts question this automaticity, arguing that such transmission is moderated by social context (Wróbel & Imbir, 2019). Informational social influence occurs when an operator updates their trust judgment by treating a teammate's assessment as credible evidence (Deutsch & Gerard, 1955).

Rojas et al. (2025) provided the first empirical evidence of trust contagion in a co-located human-human-agent team, showing that individuals paired with a highly agent-trusting confederate expressed significantly greater trust in and reliance on the agent than those paired with a distrusting confederate. Such propagation occurs even among ad hoc teams of strangers (Hidalgo et al., 2025) and can cross between interconnected teams (Yin et al., 2025). Duan et al. (2026) showed that (dis)trust spread through direct interaction within a team but traveled by word-of-mouth between teams. This spread is not always symmetric. Duan et al. (2025) found distrust spread more easily than trust within a HAT, consistent with evidence that trust and distrust are distinct constructs with different brain activation (Dimoka, 2010).

## Trust Transfer Between Agents

Trust transfer refers to how trust in one agent shapes trust in another within the same operator. Keller and Rice (2009) proposed two competing theories. Component-Specific Trust (CST) suggests that operators maintain separate trust evaluations for each agent based on its individual performance and characteristics. In contrast, System-Wide Trust (SWT) suggests that operators develop holistic trust attitudes that generalize across related agents even when they function independently. Early empirical evidence largely supported negative SWT: when one agent's reliability declined in a multi-agent context, performance and trust declined across both reliable and unreliable agents (Geels-Blair et al., 2013; Keller & Rice, 2009; Walliser et al., 2016). However, more recent work suggested CST may dominate when operators have high task familiarity, decisional freedom, or lower workload (Lopez & Pak, 2020; Lyons et al., 2024). Chung & Yang, 2026b further showed that these patterns vary across individuals, with some operators generalizing trust across agents and others evaluating each separately. The relative dominance of SWT versus CST in multi-human multi-agent team contexts remains underexplored. Whether and how trust contagion modulates trust transfer has yet to be examined.

## The Present Study

The present study addressed three gaps. First, we manipulated trust through a pre-task performance demonstration and let it

spread through unscripted communication, better reflecting how attitudes form and spread in operational teams. Prior contagion research manipulated teammates' attitudes through confederates and scripted interactions, which afford experimental control but have lower ecological validity. Real operational HATs also involve trust attitudes shaped by briefings, demonstrations, and training before teams interact. Our study addressed this by priming one operator's trust through a pre-task UAV performance demonstration, then letting the team communicate naturally to examine whether the primed attitude propagates to a teammate. Trust is susceptible to such priming: training experiences shape operators' expectations and mental models of automation (Hoff & Bashir, 2015; Körber et al., 2018; Lee & See, 2004). Second, we isolated each teammate's influence on the other using APIM, modeling the operator pair as the unit of analysis (Cook & Kenny, 2005; Kenny, 1996). Contagion is inherently dyadic: each human operator both influences and is influenced by their partner. Linear mixed models accommodate this nonindependence through team-level random intercepts, but they estimate each operator's trust from their own predictors and do not isolate the influence of a partner's attitudes. APIM decomposes each operator's outcome into actor effects (how one's own characteristics predict one's own outcome) and partner effects (how a partner's characteristics predict one's outcome), operationalizing trust contagion by estimating the simultaneous, bidirectional influence between teammates. Third, we examined trust contagion and trust transfer within a single team. These processes have been studied separately, leaving open whether and how they interact.

Overall, we hypothesized:

**H1: Trust Priming**: Participants exposed to a high-performing UAV priming demonstration will report significantly higher post-priming trust than those exposed to a low-performing UAV demonstration.

**H2a: Pre-Task Trust Contagion**: Before task experience, participants exposed to a trust-primed teammate will report greater trust in UAVs than those exposed to a distrust-primed teammate.

**H2b: Post-Task Trust Contagion**: After task experience, the effect of a teammate's primed trust attitude on UAV trust will be moderated by UAV reliability.

**H3: Post-Task Trust Transfer**: When two UAVs show different reliability levels, the less reliable agent will reduce trust in the co-present more reliable agent, consistent with negative SWT.

## Method

This study employed a 3 (reliability: High-High, Mixed, Low-Low; within-subjects) × 2 (priming: Trust, Distrust; between-subjects) mixed design. Each team comprised two human operators (Commander, Intelligence Support) and two UAVs (North, South). Both UAVs performed the same function, differing only in reliability, consistent with prior trust transfer research that compares multiple aids serving the same function (Keller & Rice, 2009; Walliser et al., 2016). UAV reliability was manipulated at two levels: high (~95% accuracy) and low (~60% accuracy), consistent with established reliability thresholds (Chavaillaz et al., 2016; see Table 1). The reliability conditions were counterbalanced using a Latin square design. This research complied with the American Psychological Association Code of Ethics and was approved by the Institutional Review Board at Georgia Institute of Technology. Informed consent was obtained from each participant.

Participants were informed that they would work with two UAVs, each equipped with distinct algorithms, throughout the study. Details about the reliability of each UAV were not disclosed. The roles of Commander and Intelligence Support differed in two ways. First, only the Intelligence Support received the trust priming manipulation: a 30-second

**Table 1.** UAV Reliability and Detection Count

| | UAV | | | |
|---|---|---|---|---|
| | North | | South | |
| Reliability | Accuracy | Errors | Accuracy | Errors |
| High-High | 96% (27/28) | 1 | 94% (50/53) | 3 |
| Mixed | 95% (61/64) | 3 | 58% (25/43) | 18 |
| Low-Low | 58% (23/40) | 17 | 63% (31/49) | 18 |

*Note.* Accuracy is the proportion of correct UAV identifications out of total vehicles presented (shown in parentheses). Errors are the number of vehicles incorrectly identified or missed.

priming video of a UAV performing the experimental task at either 100% (Trust condition) or 57% (Distrust condition) reliability. Second, the communication instructions differed: both the Commander and the Intelligence Support were instructed to verbally express their experiences when interacting with the UAVs, but the Intelligence Support was directed to share first during team discussions. The Commander was additionally instructed to observe the Intelligence Support's UAV interactions to inform their own UAV assessments. No other role differences were imposed.

## Testbed: Reconnaissance, Surveillance, and Target Acquisition (RSTA)

The RSTA task required human operators working with two UAVs that autonomously scanned adjacent geographic areas to detect, locate, and identify a specified target (e.g., white van). As shown in Figure 1, each UAV categorized objects by marking potential targets with green frames and non-targets with white frames while displaying detection statistics (Total and Target counts). Participants needed to monitor both UAVs' identifications and veto incorrect identifications using keyboard inputs ("D" for North and "J" for South) to contribute to algorithm refinement. UAVs could commit three types of errors: false positives, false negatives, and misses. The tasks varied in landscape and traffic (e.g., a roundabout or a straight road with differing numbers of vehicles) to approximate the varied conditions of real reconnaissance. Because scenarios were generated to be realistic rather than matched on event count, the number of vehicles differed across tasks and UAVs (Table 1). Reliability was therefore defined as the proportion of correct identifications.

## Participants

A priori power analysis ($\alpha = .05$, power = .80, effect size = 0.25) determined a minimum sample size of $N = 44$ dyads. We recruited 60 dyads. Seven dyads were excluded due to missing data, yielding a final sample of 53 dyads ($N = 106$) randomly assigned to Trust ($n = 27$) or Distrust ($n = 26$) conditions. Of the sample, 37% identified as female, with ages ranging from 18 to 64 years ($M = 19.73$, $SD = 4.73$).

## Measures

*One-Item Trust.* Before the RSTA task, trust was assessed using a one-item 7-point scale: "Please rate how much you trust your UAV teammate in identifying the target on a scale of 1 (*Not at all*) to 7 (*Very*)." The item assessed trust in the UAVs collectively, as no UAV-specific differentiation was expected before the task. A one-item scale was used to minimize administration time and to capture trust prior to direct task experience.

*Multi-Dimensional Measure of Trust (MDMT).* Trust in each UAV was rated after each RSTA task using an adapted MDMT (Ullman & Malle, 2019) on an 8-point scale from 0 (*Not at all*)

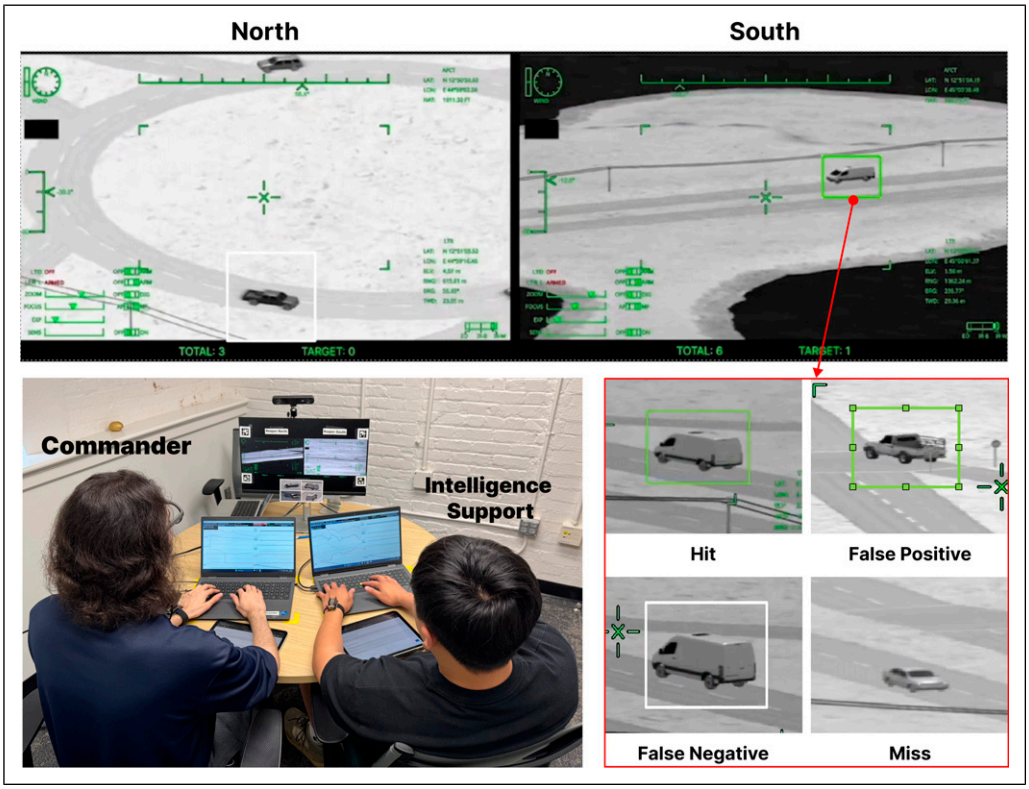


**Figure 1.** RSTA stimuli and experimental setup.
*Note.* Both operators were seated side by side and viewed a single shared display, with UAV North's video feed on the left and UAV South's on the right. Green frames mark vehicles each UAV classified as targets; white frames mark non-targets

to 7 (*Very*), with an additional *Does not Fit* option to prevent forced responses.

*Workload.* The NASA Task Load Index (NASA-TLX) (Hart & Staveland, 1988) was administered after each task on a scale of 0–20 to verify that differences in task scenarios did not impose differential task demands.

*Dispositional Measures.* Four measures were administered at the start of the study to explore whether dispositional factors moderated trust outcomes: a three-item individualism-collectivism (IC) scale on a 5-point Likert scale (Wagner, 1995), a six-item propensity to trust scale on a 5-point Likert scale (Jessup et al., 2019), the Big Five Inventory-10 (BFI-10) questionnaire (Rammstedt & John, 2007), and the Negative Attitude Toward Robots Scale (NARS) (Bartneck et al., 2005; Chen & Barnes, 2013; Nomura et al., 2006) (see Appendix A). Behavioral, conversational, and physiological data were also collected and will be analyzed in a future paper.

## Procedures

Participants first signed consent forms, completed demographic surveys, and were randomly assigned roles. They were then brought to separate rooms for training, where each watched a 4-minute introductory video explaining the task and each participant's assigned role responsibility: "Intelligence Support should be vocal about your experiences while the Commander should closely observe the

interaction." Task comprehension was verified through a mandatory quiz requiring 100% accuracy, after which participants rated their baseline trust in the UAVs using the one-item 7-point scale.

Intelligence Support then received the trust priming manipulation. In the *Trust* condition, they were told the UAV algorithm was highly accurate and watched a 100% accuracy video. In the *Distrust* condition, they were told the algorithm occasionally struggled with target identification and watched the 57% accuracy video. Intelligence Support then rated their post-priming trust using the one-item trust scale as a manipulation check.

Both participants were then brought together for a pre-task discussion: "Before the mission begins, please discuss how well you think the UAVs will perform in the mission." With the Intelligence Support directed to share first, the discussion proceeded without a time limit and typically took under three minutes, after which each operator independently rated post-discussion trust using the one-item scale. They then completed three blocks. Each block consisted of an 8-minute RSTA task followed by the MDMT for each UAV (rated North, then South) and the NASA-TLX. Blocks 1 and 2 were each followed by a between-block discussion in the same format as the pre-task discussion, except that no one-item trust rating was collected afterward. Participants were then debriefed and compensated. The study lasted approximately 90 minutes (Figure 2).

## Inferential Analysis

Data were exported from Qualtrics and analyzed in R using packages *lme4* and *emmeans* (Bates et al., 2015; Lenth & Piaskowski, 2026). The measures entering each analysis are summarized in Table 2. Primary analyses used linear mixed models with Bonferroni-corrected pairwise comparisons, accommodating individual- and team-level variation through random effects to address the non-independence inherent in repeated measures (Muhammad, 2023). To identify individual differences, we conducted exploratory backward elimination using *lmerTest* (Kuznetsova et al., 2017), which sequentially removed fixed effects based on *F*-tests with Satterthwaite degrees of freedom. Primary predictors were protected from elimination.

*APIM.* We followed the two-intercept approach (Cook & Kenny, 2005), omitting the overall intercept and including an indicator for each role, so that intercepts, actor effects, partner effects, and UAV effects were estimated separately for the Commander and the Intelligence Support without treating either as a reference category. Models were fitted by generalized least squares via the *nlme* package (Pinheiro et al., 2026). We fitted a separate model for each of the six priming condition × reliability combinations, with post-discussion one-item trust as the predictor and post-task MDMT as the outcome. Post-discussion trust was used as the predictor to focus on the post-task period. This aligns the APIM with the post-task linear mixed model used to test trust contagion and trust transfer so that differences between the analyses reflect the modeling of interdependence between operators rather than a difference in measurement.

Within each model, every team contributed four MDMT ratings: the Commander's rating of each UAV and the Intelligence Support's rating of each UAV (two roles × two UAVs). These four observations are correlated in three ways. Two ratings from the same operator share individual-level variation specific to that operator. Two ratings of the same UAV by different operators share that UAV's observed performance and the preceding discussion of it. Two ratings by different operators of different UAVs share only the team context. Each source, however, includes two distinct pairs of ratings. For example, the same-operator source comprises the Commander's pair (North, South) and the Intelligence Support's pair. Because the two operators were different individuals and the two UAVs differed in reliability, the two pairs within a source need not correlate to the same degree. We therefore estimated all six pairwise correlations freely, using an unstructured

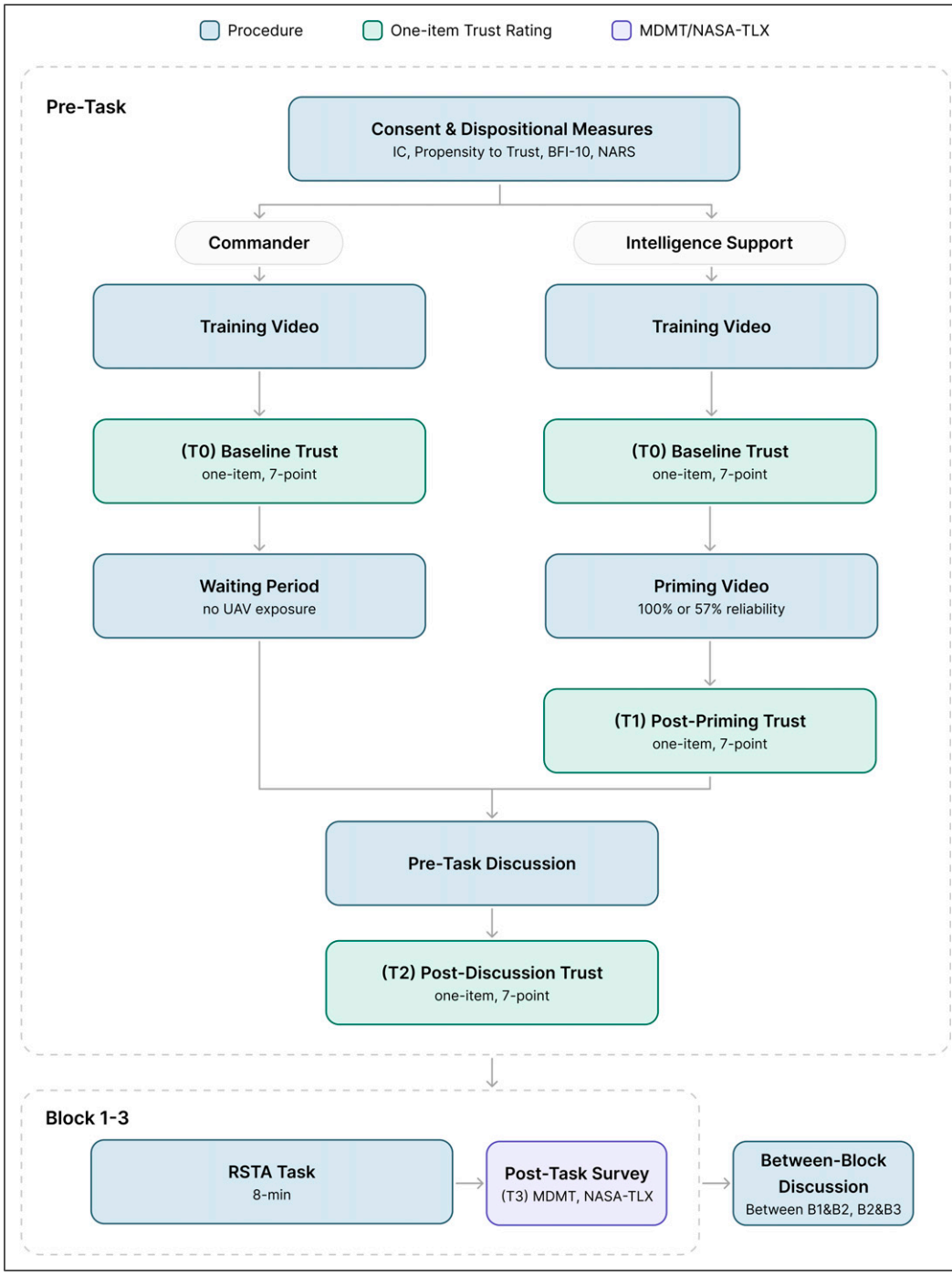


**Figure 2.** Experimental procedure

**Table 2.** Corresponding Measures and Hypotheses for Each Analysis

| Analysis | Model | Operator | Data |
|---|---|---|---|
| H1: Manipulation Check | LMM | Intelligence Support | T0 vs. T1 |
| H2a: Pre-Task Trust Contagion | LMM | Commander | T0 vs. T2 |
| | | Intelligence Support | T1 vs. T2 |
| H2b & H3: Post-Task Trust Contagion & Transfer | LMM | Both operators | T3 |
| Exploratory APIM | GLS (two-intercept) | Both operators | T2 → T3 (per priming × reliability cell) |

*Note.* LMM = linear mixed model; GLS = generalized least squares; T0 = baseline trust; T1 = post-priming trust; T2 = post-discussion trust; T3 = post-task trust (MDMT) after each of the three blocks.

correlation rather than assuming a common value. Residual variances were allowed to differ between the two roles, consistent with the two-intercept specification, which estimates separate effects for each role. Trust ratings were standardized across the full sample. Bonferroni correction was applied across the six models, and six effects were estimated within each model (see Figure 3).

## Results

### Manipulation Check

To verify the trust manipulation (H1), we fit a linear mixed model on the Intelligence Support's one-item trust ratings to assess the effect of *Priming Condition* and *Timepoint* (baseline vs. post-priming). The analysis revealed a significant *Priming Condition* × *Timepoint* interaction ($\beta = -2.27$, $t(51) = -6.85$, $p < .001$, $\omega^2 = .46$). As shown in Figure 4, trust increased in the Trust condition (baseline: $M = 4.87$, $SE = 0.21$; post: $M = 5.95$, $SE = 0.21$), $p_{adj} < .001$, $d = -1.25$, and decreased in the Distrust condition (baseline: $M = 4.71$, $SE = 0.21$; post: $M = 3.52$, $SE = 0.21$), $p_{adj} < .001$, $d = 1.36$. The two priming conditions did not differ at baseline but diverged significantly post-priming, $p_{adj} < .001$, $d = 1.70$, confirming a successful trust manipulation and supporting H1.

### Pre-Task Trust Contagion

We next tested whether the Intelligence Support's primed attitude propagated to the Commander before either operator had task experience (H2a). We first examined the Commander's baseline trust, which was not subject to priming. Commanders in the Trust condition entered with significantly higher baseline trust than Commanders in the Distrust condition (Trust: $M = 5.78$, $SE = 0.22$; Distrust: $M = 4.96$, $SE = 0.22$), $p_{adj} < .01$, $d = 0.43$. To explore this difference, we compared Commanders' propensity to trust and NARS across priming conditions using independent-samples *t*-tests, and neither differed ($p > .05$). The gap could not be attributed to these measured dispositional differences and most likely reflects chance variation in random assignment. Because Commander baseline trust was not equivalent across priming conditions, we interpret change within each priming condition over time rather than compare absolute trust across priming conditions.

We fit a linear mixed model on one-item trust with *Role*, *Priming Condition*, and *Timepoint* as predictors. *Timepoint* contrasted each operator's last rating before discussion (pre-discussion rating: baseline trust for the Commander; post-priming trust for the Intelligence Support) with their post-discussion rating. The model revealed a significant three-way interaction ($\beta = 0.94$, $t(102) = 2.38$, $p = .02$, $\omega^2 = .04$), a significant *Role* × *Priming Condition* interaction ($\beta = -1.50$, $t(151.73) = -3.43$, $p < .001$, $\omega^2 = .05$), and a significant effect of *Priming Condition* ($\beta = -0.82$, $t(151.73) = -2.64$, $p = .01$, $\omega^2 = .28$).

As shown in Figure 5, post-hoc comparisons of the three-way interaction showed that in the Trust condition, neither Commander

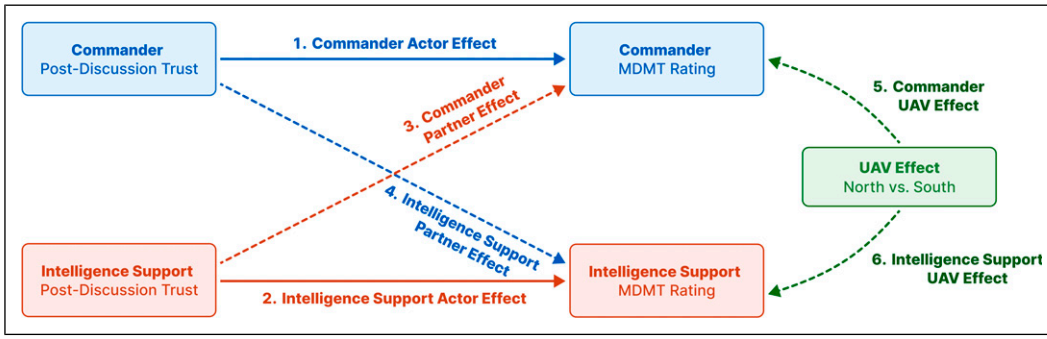


**Figure 3.** Effects estimated by APIM.
*Note.* Six estimated effects in the APIM. Actor effects (solid arrows) capture how each operator's post-discussion trust predicts their own MDMT rating. Partner effects (dashed diagonal arrows) capture cross-operator influence: Commander post-discussion trust predicting Intelligence Support MDMT and vice versa. UAV effects (dashed curved arrows) capture the mean MDMT difference of South relative to North for each operator

(pre: $M = 5.78$, $SE = 0.22$; post: $M = 6.15$, $SE = 0.22$) nor Intelligence Support (pre: $M = 5.89$, $SE = 0.22$; post: $M = 5.78$, $SE = 0.22$) changed significantly across discussion ($p_{adj} > .05$). In the Distrust condition, both roles increased: Commander (pre: $M = 4.96$, $SE = 0.22$; post: $M = 5.46$, $SE = 0.22$, $p_{adj} = .01$, $d = -0.49$) and Intelligence Support (pre: $M = 3.58$, $SE = 0.22$; post: $M = 4.54$, $SE = 0.22$, $p_{adj} < .001$, $d = -0.95$). Rather than the primed distrust attitude spreading downward to the unprimed Commander, both operators' trust rose. H2a was not supported.

Post-hoc comparisons of the *Role* × *Priming Condition* interaction indicated that in the Trust condition, Commander ($M = 5.96$, $SE = 0.19$) and Intelligence Support ($M = 5.83$, $SE = 0.19$) did not differ ($p_{adj} > .05$). In the Distrust condition, Commander ($M = 5.21$, $SE = 0.20$) reported significantly higher trust than Intelligence Support ($M = 4.06$, $SE = 0.20$), $p_{adj} < .001$, $d = 0.82$.

## Post-Task Trust in UAVs

To examine post-task trust contagion and trust transfer, we fit a linear mixed model on

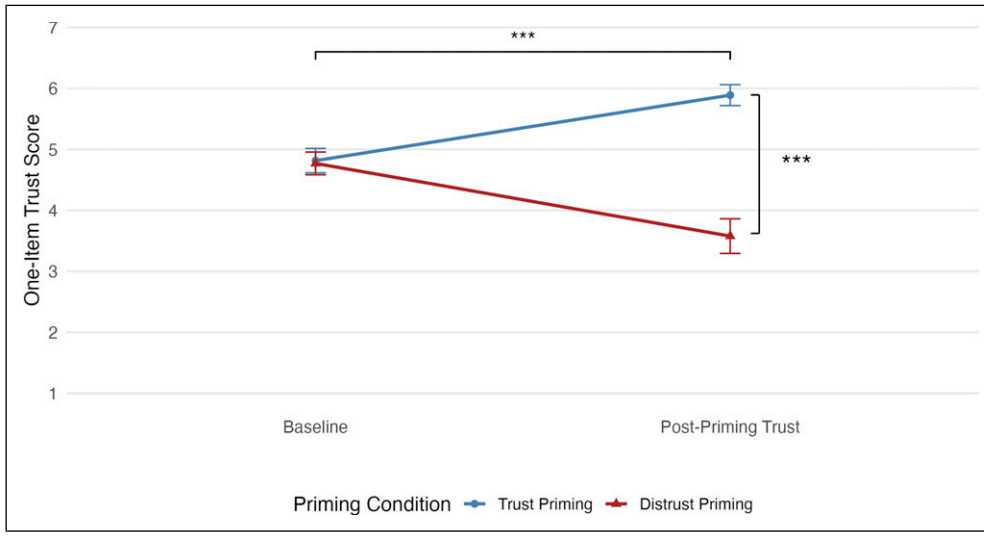


**Figure 4.** Intelligence Support one-item trust score before vs. after trust priming manipulation

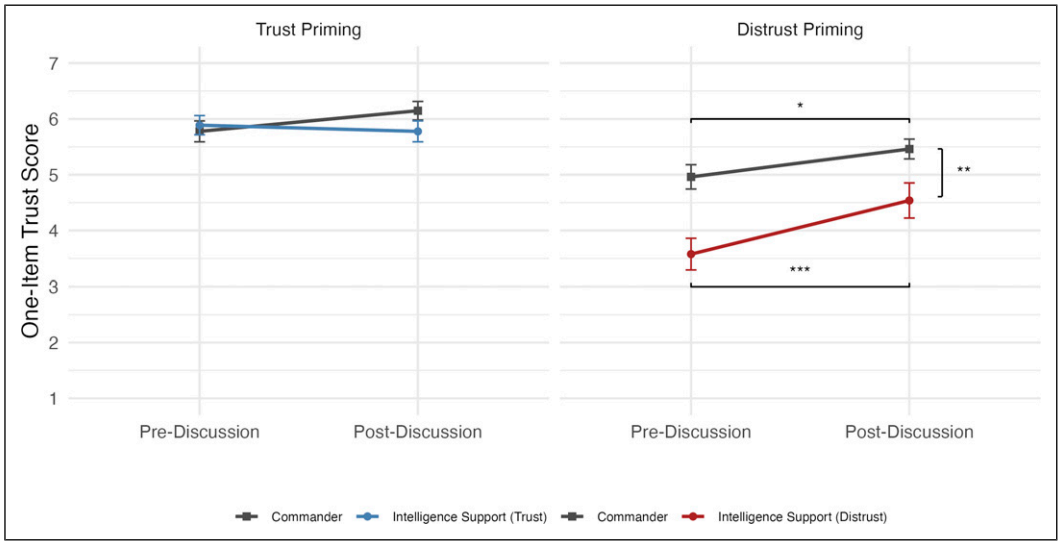


**Figure 5.** One-item trust score before vs. after discussion for each condition

post-task MDMT ratings with *Priming Condition*, *Reliability*, *UAV*, and *Role* as predictors. Because this analysis compares trust across reliability blocks, we first confirmed that the blocks imposed comparable workload demands. Reported NASA-TLX indicated low-to-moderate workload (High-High: $M = 5.34$, $SD = 2.16$; Mixed: $M = 5.56$, $SD = 1.97$; Low-Low: $M = 5.43$, $SD = 2.02$). A linear mixed model with *Reliability* as the predictor showed no significant effect of *Reliability*, $p > .05$, indicating that the varying scenarios did not produce differential workload. Trust contagion was evaluated through pairwise comparisons of *Priming Condition* within each *Reliability* level, averaged over UAV and Role. A significant priming condition effect would indicate that trust-primed and distrust-primed teams differed in post-task MDMT. Trust transfer was evaluated through pairwise comparisons of *Reliability* within each *UAV*; downward contamination would be indicated by lower MDMT for North in Mixed than in High-High. NARS emerged as a significant covariate ($\beta = -0.43$, $t(99.16) = -2.39$, $p = .02$, $\omega^2 = .04$), with more negative attitudes associated with lower UAV trust, and was retained in the final model (Table 3).

*Trust Contagion Between Human Teammates.* No significant *Priming Condition* × *Reliability* interaction or *Priming Condition* main effect emerged ($p > .05$). Participants in trust- and distrust-primed teams did not differ significantly in MDMT at any reliability level. H2b was not supported. Instead, a significant *Reliability* main effect emerged ($F(2, 520) = 404.07$, $p < .001$, $\omega^2 = .61$), with MDMT decreasing monotonically as reliability declined. Post-hoc comparisons (Figure 6A) confirmed that High-High ($M = 5.39$, $SE = 0.11$) exceeded Mixed ($M = 4.50$, $SE = 0.11$), $p_{adj} < .001$, $d = 0.87$, and Low-Low ($M = 2.90$, $SE = 0.11$), $p_{adj} < .001$, $d = 2.46$, and Mixed exceeded Low-Low, $p_{adj} < .001$, $d = 1.59$. UAV reliability, not teammates' primed attitude, was the primary driver of post-task trust.

*Trust Transfer Between UAVs.* The analysis revealed a significant *Reliability* × *UAV* interaction ($\beta = -2.43$, $t(520) = -9.80$, $p < .001$, $\omega^2 = .25$), indicating that the effect of reliability on trust differed by UAV. Post-hoc comparisons (Figure 6B) revealed different patterns for the two UAVs. For North, trust did not differ between High-High ($M = 5.25$, $SE = 0.12$) and Mixed ($M = 5.50$, $SE = 0.12$), $p_{adj} > .05$, but both exceeded

**Table 3.** Post-Task MDMT Analysis Summary

| *Predictors* | *Estimates* | *CI* | *p* |
|---|---|---|---|
| (Intercept) | 6.41 | [5.38, 7.44] | <.001 |
| Condition [Distrust Priming] | 0.11 | [−0.37, 0.60] | .647 |
| Reliability [Mixed] | 0.36 | [0.01, 0.70] | **.043** |
| Reliability [Low-Low] | −2.39 | [−2.73, −2.04] | **<.001** |
| UAV [South] | 0.32 | [−0.03, 0.66] | .070 |
| Role [Intelligence Support] | −0.04 | [−0.30, 0.23] | .795 |
| NARS | −0.43 | [−0.78, −0.08] | **.017** |
| Condition [Distrust Priming] × Reliability [Mixed] | −0.20 | [−0.69, 0.29] | .425 |
| Condition [Distrust Priming] × Reliability [Low-Low] | 0.34 | [−0.15, 0.84] | .172 |
| Condition [Distrust Priming] × UAV [South] | −0.08 | [−0.57, 0.41] | .755 |
| Reliability [Mixed] × UAV [South] | −2.43 | [−2.92, −1.95] | **<.001** |
| Reliability [Low-Low] × UAV [South] | −0.49 | [−0.97, 0.00] | .051 |
| (Condition [Distrust Priming] × Reliability [Mixed]) × UAV [South] | 0.31 | [−0.39, 1.00] | .386 |
| (Condition [Distrust Priming] × Reliability [Low-Low]) × UAV [South] | −0.12 | [−0.82, 0.58] | .735 |
| Random effects | | | |
| $\sigma^2$ | 0.83 | | |
| $\tau_{00\ ID}$ | 0.36 | | |
| $\tau_{00\ Team_ID}$ | 0.21 | | |
| $N_{Team_ID}$ | 53 | | |
| $N_{ID}$ | 106 | | |
| Observations | 636 | | |
| Marginal $R^2$/conditional $R^2$ | .510/.709 | | |

*Note*. Bold values indicate statistically significant predictors ($p < .05$).

Low-Low ($M = 3.03$, $SE = 0.12$), $p_{adj} < .001$, $d = 1.55$ and $d = 1.73$, respectively. For South, all comparisons were significant: High-High ($M = 5.53$, $SE = 0.12$) exceeded Mixed ($M = 3.50$, $SE = 0.12$), $p_{adj} < .001$, $d = 1.41$, and Low-Low ($M = 2.77$, $SE = 0.12$), $p_{adj} < .001$, $d = 1.93$, and Mixed exceeded Low-Low, $p_{adj} < .001$, $d = 0.51$. Because trust in North did not decline from High-High to Mixed, negative system-wide trust, as predicted by H3, was not supported.

## *Actor-Partner Interdependence Model (APIM)*

To model interpersonal trust transmission between the two operators, we then employed APIM. Standardized regression coefficients and significance tests are presented in Table 4. No actor effects were significant in any condition, indicating that post-discussion trust did not reliably predict participants' own post-task UAV trust.

One partner effect reached significance: under Distrust priming in the High-High reliability condition, Commander's post-discussion trust was negatively associated with Intelligence Support's post-task MDMT ($\beta = -0.27$, $p = .007$), such that higher Commander trust after discussion predicted lower Intelligence Support post-task trust.

UAV effects were largely consistent with the designed reliability manipulation. In the Mixed condition, both operators rated South (58%) significantly lower than North (95%) under both Trust (Commander: $\beta = -1.35$, $p < .001$; Intelligence Support: $\beta = -1.17$, $p < .001$) and Distrust priming conditions (Commander: $\beta = -1.10$, $p < .001$; Intelligence Support: $\beta = -1.15$, $p < .001$). The remaining UAV effects were non-significant, except that, under Distrust priming in the Low-Low condition, Intelligence Support rated South lower than North ($\beta = -0.22$, $p = .005$), even though the two UAVs performed comparably (63% vs. 58%).

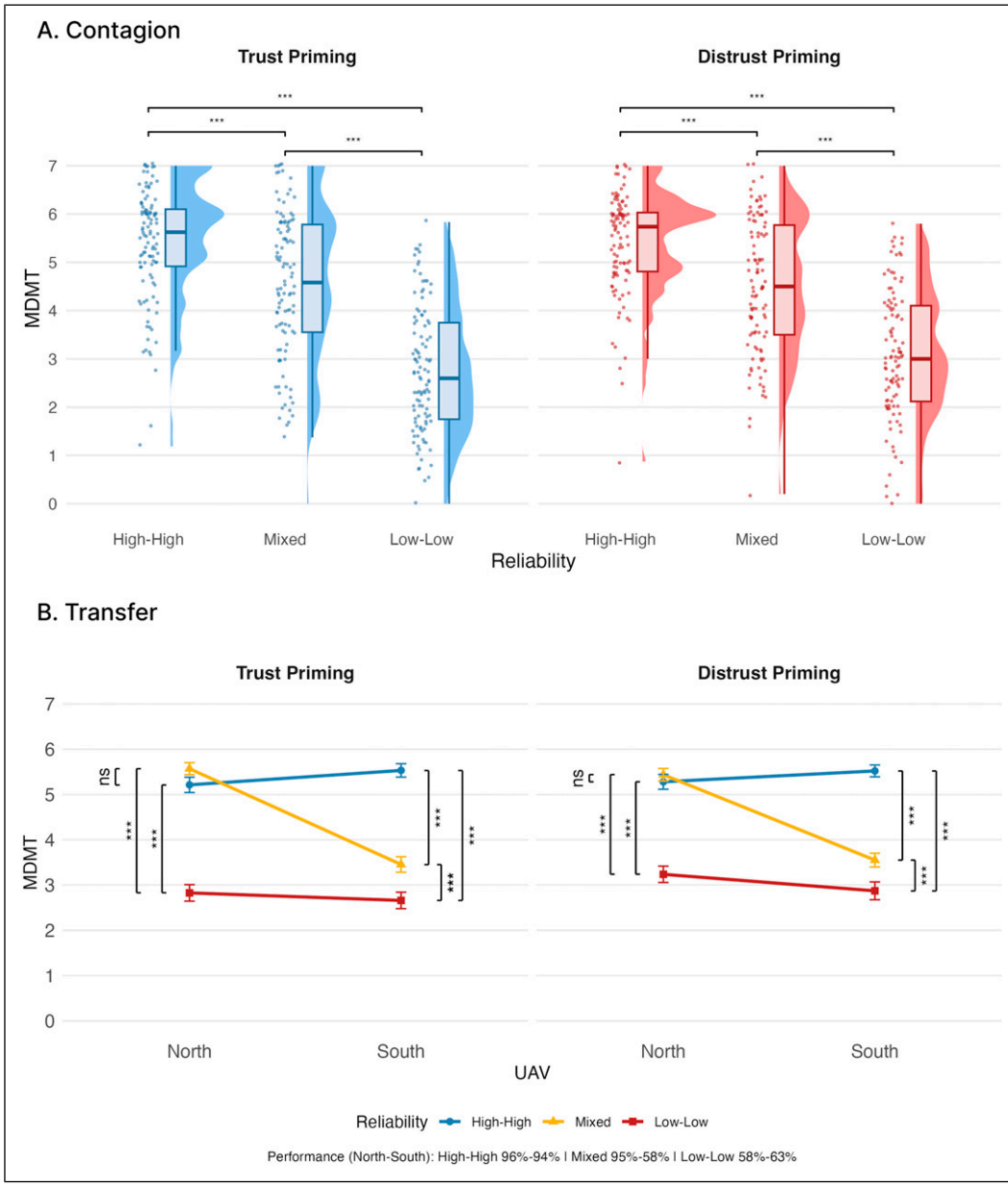


**Figure 6.** Post-task trust by reliability, UAV, and priming condition.
*Note.* (A) Trust Contagion: MDMT rating across reliability levels in each condition. (B) Trust Transfer: MDMT rating across reliability for each UAV in each condition

## Discussion

### *Trust Priming Shifted Initial Trust*

The trust priming hypothesis (H1) was supported: brief performance demonstrations shaped trust. Intelligence Support's trust ratings rose following the highreliability video and fell following the low-reliability video. A 30-second performance demonstration was therefore sufficient to shift trust in opposite directions, consistent with prior work showing that training experiences shape operators'

**Table 4.** Standardized Regression Coefficients for Commander and Intelligence Support by Condition

| Priming | Reliability | CO-Actor (β/*p*) | CO-Partner (β/*p*) | IS-Actor (β/*p*) | IS-Partner (β/*p*) | CO-UAV effect (β/*p*) | IS-UAV effect (β/*p*) |
|---|---|---|---|---|---|---|---|
| Trust | HH | 0.09/.46 | 0.00/.99 | 0.05/.82 | 0.27/.08 | 0.14/.23 | 0.24/.01 |
| Trust | M | 0.15/.20 | −0.14/.38 | 0.12/.30 | 0.10/.25 | **−1.35/<.001** | **−1.17/<.001** |
| Trust | LL | 0.01/.96 | 0.23/.31 | −0.002/.99 | 0.30/.05 | −0.12/.21 | −0.08/.12 |
| Distrust | HH | 0.09/.47 | 0.08/.42 | 0.10/.24 | **−0.27/.007** | 0.22/.14 | 0.07/.34 |
| Distrust | M | −0.13/.27 | −0.15/.13 | 0.04/.61 | −0.19/.04 | **−1.10/<.001** | **−1.15/<.001** |
| Distrust | LL | 0.12/.47 | 0.08/.59 | 0.28/.05 | 0.01/.94 | −0.22/.08 | **−0.22/.005** |

*Note.* CO = Commander; IS = Intelligence Support; HH = High-High; M = Mixed; LL = Low-Low. Each row is a separate model. Effects are labeled by the outcome operator (e.g., IS-Partner = Commander's trust predicting Intelligence Support's trust). UAV effects compare South with North (negative values indicate lower trust in South). Statistical significance determined using Bonferroni-corrected $\alpha$ = .008 to control for multiple testing across six experimental conditions.

expectations of automation (Körber et al., 2018; Sauer et al., 2016).

## *Trust Contagion Did Not Occur as Predicted*

Neither pre-task (H2a) nor post-task (H2b) trust contagion was supported. The primed attitude did not spread from Intelligence Support to Commander. This contrasts with prior studies that observed trust contagion (Erengin et al., 2026; Hidalgo et al., 2025; Rojas et al., 2025) and suggests that features of our design may have acted as boundary conditions that limit trust contagion. We discuss each hypothesis in turn, then propose three such design features.

*Pre-Task (H2a).* In the pre-task discussion period, neither operator's trust changed significantly in the Trust condition, whereas both increased in the Distrust condition. In the Trust condition, the Commander's baseline trust was already comparable to the trust-primed Intelligence Support's, leaving little gap for the discussion to close. In the Distrust condition, Intelligence Support's trust shifted upward, suggesting that the brief priming Intelligence Support received did not establish a stable trust level that persisted. The primed attitude therefore had little chance to spread to the Commander. We cannot, however, account for why the Commander's trust also increased in the Distrust condition, a shift opposite to the direction contagion would predict.

*Post-Task (H2b).* Participants in the two conditions did not differ in trust ratings at any reliability level. Trust was instead driven by UAV reliability, consistent with prior literature identifying reliability as fundamental to trust development (Lee & See, 2004). Higher NARS scores additionally predicted lower trust, in line with prior work on individual differences in trust (Hancock et al., 2011, 2021; Kaplan et al., 2023). Once the task began, each operator observed both UAVs firsthand, so direct experience of each UAV became the dominant basis for trust, outweighing any expressed teammate attitude.

*Boundary Conditions.* These results indicate that trust contagion does not occur automatically when members have different trust levels. We identified three key features that may have constrained contagion. First, stability of trust priming: we established the trust difference through brief priming rather than the extended hands-on training typically used to establish trust differences (Körber et al., 2018; Sauer et al., 2016). This may have produced an unstable attitude that did not persist or spread between teammates. Second,

explicit versus implicit communication channel: prior studies conveyed a teammate's attitude through explicit communication, including confederates (Rojas et al., 2025), scripted interactions or vignettes (Erengin et al., 2026), or numeric trust ratings (Hidalgo et al., 2025). Operators in our study had to infer one another's trust from unscripted conversation. The absence of explicit communication, especially over the short experimental session, may have left little opportunity for trust contagion. Third, interdependency between human operators: both operators independently monitored both UAVs and could observe their performance firsthand, so a teammate's opinion added little beyond direct experience. Operators in Hidalgo et al. (2025) instead worked with different agents separately, so a teammate's assessment conveyed reliability information that the operator could not obtain directly. This is consistent with evidence that operators calibrate trust to an agent mainly through direct observation rather than by adopting a teammate's expressed attitude (Duan et al., 2026). Future studies should explicitly test how trust priming stability, communication channels, and task interdependency influence trust contagion in teams.

## *No Negative Trust Transfer Between UAVs: Evidence for Component-Specific Trust*

We found no support for negative system-wide trust (H3). Trust in the reliable North UAV did not decline when it was paired with an unreliable South UAV. Instead, trust in North tracked its own reliability regardless of South's performance, more consistent with a component-specific trust account.

Trust in the less reliable South, however, was higher when paired with a reliable North (Mixed) than when North was also unreliable (Low-Low), even though South's performance was comparable across the two conditions (58% vs. 63%; Table 1). This "pull-up" may reflect positive trust transfer, in which North's reliability raised trust in the co-present South through a halo effect (Chung & Yang, 2026b; Nisbett & Wilson, 1977). However, because North was always rated before South in the post-task surveys, a high North rating may have carried over to the South rating that followed. As these explanations are confounded in our design, we do not interpret the pull-up as evidence for system-wide trust. Our support for the component-specific pattern rests on North, whose trust tracked its own reliability and was unaffected by rating order.

This account contrasts with prior literature, in which system-wide trust is more commonly found (Lopez & Pak, 2020). One factor that may distinguish our study is workload. Lopez and Pak (2020) noted that system-wide trust tends to be adopted when workload is high. Because reported workload in our task was low to moderate and did not vary across reliability conditions, operators likely had the spare capacity to evaluate each UAV independently. Future studies should explicitly test the relationship between workload and the trust strategies operators adopt when working with multiple agents.

## *Exploratory Analysis: Actor-Partner Effects*

An additional exploratory analysis using APIM examined interpersonal influence at the level of the operator dyad. A significant actor effect reflects within-person stability: an operator's post-discussion trust predicting their own post-task trust. A significant partner effect reflects interpersonal influence: a teammate's post-discussion trust predicting an operator's post-task trust. We found no significant actor effects in any condition. Post-discussion trust therefore did not carry over to post-task trust, indicating that trust was shaped by the task itself rather than by trust held beforehand. This is consistent with our main finding that direct experience with UAV performance was the primary driver of trust development.

One partner effect reached significance: under Distrust priming in the High-High reliability condition, higher Commander post-discussion trust predicted lower Intelligence Support post-task trust. This divergence suggests that when the Commander expressed relatively high trust, the distrust-primed Intelligence Support responded with greater skepticism. This interpretation warrants caution: although it survived Bonferroni correction, it was the only significant partner effect among those tested. The largely non-significant partner effects are consistent with the linear mixed model results, which found no evidence of trust contagion. The two approaches converge on an overall absence of interpersonal trust spread. We therefore treat this single directional association as a hypothesis for future work.

One UAV effect also reached significance beyond the reliability differences expected in the Mixed condition. Under Distrust priming in the Low-Low reliability condition, Intelligence Support rated the South UAV lower than the North even though the two performed comparably in that condition (63% vs. 58%). One possibility is that Distrust priming heightened sensitivity to accumulated performance. Across the other reliability conditions, South had underperformed North, and priming toward distrust may have made that history more salient. As with the partner effect, however, this effect survived correction but was isolated, and we present it only as a hypothesis for future work. By modeling the dyad directly, APIM surfaced directional, role-specific associations. Even as exploratory results, they suggest that APIM may help future multi-human multi-agent research model the direction of influence between specific operators.

## Implications

Theoretically, our findings partly diverged from prior literature on trust contagion and trust transfer. Rather than occurring automatically, both processes appeared to depend on the team and task context. Because neither reliably occurred in our task, we could not examine how the two interact—a question this study had set out to address. A framework specifying the boundary conditions that shape these dynamics may therefore be needed. Our study points to a few candidate conditions worth investigating in future research: for contagion, attitude stability, channel explicitness, and role interdependency; for trust transfer, operator workload. Testing them will require designs that manipulate these factors directly. The APIM results illustrate the value of modeling influence that flows between operators, and further research may clarify the role of dyad-level models in multi-human multi-agent teams. Practically, when operators can directly observe each agent, trust calibrates to agent-specific performance. This has implications for both training and agent design. Training should include direct interaction with each agent to support trust calibration toward each agent, and agent interfaces can display each agent's performance history to further support calibration.

## Limitations

A few limitations exist in the present study. First, the low task interdependency in our RSTA design limited social interaction. Because each operator could independently verify UAV performance through direct observation, communication between participants may have been limited, reducing opportunities for contagion effects to emerge. Future research can systematically vary task interdependency to examine how it moderates trust contagion in HATs. Second, the order of post-task trust ratings was fixed, with the North UAV always rated before the South. This introduced a potential order effect that confounds the apparent positive trust transfer we observed. We cannot fully separate a genuine transfer effect from the order effect. Third, our trust measurement did not fully distinguish trust from distrust. The distrust priming condition was designed to induce negative expectations toward the UAVs, more consistent with distrust as an active negative belief. The MDMT measures, however, did

not separate trust and distrust as two distinct constructs. Future studies should consider adopting measures that directly measure trust versus distrust. Finally, two features limit the generalizability of our findings. The sample was predominantly college students rather than experienced military or professional UAV operators. The two-human, two-UAV configuration may also differ substantially from larger teams. Future studies can collect data from domain-experienced operators to assess how expertise influences trust dynamics and can examine larger teams to test how these dynamics scale.

## Conclusion

This study examined trust dynamics in co-located multi-human multi-UAV teams, investigating trust contagion between human teammates and trust transfer between UAVs. UAV reliability emerged as the primary determinant of operators' trust, with no evidence that primed trust attitudes propagated through social interaction. Operators under low-to-moderate workload maintained separate trust in each UAV according to its reliability, more consistent with component-specific trust than system-wide trust. The APIM illustrated the value of dyad-level modeling for examining interpersonal influence in multi-human autonomy teams. Together, these findings suggest that trust contagion and transfer may not occur automatically but instead depend on team and task factors, identifying which is a promising direction for future research.

## Key Points

- UAV reliability was the primary driver of operators' trust
- Primed trust attitudes did not reliably spread between human operators
- Operators calibrated trust to each UAV, consistent with component-specific trust
- Actor-Partner Interdependence Model offers a way to model trust influence between operators in multi-human teams

### Acknowledgments

This work was supported by the Georgia Tech Research Institute under the Independent Research and Development (IRAD) program. We would also like to thank Emanuel Rojas, Trevor Patten, and JingYan Xu for their assistance in data collection and administration.

### ORCID iDs

Debbie Yutung Hsu https://orcid.org/0009-0003-2027-2353

Mengyao Li https://orcid.org/0000-0002-0819-4693

### Funding

The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This research was supported by Georgia Tech Research Institute Independent Research and Development (IRAD) program.

### Declaration of Conflicting Interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

### Supplemental Material

Supplemental material for this article is available online.

### References

Arnold, T., & Scheutz, M. (2017). Beyond moral dilemmas: Exploring the ethical landscape in HRI. In *Proceedings of the 2017 ACM/IEEE international conference on human-robot interaction* (pp. 445–452). Association for Computing Machinery. https://doi.org/10.1145/2909824.3020255

Bartneck, C., Nomura, T., Kanda, T., Suzuki, T., & Kato, K. (2005). A cross-cultural study on attitudes towards robots. In Proceedings of the 11th international conference on human-computer interaction (HCI international 2005), Las Vegas, NV, USA, July 22-27, 2005.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Chavaillaz, A., Wastell, D., & Sauer, J. (2016). System reliability, performance and trust in adaptable automation. *Applied Ergonomics*, *52*(2016), 333–342. https://doi.org/10.1016/j.apergo.2015.07.012

Chen, J. Y. C., & Barnes, M. J. (2013). *Human-agent teaming for multi-robot control: A literature review (Report no. ARL-TR-6328)*. U.S. Army Research Laboratory. https://doi.org/10.21236/ADA583900

Chung, H., & Yang, X. J. (2026a). Multilevel, multireferent trust, communication, and performance in human–agent teams: Evaluating the effects of agent reliability pairing. *The American Journal of Psychology*, *139*(1), 113–133. https://doi.org/10.5406/19398298.139.1.07

Chung, H., & Yang, X. J. (2026b). System-wide trust (SWT) versus component-specific trust (CST) in multi-agent human–agent teams: Individual variability in trust bias. *Human Factors*, *68*(12), 1624-1642. Advance online publication. https://doi.org/10.1177/00187208261464006

Cook, W. L., & Kenny, D. A. (2005). The actor–partner interdependence model: A model of bidirectional effects in developmental studies. *International Journal of Behavioral Development*, *29*(2), 101–109. https://doi.org/10.1080/01650250444000405

Deutsch, M., & Gerard, H. B. (1955). A study of normative and informational social influences upon individual judgment. *The Journal of Abnormal and Social Psychology*, *51*(3), 629–636. https://doi.org/10.1037/h0046408

Dimoka, A. (2010). What does the brain tell us about trust and distrust? Evidence from a functional neuroimaging study. *MIS Quarterly*, *34*(2), 373–396. https://doi.org/10.2307/20721433

Duan, W., Zhou, S., Scalia, M. J., Freeman, G., Gorman, J., Tolston, M., McNeese, N. J., & Funke, G. (2025). Understanding the processes of trust and distrust contagion in human–AI teams: A qualitative approach. *Computers in Human Behavior*, *165*(2025), 108560. https://doi.org/10.1016/j.chb.2025.108560

Duan, W., Zhou, S., Scalia, M. J., Weng, N., Yin, X., Hao, R., Freeman, G., Funke, G., Tolston, M., Gorman, J., & McNeese, N. (2026). "If they don't even trust their AI teammate, why should we?": Understanding trust contagion across multiple human-AI teams. *Proceedings of the ACM on Human-Computer Interaction*, *10*(2), 1–31. https://doi.org/10.1145/3788055

Erengin, T., Briker, R., & de Jong, S. B. (2026). You, me, and the AI: The role of third-party human teammates for trust formation toward AI teammates. *Journal of Organizational Behavior*, *47*(8), 1397–1422. https://doi.org/10.1002/job.2857

Fraune, M. R., Nishiwaki, Y., Sabanović, S., Smith, E. R., & Okada, M. (2017). Threatening flocks and mindful snowflakes: How group entitativity affects perceptions of robots. In *Proceedings of the 2017 ACM/IEEE international conference on human-robot interaction* (pp. 205–213). Association for Computing Machinery. https://doi.org/10.1145/2909824.3020248

Fraune, M. R., Sherrin, S., Sabanović, S., & Smith, E. R. (2015). Rabble of robots effects: Number and type of robots modulates attitudes, emotions, and stereotypes. In *Proceedings of the tenth annual ACM/IEEE international conference on human-robot interaction* (pp. 109–116). Association for Computing Machinery. https://doi.org/10.1145/2696454.2696483

Geels-Blair, K., Rice, S., & Schwark, J. (2013). Using system-wide trust theory to reveal the contagion effects of automation false alarms and misses on compliance and reliance in a simulated aviation task. *The International Journal of Aviation Psychology*, *23*(3), 245–266. https://doi.org/10.1080/10508414.2013.799355

Guo, Y., Yang, X. J., & Shi, C. (2023). TIP: A trust inference and propagation model in multi-human multi-robot teams. In *Companion of the 2023 ACM/IEEE international conference on human-robot interaction* (pp. 639–643). Association for Computing Machinery. https://doi.org/10.1145/3568294.3580164

Hancock, P. A., Billings, D. R., Schaefer, K. E., Chen, J. Y. C., de Visser, E. J., & Parasuraman, R. (2011). A meta-analysis of factors affecting trust in human-robot

interaction. *Human Factors*, *53*(5), 517–527. https://doi.org/10.1177/0018720811417254

Hancock, P. A., Kessler, T. T., Kaplan, A. D., Brill, J. C., & Szalma, J. L. (2021). Evolving trust in robots: Specification through sequential and comparative meta-analyses. *Human Factors*, *63*(7), 1196–1229. https://doi.org/10.1177/0018720820922080

Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (task load index): Results of empirical and theoretical research. In P. A. Hancock & N. Meshkati (Eds.), *Advances in psychology: Vol. 52. Human mental workload* (pp. 139–183). North-Holland. https://doi.org/10.1016/S0166-4115(08)62386-9

Hatfield, E., Cacioppo, J. T., & Rapson, R. L. (1993). Emotional contagion. *Current Directions in Psychological Science*, *2*(3), 96–100. https://doi.org/10.1111/1467-8721.ep10770953

Hidalgo, D. T., Guo, Y., Lyons, J. B., Shi, C., & Yang, X. J. (2025). Trust propagation in multi-operator multi-autonomy (MOMA) teams: Friends versus strangers. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *69*(1), 1742–1748. https://doi.org/10.1177/10711813251370736

Hoff, K. A., & Bashir, M. (2015). Trust in automation: Integrating empirical evidence on factors that influence trust. *Human Factors*, *57*(3), 407–434. https://doi.org/10.1177/0018720814547570

Huang, L., Cooke, N. J., Gutzwiller, R. S., Berman, S., Chiou, E. K., Demir, M., & Zhang, W. (2021). Distributed dynamic team trust in human, artificial intelligence, and robot teaming. In C. S. Nam & J. B. Lyons (Eds.), *Trust in human-robot interaction* (pp. 301–319). Academic Press. https://doi.org/10.1016/B978-0-12-819472-0.00013-7

Hutchins, E. (1995). *Cognition in the wild*. MIT Press.

Jessup, S. A., Schneider, T. R., Alarcon, G. M., Ryan, T. J., & Capiola, A. (2019). The measurement of the propensity to trust automation. In J. Y. C. Chen & G. Fragomeni (Eds.), *Virtual, augmented and mixed reality: Applications and case studies* (pp. 476–489). Springer. https://doi.org/10.1007/978-3-030-21565-1_32

Kaplan, A. D., Kessler, T. T., Brill, J. C., & Hancock, P. A. (2023). Trust in artificial intelligence: Meta-analytic findings. *Human Factors*, *65*(2), 337–359. https://doi.org/10.1177/00187208211013988

Keller, D., & Rice, S. (2009). System-wide versus component-specific trust using multiple aids. *The Journal of General Psychology*, *137*(1), 114–128. https://doi.org/10.1080/00221300903266713

Kenny, D. A. (1996). Models of non-independence in dyadic research. *Journal of Social and Personal Relationships*, *13*(2), 279–294. https://doi.org/10.1177/0265407596132007

Körber, M., Baseler, E., & Bengler, K. (2018). Introduction matters: Manipulating trust in automation and reliance in automated driving. *Applied Ergonomics*, *66*(2018), 18–31. https://doi.org/10.1016/j.apergo.2017.07.006

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software*, *82*(13), 1–26. https://doi.org/10.18637/jss.v082.i13

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

Lenth, R, & Piaskowski, J. (2026). emmeans: Estimated Marginal Means, aka Least-Squares Means. R package version 2.0.5, https://rvlenth.github.io/emmeans/.

Lopez, J., & Pak, R. (2020). Does one bad machine spoil the bunch? A review of trust in multiple-component systems. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *64*(1), 1546–1550. https://doi.org/10.1177/1071181320641370

Lyons, J. B., Mator, J. D., Orr, T., Alarcon, G. M., & Barrera, K. (2024). Is the pull-down effect overstated? An examination of trust propagation among fighter pilots in a high-fidelity simulation. *Journal of Cognitive Engineering and Decision Making*, *18*(2), 99–113. https://doi.org/10.1177/15553434231225909

Muhammad, L. N. (2023). Guidelines for repeated measures statistical analysis approaches with basic science research considerations. *The*

*Journal of Clinical Investigation*, *133*(11), e171058. https://doi.org/10.1172/JCI171058

Nisbett, R. E., & Wilson, T. D. (1977). The halo effect: Evidence for unconscious alteration of judgments. *Journal of Personality and Social Psychology*, *35*(4), 250–256. https://doi.org/10.1037/0022-3514.35.4.250

Nomura, T., Suzuki, T., Kanda, T., & Kato, K. (2006). Measurement of negative attitudes toward robots. *Interaction Studies*, *7*(3), 437–454. https://doi.org/10.1075/is.7.3.14nom

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

Pinheiro, J., Bates, D., & Core Team, R. (2026). *Linear and Nonlinear Mixed Effects Models*. https://doi.org/10.32614/CRAN.package.nlme. R package version 3.1-172, https://CRAN.R-project.org/package=nlme

Ramchurn, S. D., Huynh, D., & Jennings, N. R. (2004). Trust in multi-agent systems. *The Knowledge Engineering Review*, *19*(1), 1–25. https://doi.org/10.1017/S0269888904000116

Rammstedt, B., & John, O. P. (2007). Measuring personality in one minute or less: A 10-item short version of the big five inventory in English and German. *Journal of Research in Personality*, *41*(1), 203–212. https://doi.org/10.1016/j.jrp.2006.02.001

Rojas, E., Hsu, D., Huang, J., & Li, M. (2025). Interpersonal influence matters: Trust contagion and repair in human-human-AI team. *Computers in Human Behavior: Artificial Humans*, *5*(2025), 100194. https://doi.org/10.1016/j.chbah.2025.100194

Sauer, J., Chavaillaz, A., & Wastell, D. (2016). Experience of automation failures in training: Effects on trust, automation bias, complacency and performance. *Ergonomics*, *59*(6), 767–780. https://doi.org/10.1080/00140139.2015.1094577

Schelble, B. G., Flathmann, C., Scalia, M., Zhou, S., Myers, C., McNeese, N. J., Gorman, J., & Freeman, G. (2022). *Addressing the spread of trust and distrust in distributed human-AI teaming constellations [Paper presentation]*. CHI 2022 Workshop on Trust and Reliance in AI-Human Teams (TRAIT). https://beauschelble.com/papers/schelble-2022-trust-constellations.

Schneiders, E., Cheon, E., Kjeldskov, J., Rehm, M., & Skov, M. B. (2022). Non-dyadic interaction: A literature review of 15 years of human-robot interaction conference publications. *ACM Transactions on Human-Robot Interaction*, *11*(2), 1–32. https://doi.org/10.1145/3488242

Ullman, D., & Malle, B. F. (2019). Measuring gains and losses in human-robot trust: Evidence for differentiable components of trust. In *2019 14th ACM/IEEE international conference on human-robot interaction (HRI)* (pp. 618–619). IEEE. https://doi.org/10.1109/HRI.2019.8673154

Wagner, J. A., III. (1995). Studies of individualism-collectivism: Effects on cooperation in groups. *Academy of Management Journal*, *38*(1), 152–172. https://doi.org/10.2307/256731

Walliser, J. C., de Visser, E. J., & Shaw, T. H. (2016). Application of a system-wide trust strategy when supervising multiple autonomous agents. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *60*(1), 133–137. https://doi.org/10.1177/1541931213601031

Wróbel, M., & Imbir, K. K. (2019). Broadening the perspective on emotional contagion and emotional mimicry: The correction hypothesis. *Perspectives on Psychological Science*, *14*(3), 437–451. https://doi.org/10.1177/1745691618808523

Yin, X., Zhou, S., Scalia, M. J., Hao, R., Duan, W., Tuttle, J., Weng, N., Tolston, M. T., Funke, G. J., Freeman, G., McNeese, N. J., & Gorman, J. C. (2025). Trust contagion in team of teams (ToT) for human-autonomy teaming. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *69*(1), 442–446. https://doi.org/10.1177/10711813251374545

## Author Biographies

**Debbie Yutung Hsu** is a second-year engineering psychology Ph.D. student at Georgia Institute of Technology. She received her master of science in human-computer

interaction at the Georgia Institute of Technology in 2025.

**Theresa Kessler** is a senior research scientist in the Applied Systems Laboratory at Georgia Tech Research Institute. She received her Ph.D. in human factors and cognitive psychology from the University of Central Florida in 2020.

**Elizabeth Weldon** is a senior research scientist in the Applied Systems Laboratory at Georgia Tech Research Institute. She received her master of science in human-computer interaction at the Georgia Institute of Technology in 2010.

**Mengyao Li** is an assistant professor in the School of Psychological and Brain Sciences at Georgia Institute of Technology. She received her Ph.D. in industrial and systems engineering from the University of Wisconsin-Madison in 2023.