An Integrative Review of Trust in Automation

David T. Feltner

Department of Psychology, NC State University

Author Note

David Feltner is a Ph.D. student in the Human Factors and Applied Cognition Program at NC State. I have no known conflicts of interest to disclose.

Correspondence concerning this review should be addressed to Dave Feltner, Department of Psychology, NC State, Raleigh, NC 27695, United States. Email: dtfeltne@ncsu.edu or david.t.feltner2.mil@army.mil.

Abstract

Trust in automation was a critical factor in human-machine collaboration, particularly in high-stakes environments where automated systems must perform reliably under pressure. While Lee and See's (2004) model of dynamic trust has served as a foundational framework for understanding how trust evolves, research over the past two decades has identified additional factors that shape trust calibration between human operators and automation. This integrative review synthesizes empirical findings to refine the trust model, incorporating individual differences in attention control, pre-existing attitudes, task interdependence, perceived risk, automation reliability, transparency, and the physical characteristics of robotic systems. Notably, trust is an attitude, distinct from dependence, which is a behavior that determines whether automation is actually used. As automation expands into military operations, ensuring appropriate reliance, rather than simply fostering trust, becomes paramount. Future research must shift towards experimental methodologies with greater ecological validity, capturing decision-making in dynamic environments. These findings will inform the development of robotic teammates that align with operators' cognitive needs and integrate into mission execution. By optimizing system design and human-machine interaction strategies, the military can field automation that enhances operational effectiveness, ensuring that robotic systems are not just trusted, but actively depended upon in combat.

Key Words: automation, trust in automation, dependence, transparency, risk, human-robot teaming

An Integrative Review of Trust in Automation

State of Art

Integrating the Trust Literature

In 2004, Lee and See published their dynamic theory of trust in automation based on the human-human trust literature. They identified the contextual drivers that fed into individuals' initial beliefs, how individuals formed their intention to use the automation, and specific characteristics of that automation. They showed that a person's trust in automation was dynamic and continuously calibrated based on contextual factors (Lee & See, 2004). In the 21 years since that publication, empirical research has grown and robotic capabilities have exploded. During that time, the same trust model - shown in Figure 1 - has maintained its prominence within the human factors field and its influence is shown throughout the studies highlighted in this review. However, human factors and social robotics have refined their research and discovered new insights into human-automation teaming. The research on automation reliability, anthropomorphism, and task interdependence have demonstrated their importance, and these trust drivers should be included in future trust calibration models. A full review of trust in automation is outside the scope of this paper, but this integrated review demonstrates that these concepts should be explicitly captured under *Reliance Action*, *Automation*, and *Display*. Based on the research, I have provided an updated trust model with these drivers - seen in Figure 4 - and briefly outlined a future direction as we seek to further understand the relationship between trust (an *attitude*) and dependence (a *behavior*) on automation.

Automation on the Battlefield

Over the years, robotic systems have played significant roles for U.S. military forces. As the Army shifts focus toward Large Scale Combat Operations against near-peer threats, modernization efforts are intensifying. The Army is equipping its ground combat formations with robotic teammates designed to enhance mission execution. These systems amplify soldier capabilities and improve survivability (Endsley, 2015; Szegedi et al., 2017). The Department of Defense (2023) emphasized that integrating these technologies is essential for maintaining battlefield advantage. The Army has already fielded semi-autonomous robots capable of receiving tactical tasks and maneuvering across complex terrain. These robotic systems introduce unique capabilities intended to support soldiers on the ground such as: reconnaissance, bomb disposal, communication relays, or facilitating resupply to units on the front line (Young & Winstead, 2025). Army units will deploy these robotic enablers as essential teammates to boost performance and effectiveness. Given their pivotal role in future missions, it is essential to examine the factors influencing the trust and dependence of human soldiers on these robot teammates.

For soldiers to trust robotic teammates, they must first understand the robots' roles and responsibilities, engage with them confidently, and properly rely on them in combat. While fielding advanced automation promises operational gains, trust is not guaranteed. Soldiers must believe that robotic systems will perform reliably under pressure, make sound decisions, and enhance—not hinder—mission outcomes. Without this trust, operators may hesitate, second-guess automation decisions, or override system recommendations, diminishing the advantages these technologies offer. Developing robotic teammates that earn and maintain soldiers' trust is essential for achieving the Army's modernization objectives and ensuring dominance on the future battlefield. Automation research has explored the impacts of attitudes towards robots, task interdependence, reliability, transparency, and anthropomorphism impact a person's trust. This research has impacted, and will continue to impact, how the right tools get into soldiers' hands.

Theories of Trust in Automation

Lee and See's Conceptual Model of Trust in Automation

Lee and See (2004) offered a framework explaining how trust shapes human reliance on automation tools. Their model emphasized that trust would guide future actions, particularly when automation became too complex. They identified drivers that impacted throughout the entire process, from previously held belief up to reliance on the automation, and demonstrated how feedback from these drivers impact trust throughout (see Figure 1 below).

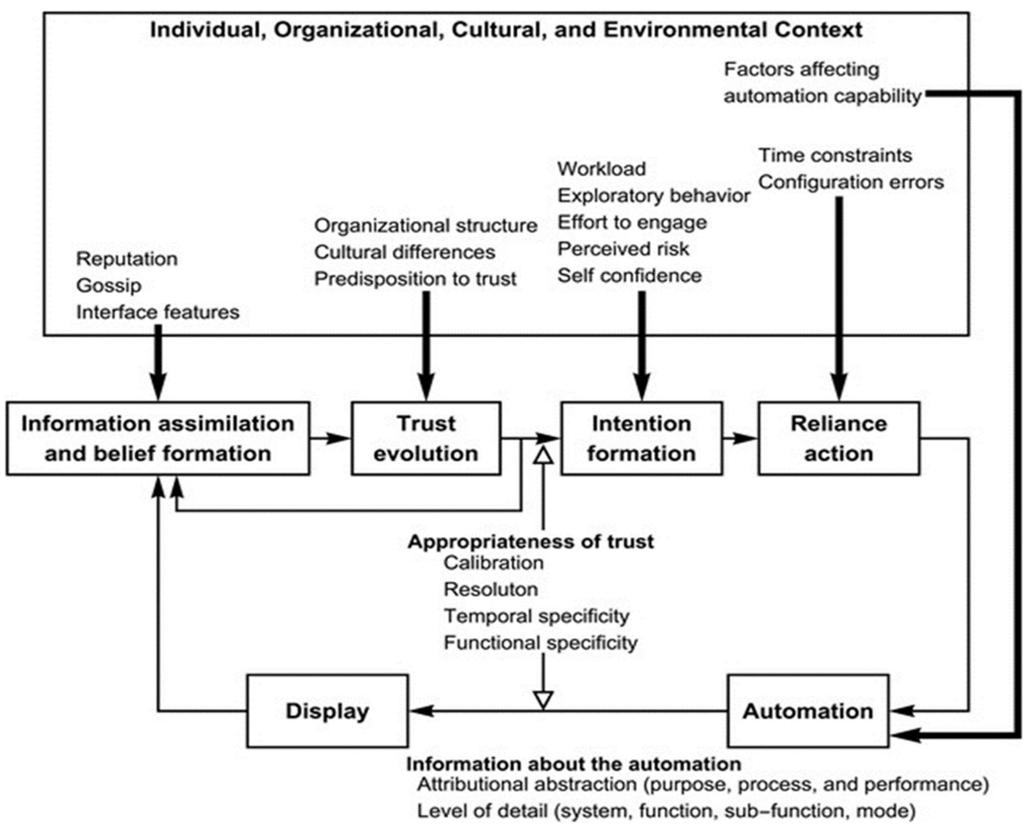

Figure 1. A conceptual model of the dynamic process that governs trust, taken from Lee and See (2004). Trust developed through a sequential process, beginning with information assimilation and belief formation, where users integrated prior knowledge to assess system

reliability. As they interacted with automation, trust evolved based on factors like workload and perceived risk. Users then determined whether to rely on automation before ultimately experiencing its actual or perceived reliability. Throughout, feedback loops shaped trust through ongoing interactions.

This model has been applied across multiple domains, as it applied to any situation that relied upon human-automation interaction. Their findings demonstrated that over-trust or under-trust often results from miscalibration rather than simple system performance (Hoff & Bashir, 2015). Within aviation, pilots' trust was highly susceptible to reliability fluctuations and minor errors caused a disproportionate loss of trust (Merritt et al., 2015). This aligned with Lee and See's (2004) proposition that trust was sensitive to deviations from expected performance in high perceived risk scenarios, reinforcing the need for automation to be both reliable and predictable to sustain trust over time. Similarly, medical providers' initial trust of decision-support systems was often influenced by external factors such as authority endorsements, but long-term trust was contingent on consistent and interpretable system performance (Dzindolet et al., 2003). The trust model explained these findings through the progression from *Information Assimilation* to *Automation* (as seen in Figure 1) and the feedback loops after using the automation (Lee & See, 2004).

Overall, the dynamic theory of trust proposed by Lee and See (2004) proved to be a valuable framework for understanding trust in various domains. Its emphasis on trust calibration, and the influence of feedback mechanisms made it particularly well-suited for applications in human-automation teaming, and decision-making under uncertainty. This model has remained the primary model to understand trust in automation because of its holistic approach towards the human and the automation.

Situational Trust

Situational trust is fluid, adapting to changes in operations, environment, and automation behavior. Hoff and Bashir's (2015) created a trust framework consisting of dispositional, situational, and learned trust. People's dispositional trust remained relatively stable and was based on an individual's general propensity to trust automation. Whereas situational trust fluctuated in response to contextual factors such as workload, perceived risk, task complexity, and system transparency. Situational trust was highly sensitive to these contextual cues, requiring rapid recalibration. Situational trust played a pivotal role by governing real-time human-automation interactions. The adaptive nature of situational trust meant that robotic teammates needed to dynamically adjust transparency levels and communication strategies to align with their human counterparts requirements (Hoff & Bashir, 2015).

In high-risk military environments, operators continually monitored automation performance, especially under elevated workloads where trust became a critical factor (Sato et al., 2020). Their research revealed that participants exhibited greater trust in automation when workload and operational risks were significant. This indicated that under pressure, operators were more willing to rely on automation. However, this reliance was conditional; people were influenced by factors such as historical system performance, real-time behavior, and transparent decision-making processes. Together, these factors underscored that situational trust was a strategic calibration process in which operators balanced perceived automation benefits against operational risks.

A person's perceived risk plays a significant role in how people calibrate their situational trust with an automation support tool. Research showed that under high-risk conditions, people tended to verify automation outputs, even when their cognitive resources were stretched. Bhaskara et al. (2021) found that participants, despite being engaged in reconnaissance tasks, took the time to verify which unmanned reconnaissance vehicle (UV) to deploy—an effort that required completing a tedious mathematical calculation. This verification occurred even though their attention was needed elsewhere, primarily due to the potential financial penalty if the AI-selected UV was incorrect. In high-stakes scenarios where errors carried significant consequences, individuals instinctively engaged in more rigorous verification processes, highlighting the need for explainable automation outputs. These findings aligned with Lee and See's (2004) concept of trust appropriateness, which emphasized that trust should be calibrated to match an automation system's capabilities relative to the risks involved.

Task interdependence played a critical role in shaping situational trust, particularly in environments requiring close collaboration between humans and automation. When tasks were highly interdependent, automation that performed reliably helped reduce operator stress and fostered greater trust (Zhao et al., 2020). Repeated success reinforced confidence in the system, making users more willing to rely on automation for decision-making. The role of explainable communication became even more vital in these scenarios, ensuring that operators understood how and why automation reached its conclusions (Verhagen et al., 2021). Without this clarity, even a reliable system risked eroding trust over time. Additionally, research indicated that in high-interdependence tasks, people perceived robotic teammates as true collaborators rather than mere tools, which strengthened overall team cohesion and trust (O'Neill et al., 2022).

Collectively, these insights illustrated that situational trust in military automation was a multifaceted construct influenced by workload, risk perception, transparency, emotional expressiveness, and task interdependence. Therefore, automation systems that adjust transparency and decision-support levels in real-time are better positioned to sustain appropriate trust levels. By leveraging adaptive transparency mechanisms, and designing automation systems that support high levels of interdependence, military organizations can create great robotic teammates capable of fighting on any battlefield. Such efforts will ensure that our military forces are best positioned to fight and win in a complex world.

Information Assimilation and Belief Formation

Based on Lee and See (2004), users formed trust in automation by drawing on prior knowledge and observations to shape their expectations. This process relied on a mix of analytical reasoning, pattern recognition, and emotional responses to system behavior. Trust developed gradually, influenced by the clarity and consistency of the system's performance and the way it communicated reliability. When automation behaved predictably and conveyed useful feedback, users became more confident in its capabilities over time. On the other hand, inconsistent or unexpected performance made them more cautious and less likely to rely on the system over time (Lee & See, 2004).

The aforementioned trust model did not discuss individual differences between users and how those differences impact their trust in different automation tools. For example, attention control capabilities impacted people's initial trust and propensity to use different levels of automation (Rovira et al., 2017); additionally, the explicit and implicit attitudes people held towards robots shaped trust (Spatola & Wudarczyk, 2021). These attitudes, whether consciously held (explicit) or subconsciously formed (implicit), influenced how operators perceive, engage with, and ultimately depend on automated systems (Elsbach & Stigliani, 2019; Han et al., 2020). Within Lee and See's (2004) model, people's individual differences and attitudes would be classified as an organizational or environmental context that fed into the *Information Assimilation and Belief Formation* box (see Figure 4).

Individual Differences

Working memory has been widely studied in the context of predicting performance and trust in automation. Working memory (WM) referred to a person's ability to temporarily store and manipulate information while engaged in cognitive tasks (Baddeley, 2000). It played a critical role in decision-making, particularly when individuals evaluated automation that provided both accurate and erroneous recommendations. People with higher WM capacity performed better in automation-assisted tasks because they effectively processed and verified information, even when the automation failed (Rovira et al., 2017). Conversely, individuals with lower WM capacity relied more heavily on automation, even when it was incorrect, leading to performance decrements (de Visser et al., 2010). For example, in a simulated military targeting task, people with lower WM struggled identifying errors in an automated targeting system, whereas those with a higher WM were better able to detect automation failures and adjust their decisions accordingly (Rovira et al., 2017).

While working memory has traditionally been used to predict performance in human factors research, and specifically automation, it was distinct from attention control. Working memory involved both the storage and manipulation of information (Baddeley, 2000), while attention control referred to a person's ability to regulate cognitive resources and maintain their focus on task-relevant information (Pak et al., 2023). A key issue with working memory as a predictor was that many measures conflate memory storage with attention control - this made it difficult to determine which cognitive process drove performance differences (Shipstead et al., 2014). As such, WM tasks often fail to isolate the executive attention component, which resulted in ambiguity while interpreting results.

Attention control emerged as a stronger predictor of performance in complex task environments (Durso et al., 2006; Chen & Barnes, 2012). Attention control facilitated goal-directed behavior by allowing people to maintain focus on relevant information while effectively disengaging from distractions (Engle, 2002). Unlike working memory, which was marred by its confounds, attention control measures have demonstrated more predictive power in multi-tasking and automation related tasks (Pak et al., 2023). For example, in an air traffic control simulation, WM capacity did not predict performance, but attention control measures explained the variance in task success, suggesting that the ability to maintain focus was more critical that raw capacity (Durso et al., 2006). Additionally, when people were forced to multitask, attentional control measures - rather than WM capacity - were the best predictors of an operators ability to manage multiple robotic systems in a high workload environment (Chen & Barnes, 2012). Additionally, attention control has been less susceptible to adverse impacts because attention control assessments relied on fundamental cognitive processes, such as the ability to regulate focus and resist distractions rather than acquired knowledge (Burgoyne et al., 2021). These findings suggested that attention control provided a more precise explanation of individual differences in performance, particularly in high workload or distracting environments.

Explicit Attitudes

Explicit attitudes represented conscious beliefs and evaluations about robots, including perceptions of their competence, reliability, and usefulness. Traditionally, individual attitudes towards robots have been investigated via explicit self-reports. Previous research has emphasized the notion of a strong link between people's previous knowledge about robots and associated robotic acceptance (Arras & Cerqui, 2003). When human operators believed a robotic system was highly capable, they reported higher levels of trust and reliance (Young et al., 2009). However, this sometimes led to over-dependence when the system did not perform as expected. On the other hand, negative explicit attitudes, such as skepticism toward a robot's decision-making abilities, often resulted in under-dependence, with operators underutilizing automation even when it performed reliably.

Anthropomorphic design features played a significant role in shaping people's attitudes. Robots with human-like characteristics tended to evoke positive explicit attitudes regarding competence and reliability. However, when these anthropomorphic features did not match a robot's functional capabilities, they often led to misplaced trust and over-dependence, particularly among operators with positive implicit biases toward human-like machines (Haring et al., 2021). These beliefs have been shaped by training, past experiences, and organizational culture, highlighting the need for educational interventions that accurately convey automation capabilities and limitations (Young et al., 2009; Zlotowski et al., 2018).

Implicit Attitudes

Implicit attitudes, by contrast, operate at a subconscious level. They are shaped by unconscious biases that affect split-second decisions in high-pressure situations. De Houwer et al. (2009) argued that implicit attitudes formed automatically and could be measured using techniques like implicit association tests (IATs). Research indicated that these implicit biases, which are less susceptible to deliberate self-perceptions, can be used to evaluate novel objects (Cunningham et al., 2004; de Graaf et al., 2016). These implicit biases—whether positive or negative—significantly impacted trust calibration and reliance decisions, especially under time constraints. For example, Cunningham et al. (2004) found that the amygdala activated during implicit evaluations, suggesting that emotional intensity influenced subconscious trust decisions. Operators with a positive implicit bias often over-relied on automation, bypassing critical verification steps. In contrast, those with negative biases performed redundant manual checks, which slowed decision-making. Crucially, within the domain of human-robot interaction, the implicit measure of attitude towards robots has shown to be an effective predictor of future behavior towards a robot (Spatola & Wudarczyk, 2021).

The interaction between implicit and explicit attitudes complicates the trust-dependence relationship. People's explicit attitudes frequently suggested positive attitudes toward robots, implicit measures uncovered more negative associations (Spatola & Wudarczyk, 2021). This discrepancy between conscious beliefs and subconscious perceptions sometimes led to hesitation, inconsistent reliance behaviors, or trust asymmetries in operational settings. Addressing these gaps requires automation systems capable of providing real-time, context-sensitive feedback that aligns with both conscious expectations and subconscious perceptions. For example, adaptive transparency mechanisms that tailor information delivery based on real-time user engagement could help reconcile these conflicts, promoting consistent reliance patterns.

Reliance Action

Ultimately in Lee and See's (2004) model, users decided whether to rely on automation based on their level of trust and the demands of the situation. Factors like workload, time pressure, and perceived risk all played a role in shaping these decisions. When trust was high, users tended to rely on automation more frequently, sometimes to the point of reduced oversight. When trust was low or uncertain, they were more likely to fall back on manual control, even if automation could have handled the task more efficiently. These patterns of reliance evolved over time, shaped by ongoing feedback from the system and the user's experiences with its performance (Lee & See, 2004).

However, the model did not discuss the relationship between the person and the automation. Task interdependence defined how much, and the type of work, required from both the human and automation (Zhao et al, 2020). Task interdependence plays a pivotal role in shaping trust dynamics within human-automation teams. Within Lee and See's (2004) model, task interdependence would be classified as an organizational or environmental context that fed into the *Reliance Action* box.

Task Interdependence

At its core, task interdependence refers to the degree to which the actions of one team member—whether human or automated—affect the outcomes of others. High levels of reciprocal task interdependence, where human and automated agents rely on each other's actions for successful task completion, fostered higher levels of trust and reduced operator stress (Zhao et al., 2020). Furthermore, Verhagen et al. (2022) showed that trust in human-robot teams was influenced by the type of communication employed by the robotic teammate. Specifically, explainable transparency provided insights into the rationale behind automation decisions, which reduced uncertainty and increased trust in those interdependent tasks. Conversely, in scenarios with lower interdependence, concise and transparent interactions are more beneficial, preventing information overload and maintaining efficiency. This dynamic is essential according to Lee and See's (2004) performance-based trust framework, where trust develops as automation systems consistently demonstrate reliability in these interdependent tasks

The perception of automation as a genuine team member also influenced trust levels, particularly in high-interdependence scenarios. O'Neill et al. (2022) stressed that human–autonomy teaming required autonomous systems to exhibit behaviors that signal competence, reliability, and shared intent. In situations of high interdependence, autonomous agents perceived as possessing higher levels of agency fostered stronger trust relationships, as they are viewed as collaborators rather than mere tools. Automation was more likely to be trusted if it communicated its decision-making processes transparently and aligned its actions with the team's objectives.

High levels of task interdependence between a person and a robot can amplify trust asymmetries, leading to either an over-reliance or underutilization on automation. Over-reliance occurred when operators trusted automation excessively and overlooked critical errors, whereas underutilization stemmed from a lack of confidence in the system's capabilities (Gans & Rogers, 2021). In cooperative multirobot systems, where people relied on multiple autonomous robots, effective systems extended to include distributed autonomous robots, where decisions were made collectively across multiple automated units. Over time, people developed confidence in the team's decision-making capabilities through consistent performance and transparent communication (Gans & Rogers, 2021). Predicting these trust asymmetries to dynamically create the right level of interdependence based on the task and environment would ensure that human-automation networks function cohesively in risk-intensive battlefield scenarios.

Task interdependence also intersects with both learned and situational trust. Trust calibration in interdependent tasks was influenced by the historical performance of automation systems. Prior successes reinforce trust, while past failures necessitated greater transparency and explainability to rebuild confidence (Hoff & Bashir, 2015). Situational trust fluctuated based on contextual factors, which showed situational trust to be sensitive to the nature of task interdependence (Hoff & Bashir, 2015). In high-interdependence scenarios, trust was dynamically calibrated, while automation systems provided adaptive transparency that adjusted to evolving mission requirements. Conversely, in low-interdependence scenarios, streamlined communications sufficed, preserving cognitive bandwidth if the user needed to execute any possible emergent task (Zhao et al., 2020; Verhagen et al., 2022).

In conclusion, task interdependence significantly shaped trust dynamics in human-automation teams. Empirical studies supported that adaptive levels of task interdependence and specific communication strategies enabled human operators to see their automation as robotic teammates, rather than solely tools to be used. The incorporation into being a teammate facilitated learned trust and improved human-automation collaboration in those teams.

Automation

The Lee and See (2004) model captured the effect of the automation itself through its performance, depicted in the model as an Attributional Abstraction under the information about the automation. The performance references the automation's reliability, or producing accurate outputs and responding appropriately to minimize system errors (Lee & See, 2004). When automation operated reliably people developed trust in the system. However, if the system was unpredictable, trust eroded, often leading users to eventually disengage from the automation.

Though the model discusses reliability and how crucial it was for user trust, it did not visually depict reliability as a key driver. Highly reliable systems created trust, decreased people's workload, increased efficiency, and increased performance (Onnash et al., 2015; Djuric et al., 2016; Chen et al., 2018). Also, the model did not precisely include and demonstrate the role that transparency has on trust. Explainable transparency, defined as giving a reasoned explanation for why it chose an action or is giving a certain recommendation, has shown to increase trust without impacting a person's subjective workload (Djuric et al., 2016; O'Neill et al., 2022; Verhagen et al., 2021), The reliability and transparency concepts would be drivers leading to the *Automation* box in Lee and See's (2004) trust model as depicted in Figure 4.

Automation Reliability

Automation reliability was thought to be the most important driver of a person's trust in automation. Meaning that when automation operated as expected on simple tasks, people trusted the automation's capabilities more, which encouraged operators to depend on it for more complex tasks (Dixon et al., 2007; Hutchinson et al., 2023; Lee & See, 2004). Yet, this trust between operator and automation remained delicate. Even minor discrepancies, such as unexpected behavior or inconsistent performance eroded trust, which compelled operators to reassert manual control and increased their cognitive workload (Endsley & Kiris, 1995). Eroded trust and unnecessarily increased cognitive workload in high risk environments was a potentially dangerous combination. A person's cognitive resources and oversight should be reserved for conditions that may determine mission success or failure (van de Merwe et al., 2024).

An operator's expectations of automation largely depended on its history of reliable performance. When automation consistently meets the expectations of the operator, trust strengthens; however, a single failure can destabilize trust (Endsley, 2017). Different degrees of automation reliability exerted varied effects on operator trust. Low reliability, typically around 60%, led to swift distrust, driving operators to default to manual control (Chavaillaz et al., 2016). Moderate reliability, typically seen between 70% and 80%, presented more nuanced challenges - while it fostered some trust, minor errors generated inconsistent reliance (Onnash, 2015; Guznov et al., 2020). Consistently high reliability, typically at 80% or above, fostered stable trust (Onnasch, 2015). In manufacturing, collaborative robots that performed reliably enabled workers to reduce both physical and cognitive workload, ultimately increasing productivity (Djuric et al., 2016). Likewise, in simulated military operations, dependable automation provided accurate and timely decision-making support, alleviated cognitive overload, and improved mission performance (Chen et al., 2018).

Yet, this same high reliability that increased performance also fostered complacency in people, as they became over-reliant on automation (Selkowitz et al., 2017). Users relied too heavily on automation and failed to monitor system outputs effectively, assuming that the technology would operate flawlessly (Parasuraman & Manzey, 2010). People with high autonomous systems tended to disengage from monitoring tasks, which increased errors when automation encountered unexpected situations (Gutzwiller et al., 2019). To fight against this complacency, adaptive systems incorporated periodic user-engagement prompts designed to sustain operator vigilance and prevent skill atrophy (Bhaskara et al., 2021; Stowers et al., 2021).

Automation complacency also impaired performance in collaborative, high-stakes environments. In military applications, for example, people working alongside autonomous systems exhibited reduced engagement over time, and deferred critical judgment to the system even when errors were evident (Endsley & Kaber, 1999). Similarly, in autonomous vehicle research, drivers reacted slower to system failures when they relied excessively on driver-assistance systems (Carsten & Martens, 2019). These findings reinforced the importance of designing automation with appropriate trust regulation strategies, such as real-time feedback, system transparency, and adaptive autonomy, to mitigate the risks associated with automation complacency.

Automation Transparency

Trust in automation doesn’t hinge solely on performance and the automation’s reliability. Transparency communicated the intent, performance, future plans, and reasoning process to the user in a clear and understandable presentation (Chen et al., 2014). Transparent explanations from the automation enhanced trust by aligning operators’ expectations with system capabilities. For example, Mercado et al. (2016) demonstrated explainable transparency to users forced to decide a tactical plan based on an integrated graphic with potential obstacles based on equipment capabilities, text explaining how the plan met the intent of the mission, and uncertainty information. Explainable transparency helped users understand the purpose of the automation, what it was attempting to accomplish, and potential uncertainty in the decision. Also, when systems emphasized their limitations, they can soften the blow of reliability failures. For instance, if a system makes an error but explains why it happened and what’s being done to fix it, operators are more likely to give it another chance (Van de Merwe et al., 2024).

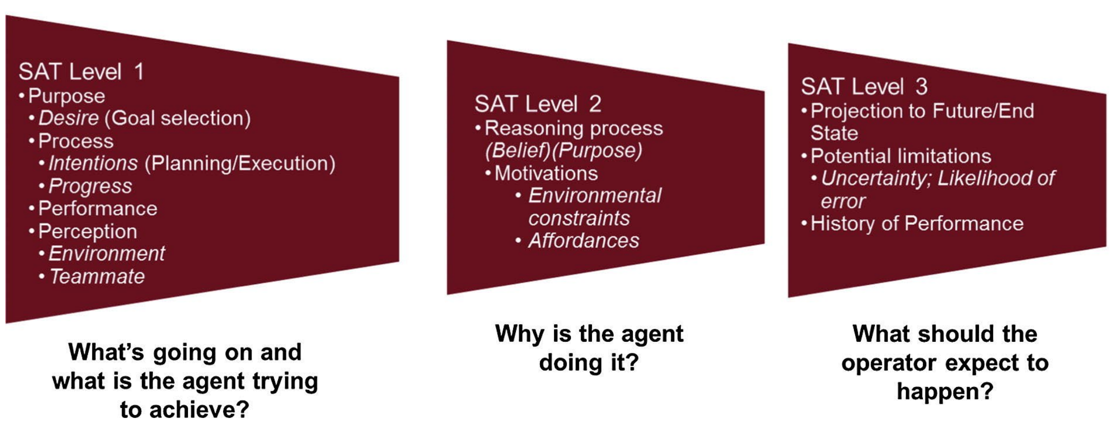

Figure 2. A model of transparency from Chen et al. (2014) based on the purpose, process, and projection framework provided by Lee and See (2004). The level of transparency increased as the automation explained what it was doing, why it was doing it with further details to the user on its reasoning, and provided future projections.

As seen above in Figure 2, Chen et al. (2014) parsed transparency into three levels based on Lee and See's (2004) model: (1) purpose, where the automation communicated its purpose; (2) reasoning process, which detailed why the automation was making decisions; (3) projection to future state, which communicated what potentially could happen. Van de Merwe et al.'s (2024) meta-analysis reinforced that adaptive transparency proved beneficial in low-workload conditions, while targeted information prevented cognitive overload in users. When explainable transparency was integrated into system designs, cognitive workload decreased by streamlining information processing and minimizing the need for manual verification (Mercado et al., 2016; Selkowitz et al., 2015; Chen et al., 2018).

Transparency is a multifaceted construct, representing the extent to which an automated system communicates its processes, reasoning, and limitations to the user. Transparency allows operators to understand not only what a system is doing but also why it is performing specific actions (Verhagen et al., 2022). This transparency is especially critical in military environments, where context changes quickly and is best understood by the commander on the ground. However, transparency exists on a spectrum, ranging from basic transparency to explainable transparency. Basic transparency is where a display provides all the data to the user for them to interpret system outputs; whereas, in explainable transparency, the automation provides analyzed data, a recommendation, and reasoning for that recommendation. As Verhagen et al. (2022) noted, explainable transparency helps users interpret a system's behavior, predict its actions, and maintain trust in the robotic teammate.

Explainable transparency can reduce the cognitive burden associated with decision verification in high risk scenarios. Unlike basic transparency, which presents surface-level outputs, explainable transparency provides detailed insights into the confidence and reasoning behind the automation's decision (Loft et al., 2023; Verhagen et al., 2021; Bashkara et al., 2021). Explainable transparency involved a step-by-step breakdown of target identification decisions, associated risk levels, and contingency plans in response to potential enemy actions. Such comprehensive explanations empowered users to balance their trust in the automation with the risk of the situation, similar to what a soldier might have to do (Loft et al., 2023). These adaptive transparency models are needed to best support the user - they tailor the information based on the environment, risk, and the commander's workload at the time. Embedded explainable transparency reconnaissance systems would ensure that commanders can assess mission risk without incurring excessive cognitive workload.

Verhagen et al. (2022) demonstrated that explainable communication significantly improved trust in highly interdependent tasks. Explainable transparency ensured that the entire team understood the rationale behind decisions, and therefore fostered coordinated action. This capability was particularly critical in cooperative multirobot systems, where distributed decision-making requires a shared mental model to ensure operational cohesion (Lyons & Stokes, 2012).

Explainable transparency facilitated this shared understanding and supported trust distribution across multiple automated agents. Also, the automation's reasoned explanations increased user's performance on the task without impacting their subjective workload (Mercado et al., 2016; Chen et al., 2018). The presented information helped the user without any discernible performance downside.

Ultimately, explainable transparency provided the right information in the right format that matched the demands of the environment. In low-risk scenarios, minimal transparency may have allowed operators to focus on broader mission objectives; however, people in high-risk scenarios wanted detailed explanations that allow them to double check decisions without compromising decision timelines (Mercado et al., 2016). Future research should explore how machine learning algorithms could predict what information military commanders would need in real-time. Having systems capable of providing the right amount of explainable transparency at the right time would foster trust and improve operational outcomes in the military.

Display

Lee and See (2004) discussed the impacts of information presentation on how users perceived and trusted automation. A well-designed display helped users understand system capabilities by providing new information, relevant feedback about past performance, and overall reliability. Trust suffered if these were poorly communicated, leading to either over-reliance or unnecessary skepticism (Lee & See, 2004). The best displays balanced abstraction with detail, giving users enough information to make informed decisions without overwhelming them.

While the presentation of information is important, the model did not take into account the emotional expressions of the automation and if it had anthropomorphic qualities. People performed best when executing complex tasks with highly anthropomorphic robots (Hinds et al., 2004) and military personnel demonstrated greater trust when robots' physical actions appeared purposeful and communicated emotional context clues (Cappuccio et al., 2021). Lee and See's trust model did not represent these drivers well, but both emotional expressions and anthropomorphism would appear as influencing the *Display* box (shown in Figure 4).

Emotional Expressions

Although not explored as often as transparency or reliability within automation research, emotional expressiveness has demonstrated a profound influence. Emotionally expressive robots facilitated greater trust and reliance, particularly in high-risk decision-making environments where human and automated recommendations conflict (Feng et al., 2019). Emotional expressions provided participants social cues that helped them interpret the intentions, reliability, and urgency of the automation. Research has shown that these cues, given in conjunction with the recommendation itself, supported more effective trust calibration between the two. Ho et al. (2017) highlighted the impacts of personal experiences, culture, and contextual risk. They found that while pilots initially hesitated, trust in the automation increased over time as the system demonstrated reliability and effectiveness in high-risk scenarios. This study showed that emotional expressiveness—whether from human operators or integrated into automation—accelerated trust calibration when aligned with real-world operational demands.

Robots with explainable transparency using emotional expressiveness when communicating enhanced trust with its human partner. Feng et al. (2019) suggested that emotional cues provided contextual signals that complement transparency by indicating confidence levels, operational urgency, or system uncertainty. For example, in potential combat operations, an automated system explaining its targeting process while simultaneously expressing appropriate levels of urgency in its recommendations more effectively earned people's trust. This combination of logical explanations and affective assurances ensures that operators are better equipped to interpret and rely on automation decisions (Ho et al., 2017; Feng et al., 2019). It also highlighted the importance of dynamic emotional adaptation in future robotic teammates, where systems modulate emotional expressiveness based on risk levels, task complexity, and operator workload. Over time, the role of emotional cues diminished, giving way to performance-based trust (Ho et al., 2017). In military settings, where uncertainty and time-sensitive decisions are prevalent, integrating emotional expressiveness into both human-agent communication and automation interfaces may help regulate trust dynamics, ensuring neither over-reliance nor undue skepticism occurs.

Anthropomorphism

Integrating emotional expressiveness into human-automation interaction models offered a promising avenue for anthropomorphic robots. Designers manipulated these physical attributes to make robots more or less human-like, or in some cases, to make a robot look like a real dog. Anthropomorphism, however, was a double-edged sword—while complex collaborative tasks were performed more efficiently with highly anthropomorphic robots, robots that appeared too human-like triggered the uncanny valley phenomenon (Hinds et al., 2004; Mori et al., 2012; Kim et al., 2022). The uncanny valley referred to the discomfort or distrust users felt when robots appear too human-like. Therefore, creating a design with the right level of anthropomorphism posed a critical challenge for robotic designers. Designers wanted soldiers to trust and feel comfortable with their automation. Boston Dynamics sought to avoid the uncanny valley problem by creating an animaloid robot which roughly resembled a dog. For example, Spot's dog-like form (seen in Figure 3) may have engendered trust from its human operator, and allowed operators to intuitively gauge its reconnaissance capabilities.

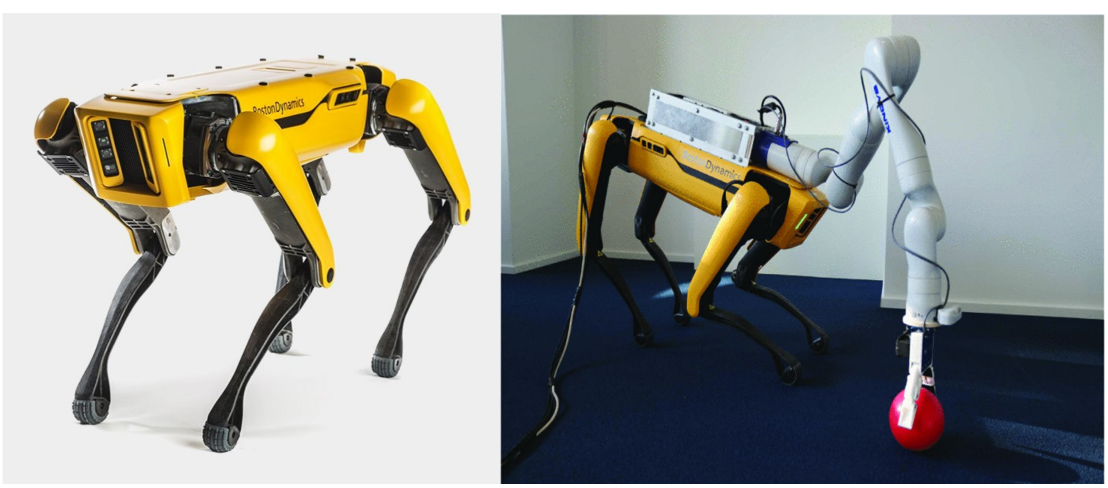

Figure 3. Boston Dynamics' Spot the robot dog resembled the overall look and movement of a physical dog. Spot's abilities were highly customizable based on the mission assigned. Spot conducted physical or map reconnaissance (Vom Hofe et al., 2023), conducted RADAR operations (Wetzel et al., 2022), or be affixed with tools like an external arm (Zimmerman et al., 2021).

However, if the robot's movements or behaviors were too lifelike without matching cognitive functions, it could trigger unease, undermining trust. MacDorman et al. (2009) emphasized that such aesthetic-functional mismatches impact both explicit and implicit attitudes. Crucially, these attitudes potentially impacted users' trust in the automation as well their future behavior with the robot (Spatola & Wudarczyk, 2021).

Military robots required simple, socially interactive features that aligned with their functional roles (Cappuccio et al., 2021; Djuric et al., 2016). For example, a robot like Spot, designed for reconnaissance missions, leveraged dog-like physical attributes that influence how users perceive its capabilities. These physical cues prompted assumptions about the robot's agility, sensor capabilities, and operational reliability. However, these assumptions may not always reflect the robot's actual functionality, leading to over-reliance based on misperceptions. The form-function attribution bias described by Haring et al. (2018) illustrated this issue, where operators incorrectly inferred advanced competencies based solely on appearance. Thus, robotic form should clearly showcase its functionality, ensuring that trust levels are appropriately calibrated and aligned with actual system performance.

Robotic teammates with an appropriate level of anthropomorphic features have shown to be especially effective on human-automation teams executing highly interdependent tasks. Human-likeness and physical embodiment in automation systems can enhance trust, particularly in high-pressure environments where rapid trust calibration is essential (Haring et al., 2021). In high-interdependence military tasks, where coordinated action is essential, moderate anthropomorphic cues fostered trust and effective collaboration. However, Rovira et al. (2024) cautioned that individual differences in pre-existing attitudes toward robots can either amplify or diminish the positive effects of anthropomorphism. Operators with negative predispositions may exhibit lower trust levels despite functional reliability, whereas those with positive biases may over-rely, even when system performance does not warrant it. These dynamics underscored the need for customizable anthropomorphic features, allowing operators to adjust robotic behaviors and appearances according to personal comfort levels and operational requirements.

In conclusion, anthropomorphism and emotional expressiveness represent a component of trust adaptation in human-automation teams. When combined with explainable transparency and adaptive communication strategies, these offer a more holistic approach to fostering robust trust dynamics capable of withstanding the complexities of modern military operations.

Revised Theory of Trust

Since the initial trust theory was published, empirical research has continued to define and refine its constructs. As such, some of the drivers of trust calibration discussed above were not explicitly mentioned by Lee and See (2004). Based on the literature reviewed previously, there are some important characteristics of both the form and function of automation that govern a person's trust. Below in Figure 4 is an updated version of the original model with additional inputs added to the *Information assimilation and belief formation, Reliance, Automation, and Display* boxes of the flowchart. A greater understanding of the person (attitudes, attentional control), the task (task interdependence), how the automation works (reliability, transparency), and how it looks and communicates (emotional expressions, anthropomorphism) will refine our trust framework across all fields of study.

Eventually, robots will conceptually change from just a tool to be used towards being a robotic teammate that helps accomplish the mission. As they make this transformation, the points of interaction between the robotic teammate and the human operator becomes that much more important. Creating the right robot that soldiers would embrace within their formation and trust on the battlefield through how it looks, acts, and communicates gives our military a significant advantage. Additionally, the ability to dynamically govern task interdependence and explainable transparency enables the automation to understand the environmental context and best support a military commander on the ground.

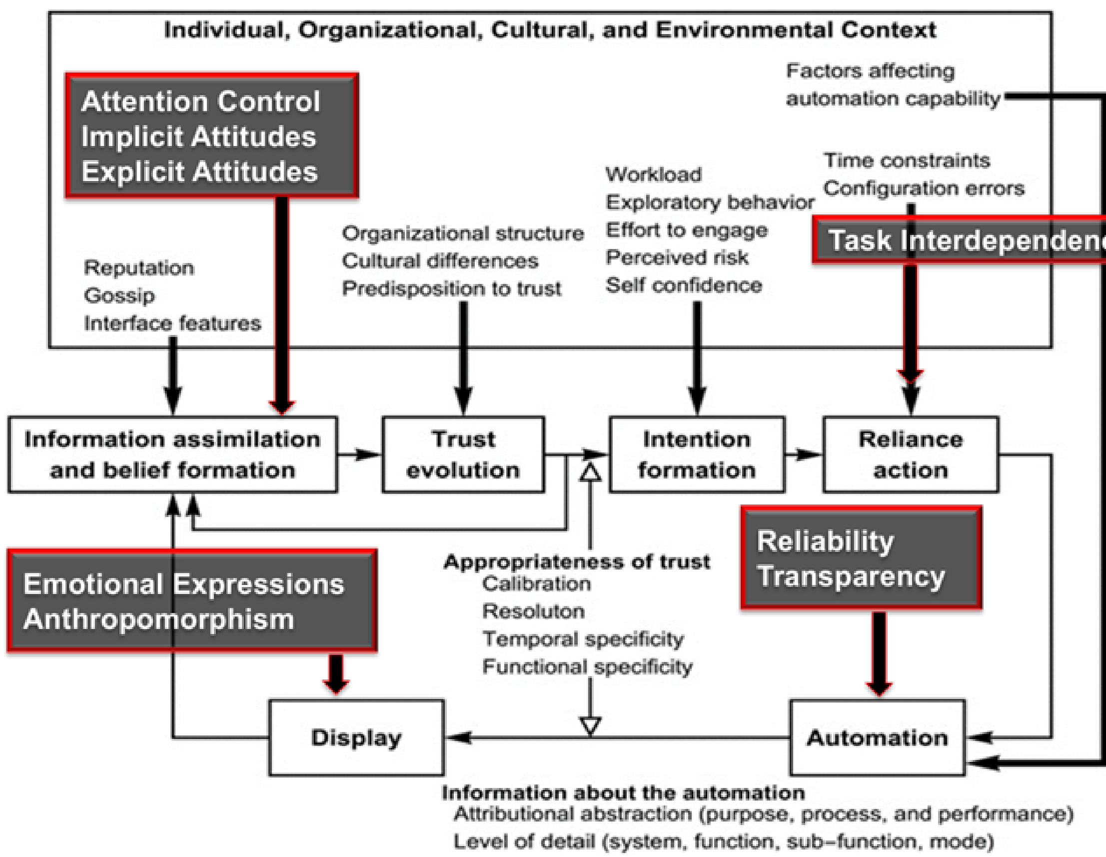

Figure 4. A revised version of Lee & See's (2004) conceptual model of the dynamic process that governs trust. The additional drivers of trust have been added throughout the model, in line with how they impact the human or the automation.

Dependence: A Critical Gap in the Literature

Despite substantial research on trust in human-automation interactions, a notable gap remained in understanding automation dependence. Automation dependence was not a direct output from trust, instead it was a specific behavior shaped by the situation, environmental demands, and perceived system performance (Parasuraman & Riley, 1997). To fully grasp how reliance on automation influenced mission success and operator efficiency, it was essential to explore this nuanced relationship.

Automation Dependence

Dependence was a behavioral measure used to assess how often individuals utilized automation when it was available. For alerting systems or diagnostic tools, dependence was often broken down into reliance and compliance (Meyer, 2004). Reliance occurred when a person accepted an "all is well" signal from the automation. This behavior was influenced by automation; when the system failed to signal real issues, operators tended to rely on it less (Dixon et al., 2007). Reliance was also closely linked to complacency, which happened when an operator over-relied on the automation (Parasuraman & Manzey, 2010). Compliance, on the other hand, referred to when an operator followed an alert from the automation, even if the alert did not match the true state of the situation. Compliance generally decreased with false alarm-prone automation, as frequent inaccurate alerts caused operators to delay their responses or ignore alerts altogether. Dependence also applied to non-alerting systems, where it was measured by how often the system was used (Meyer, 2004).

Automation dependence and trust were intuitively related, but not the same thing. In a recent metaanalysis, trust and dependence only had a correlation of $r = .13$ (Patton & Wickens, 2024). While trust determined whether an operator was willing to rely on automation, dependence reflected actual reliance behavior, which did not always align with trust levels (Dzindolet et al., 2003). As covered previously, over-dependence led to automation bias, where users blindly trusted the automation, potentially missing crucial errors. Conversely, under-dependence forced operators to assume excessive manual tasks, resulting in cognitive overload and an inefficient use of the automation (Parasuraman & Riley, 1997). Perhaps these automation systems were not able to dynamically adapt to the changing environment or they were unable to communicate to the human operator.

Situational factors such as workload, risk perception, and time constraints played significant roles in shaping automation dependence. Parasuraman et al. (2000) observed that during high-workload scenarios, operators were inclined to delegate decisions to automation without thorough validation, risking over-reliance. In high-risk conditions, even highly reliable systems were underutilized as operators became overly cautious, fearing potential failures which led to consistently verifying the automation's decision (Bashkara et al., 2021; Muir & Moray, 1996). These clearly show a need for adaptive transparency and context-aware explainability, so operators could assess the trustworthiness of automated systems in real-time.

Task interdependence also influenced automation dependence. In cooperative multirobot systems, operators needed to trust not only individual robots but also the collective decision-making of distributed networks (Gans & Rogers, 2021). In these cooperative tasks, the level of interdependence was crucial. In highly interdependent tasks, operators were more likely to rely on automation—provided the system had demonstrated consistent performance and communicated decision-making processes clearly (Zhao et al., 2020). However, a single failure within a multirobot system diminished trust in the entire network, resulting in widespread under-dependence (Riley, 1995). Based on these, the incorporation of dynamic explainable transparency mechanisms to explain why automation teammates made their decision would restore trust while maintaining the required operational tempo.

Ongoing research must investigate how machine learning algorithms can predict dependence patterns in real time, enabling automated systems to adjust transparency levels and operational autonomy accordingly (Parasuraman & Riley, 1997; Madhavan & Wiegmann, 2007). As robotic teammates gain technological capabilities and are mass fielded in the Army, research will continue to bridge the gap between the attitude that is trust and the behavior that dependence represents.

Measures of Automation Dependence

Accurately measuring trust in automation and dependence behaviors remained essential for understanding how human-automation teams operate. Trust was “the attitude that an agent will help achieve an individual’s goals in a situation characterized by uncertainty and vulnerability” (Lee & See, 2004). Dependence, in contrast, referred to the actual behavioral reliance demonstrated during interactions with automation. It included reliance and compliance (Meyer, 2004). Therefore, when compared to self-report measures, behaviorally based measures provided clearer insights on how people depended on automation.

Self-report surveys and Likert-scale questionnaires served as foundational tools for assessing trust attitudes. Instruments such as the Trust in Automation Scale (Jian et al., 2000) remained widely adopted due to their simplicity and ease of use. Studies relied on self-report methods to gauge participants' perceived trust levels after interacting with automated systems. These measures provided direct insights into users’ subjective trust evaluations and can be administered repeatedly to track trust development over time. However, trust is dynamic and changes across time (Yang et al., 2021), but since these self-report measures typically represent a single moment in time, they were unable to capture real-time trust fluctuations during the mission. Holistically, these self-report surveys measured peoples intent to rely - or intent to depend - rather than dependence behaviors. Given the weak correlation between trust and dependence (Patton & Wickens, 2024), there was a gap not addressed by the majority of empirical literature.

Behavioral measures offered a more objective assessment of dependence by tracking how users interacted with automation and incorporated both behavioral and physiological measures. Metrics such as frequency of system use, override rates, and response times indicated when and the extent to which operators rely on automated systems. Override rates, communication patterns between the user and the automation, and reliance on the automation's suggestions all provide specific information on the interaction between the person and the automation (Pak et al., 2012; Zhao et al., 2020; Verhagen et al., 2022). For instance, Pak et al. (2012) investigated checking raw data as an opposite behavior to compliance, and reinforced the idea that automation dependence varied based on user trust and confidence in system accuracy. Additionally, eye-tracking metrics such as fixation frequency and duration were used to infer dependence, with more frequent glances at automation being interpreted as lower dependence. For instance, Dixon et al. (2007) found that automation with frequent misses led to more operator glances at the automated task, suggesting decreased reliance. These measures reflected actual user behaviors while using the automation, minimizing self-report biases or people guessing their own intent to rely on the automation.

Interpreting these dependence behaviors was not always straightforward. High usage rates might have suggested an appropriate level of reliance or, conversely, it might have meant complacency and over-dependence. Similarly, frequent overrides may have indicated healthy skepticism or a lack of trust that undermines performance. The context of the task, associated risks, and operator workload were all considered when interpreting behavioral data (Parasuraman & Riley, 1997. So, despite the variety of available measures, no single approach comprehensively captured the multifaceted nature of trust and dependence.

Conclusion and Way Forward

Trust in automation is not a fixed state, rather it shifts and evolves based on experience, context, and system performance. In military operations, where the stakes are high, trust must be carefully calibrated. Operators need confidence that automation will perform reliably, make sound decisions, and ultimately enhance mission effectiveness rather than complicate it. But trust isn't just about reliability; it's shaped by transparency, emotional expressiveness, and the level of interdependence between human and machine. Robotic teammates must do more than function correctly—they must communicate their reasoning, adapt to changing conditions, and earn their place within a team. As automation grows more advanced, its ability to provide clear explanations, adjust to operator needs, and align with mission goals will be crucial for fostering lasting trust.

Looking ahead, the challenge with trust is ensuring that trust leads to appropriate reliance. Trust is an attitude, but dependence is a behavior, and the two rarely match up. In high-risk, high-pressure environments, operators may either lean too heavily on automation or hesitate to use it at all, both of which can undermine performance. To bridge this gap, future systems must recognize when and how operators are relying on them through experiments with increased ecological validity. Within these scenarios, people would make situational judgements based on an autonomous teammate's recommendation (Manniste et al., 2019); this would provide clear insights on how users would depend on automation. As these technologies become more embedded in military operations, ongoing research must explore how to establish, maintain, and rebuild trust when failures occur. The goal is not just to create functional automation, but to build systems that soldiers trust, depend on, and ultimately see as partners in mission success.

References

Arras, K. O., & Cerqui, D. (2005). *Do we want to share our lives and bodies with robots? A 2000-people survey* (Report No. 0605-001). Lausanne, Switzerland: Autonomous Systems Laboratory.

Baddeley, A. (2000). The episodic buffer: a new component of working memory?. *Trends in cognitive sciences*, *4*(11), 417-423.

Bhaskara, A., Duong, L., Brooks, J., Li, R., McInerney, R., Skinner, M., & Loft, S. (2021). Effect of automation transparency in the management of multiple unmanned vehicles. *Applied Ergonomics, 90*, 103243.

Burgoyne, A. P., Mashburn, C. A., & Engle, R. W. (2021). Reducing adverse impact in high-stakes testing. *Intelligence, 89*, 101589.

Cappuccio, M. L., Galliott, J. C., & Sandoval, E. B. (2021). Saving private robot: risks and advantages of anthropomorphism in agent-soldier teams. *International Journal of Social Robotics, 14*(10), 2135-2148.

Carsten, O., & Martens, M. H. (2019). How can humans understand their automated cars? HMI principles, problems, and solutions. *Cognition, Technology & Work, 21*(1), 3-20.

Chavaillaz, A., Wastell, D., & Sauer, J. (2016). System reliability, performance, and trust in adaptable automation. *Applied Ergonomics, 52*, 333-342.

Chen, J. Y. C., & Barnes, M. J. (2012). Supervisory control of multiple robots: Effects of automation and decision aid reliability. *Human Factors*, *54*(2), 157–174.

Chen, J. Y., Procci, K., Boyce, M., Wright, J., Garcia, A., & Barnes, M. (2014). *Situation awareness-based agent transparency* (Report No. ARL-TR-6905). Aberdeen Proving Ground, MD: U.S. Army Research Laboratory.

Chen, J. Y. C., Lakhmani, S. G., Stowers, K., Selkowitz, A. R., Wright, J. L., & Barnes, M. (2018). Situation awareness-based agent transparency and human-autonomy teaming effectiveness. *Theoretical Issues in Ergonomics Science, 19*, 259–282.

Cunningham, W. A., Raye, C. L., & Johnson, M. K. (2004). Implicit and explicit evaluation: fMRI correlates of valence, emotional intensity, and control in the processing of attitudes. *Journal of Cognitive Neuroscience, 16*(10), 1717-1729.

de Graaf, M. M., Allouch, S. B., & Lutfi, S. (2016). What are people's associations of domestic robots? Comparing implicit and explicit measures. In *2016 25th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*, (pp. 1077-1083).

De Houwer, J., Teige-Mocigemba, S., Spruyt, A., & Moors, A. (2009). Implicit measures: a normative analysis and review. *Psychological Bulletin, 135*(3), 347-368.

de Visser, E. J., Peeters, M. M. M., Jung, M. F., Kohn, S., Shaw, T. H., Pak, R., & Neerincx, M. A. (2010). Towards a theory of human–robot trust: the role of cognitive ability and automation reliability. *Human Factors*, *52*(3), 335–345.
Dixon, S. R., Wickens, C. D., & McCarley, J. S. (2007). On the independence of compliance and
reliance: are automation false alarms worse than misses?. *Human Factors, 49(4),* 564-572.

Djuric, A. M., Urbanic, R. J., & Rickli, J. L. (2016). A framework for collaborative robot (CoBot) integration in advanced manufacturing systems. *SAE International Journal of Materials and Manufacturing, 9*(2), 457-464.

Durso, F. T., Bleckley, M. K., & Dattel, A. R. (2006). Does working memory load affect visual vigilance? *Human Factors*, *48*(1), 77–91.

Dzindolet, M. T., Peterson, S. A., Pomranky, R. A., Pierce, L. G., & Beck, H. P. (2003). The role of trust in automation reliance. *International Journal of Human-Computer Studies, 58*(6), 697-718.

Elsbach, K. D., & Stigliani, I. (2019). New information technology and implicit bias. *Academy of Management Perspectives, 33*(2), 185-206.

Endsley, M. R. (2015). *Autonomous horizons: System autonomy in the Air Force: A path to the future* (Report No. 2015-0267). Washington, DC: U.S. Air Force Office of the Chief Scientist.

Endsley, M. R. (2017). From here to autonomy: lessons learned from human–automation research. *Human Factors, 59*(1), 5-27.

Endsley, M. R., & Kaber, D. B. (1999). Level of automation effects on performance, situation awareness and workload in a dynamic control task. *Ergonomics*, *42*(3), 462-492.

Endsley, M. R., & Kiris, E. O. (1995). The out-of-the-loop performance problem and level of control in automation. *Human Factors, 37*(2), 381-394.

Engle, R. W. (2002). Working memory capacity as executive attention. *Current Directions in Psychological Science*, *11*(1), 19–23.

Feng, J., Sanchez, J., Sall, R., Lyons, J. B., & Nam, C. S. (2019). Emotional expressions facilitate human–human trust when using automation in high-risk situations. *Military Psychology, 31*(4), 292-305.

Gans, N. R., & Rogers, J. G. (2021). Cooperative multirobot systems for military applications. *Current Robotics Reports*, *2*, 105-111.

Gutzwiller, R. S., Fugate, S. J., Sawyer, B. D., & Hancock, P. A. (2019). The human factors of trust and attention in autonomous systems. *Human Factors, 61*(4), 614-624.

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

Han, B., Wang, L., & Li, X. (2020). To collaborate or serve? Effects of anthropomorphized brand roles and implicit theories on consumer responses. *Cornell Hospitality Quarterly, 61*(1), 53-67.

Haring, K. S., Watanabe, K., Velonaki, M., Tossell, C. C., & Finomore, V. (2018). FFAB—the form function attribution bias in human–robot interaction. *IEEE Transactions on Cognitive and Developmental Systems, 10*(4), 843-851.

Haring, K. S., Satterfield, K. M., Tossell, C. C., De Visser, E. J., Lyons, J. R., Mancuso, V. F., & Funke, G. J. (2021). Robot authority in human-robot teaming: effects of human-likeness and physical embodiment on compliance. *Frontiers in Psychology, 12*, 625713.

Hinds, P. J., Roberts, T. L., & Jones, H. (2004). Whose job is it anyway? A study of human-robot interaction in a collaborative task. *Human–Computer Interaction, 19*(1-2), 151-181.

Ho, N. T., Sadler, G. G., Hoffmann, L. C., Lyons, J. B., & Johnson, W. W. (2017). Trust of a military automated system in an operational context. *Military Psychology, 29*(6), 524-541.

Hoff, K. A., & Bashir, M. (2015). Trust in automation: Integrating empirical evidence on factors that influence trust. *Human Factors, 57*(3), 407-434.

Hutchinson, J., Strickland, L., Farrell, S., & Loft, S. (2023). The perception of automation reliability and acceptance of automated advice. *Human Factors, 65 (8),* 1596–1612.

Jian, J. Y., Bisantz, A. M., & Drury, C. G. (2000). Foundations for an empirically determined scale of trust in automated systems. *International Journal of Cognitive Ergonomics, 4*(1), 53-71.

Kim, B., de Visser, E., & Phillips, E. (2022). Two uncanny valleys: re-evaluating the uncanny valley across the full spectrum of real-world human-like robots. *Computers in Human Behavior, 135*, 107340.

Lee, J. D., & See, K. A. (2004). Trust in automation: designing for appropriate reliance. *Human Factors, 46*(1), 50-80.

Loft, S., Bhaskara, A., Lock, B. A., Skinner, M., Brooks, J., Li, R., & Bell, J. (2023). The impact of transparency and decision risk on human–automation teaming outcomes. *Human Factors, 65*(5), 846-861.

Lyons, J. B., & Stokes, C. K. (2012). Human–human reliance in the context of automation. *Human Factors, 54*(1), 112-121.

MacDorman, K. F., Vasudevan, S. K., & Ho, C. C. (2009). Does Japan really have robot mania? Comparing attitudes by implicit and explicit measures. *AI & Society, 23*, 485-510.

Madhavan, P., & Wiegmann, D. A. (2007). Similarities and differences between human–human and human–automation trust: an integrative review. *Theoretical Issues in Ergonomics Science, 8*(4), 277-301.

Manniste, T., Pedaste, M., & Schimanski R. (2019). Situational judgement test for measuring military tactical decision-making skills. *Military Psychology, 31:6,* 462-473.

Mercado, J. E., Rupp, M. A., Chen, J. Y. C., Barnes, M. J., Barber, D., & Procci, K. (2016). Intelligent agent transparency in human–agent teaming for multi-UXV management. *Human Factors, 58*, 401–415.

Merritt, S. M., Heimbaugh, H., LaChapell, J., & Lee, D. (2015). I trust it, but I don't know why: effects of implicit attitudes toward automation on trust in an automated system. *Human Factors, 57*(5), 760-773.

Meyer, J. (2004). Conceptual issues in the study of dynamic hazard warnings. *Human Factors, 46*(2), 196-204.

Mori, M., MacDorman, K. F., & Kageki, N. (2012). The uncanny valley [from the field]. *IEEE Robotics & Automation Magazine, 19*(2), 98-100.

Muir, B. M., & Moray, N. (1996). Trust in automation. Part II. Experimental studies of trust and human intervention in a process control simulation. *Ergonomics, 39*(3), 429-460.

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

Onnasch, L. (2015). Crossing the boundaries of automation—function allocation and reliability. *International Journal of Human-Computer Studies, 76*, 12–21.

OpenAI. (2025). ChatGPT. (Feb 26). [Large language model]. https://chatgpt.com/share/67c0b5df-f3e0-8006-89aa-227d8fcc57c1

Pak, R., Fink, N., Price, M., Bass, B., & Sturre, L. (2012). Decision support aids with anthropomorphic characteristics influence trust and performance in younger and older adults. *Ergonomics, 55(9),* 1059-1072.

Pak, R., McLaughlin, A. C., & Engle, R. W. (2023). The relevance of attention control, not working memory, in human factors. *Human Factors*, *66*(5), 1321–1332.

Parasuraman, R., & Manzey, D. H. (2010). Complacency and bias in human use of automation: an attentional integration. *Human Factors, 52*(3), 381-410.

Parasuraman, R., & Riley, V. (1997). Humans and automation: use, misuse, disuse, abuse. *Human Factors, 39*(2), 230-253.

Parasuraman, R., Sheridan, T. B., & Wickens, C. D. (2000). A model for types and levels of human interaction with automation. *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans, 30*(3), 286-297.

Patton, C. E., & Wickens, C. D. (2024). The relationship of trust and dependence. *Ergonomics, 1*-17.

Riley, V. (1995). Issues associated with operator use of automation. *SAE Transactions, 104*, 1393-1397.

Rovira, E., McLaughlin, A. C., Comer, N., Rose, K. M., McDonald, M. D., Larkin, D., & Novitzky, M. (2024). I didn't like you before, and I like you less now: individual differences predict trust and behavior in HRI. *SSRN*.

Rovira, E., Pak, R., & McLaughlin, A. (2017). Effects of individual differences in working memory on performance and trust with various degrees of automation. *Theoretical Issues in Ergonomics Science*, *18*(6), 573–591.

Sato, T., Yamani, Y., Liechty, M., & Chancey, E. T. (2020). Automation trust increases under high-workload multitasking scenarios involving risk. *Cognition, Technology & Work, 22*, 399-407.

Selkowitz, A., Lakhmani, S., Chen, J. Y., & Boyce, M. (2015). The effects of agent transparency on human interaction with an autonomous robotic agent. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 59*(1), 806-810.

Selkowitz, A. R., Lakhmani, S. G., & Chen, J. Y. (2017). Using agent transparency to support situation awareness of the autonomous squad member. *Cognitive Systems Research, 46*, 13–25.

Shipstead, Z., Harrison, T. L., & Engle, R. W. (2014). Working memory capacity and fluid intelligence: Maintenance and disengagement. *Perspectives on Psychological Science*, *9*(1), 22–43.

Spatola, N., & Wudarczyk, O. A. (2021). Implicit attitudes towards robots predict explicit attitudes, semantic distance between robots and humans, anthropomorphism, and prosocial behavior: from attitudes to human–robot interaction. *International Journal of Social Robotics, 13*(5), 1149-1159.

Stowers, K., Brady, L. L., MacLellan, C., Wohleber, R., & Salas, E. (2021). Improving teamwork competencies in human-machine teams: perspectives from team science. *Frontiers in Psychology, 12*, 590290.

Szegedi, P., Koronvary, P., & Bekesi, B. (2017). The use of robots in military operations. *Scientific Research & Education in the Air Force-AFASES, 1*, 1-8.

U.S. Department of Defense. (2023, November 3). Deputy Secretary of Defense Kathleen Hicks announces publication of data, analytics and AI adoption strategy [Press release]. https://www.defense.gov/News/Releases/Release/Article/3577857/deputy-secretary-of-defense-kathleen-hicks-announces-publication-of-data-analyt/

van de Merwe, K., Mallam, S., & Nazir, S. (2024). Agent transparency, situation awareness, mental workload, and operator performance: a systematic literature review. *Human Factors, 66*(1), 180-208.

Verhagen, R. S., Neerincx, M. A., & Tielman, M. L. (2021). A two-dimensional explanation framework to classify AI as incomprehensible, interpretable, or understandable. In *International Workshop on Explainable, Transparent Autonomous Agents and Multi-Agent Systems* (pp. 119-138). Springer.

Verhagen, R. S., Neerincx, M. A., & Tielman, M. L. (2022). The influence of interdependence and a transparent or explainable communication style on human-robot teamwork. *Frontiers in Robotics and AI, 9*, 993997.

Vom Hofe, N., Sossalla, P., Hofer, J., Vielhaus, C. L., Rischke, J., Steinke, J., & Fitzek, F. H. (2023). Robotics meets augmented reality: Real-time mapping with Boston Dynamics Spot and Microsoft HoloLens 2. In *2023 IEEE 24th International Symposium on a World of Wireless, Mobile and Multimedia Networks (WoWMoM)*, (pp. 358-360).

Wetzel, E. M., Liu, J., Leathem, T., & Sattineni, A. (2022). The use of Boston Dynamics SPOT in support of LiDAR scanning on active construction sites. In *ISARC Proceedings of the International Symposium on Automation and Robotics in Construction* (Vol. 39, pp. 86-92). IAARC Publications.

Yang, X. J., Schemanske, C., & Searle, C. (2021). Toward quantifying trust dynamics: how people adjust their trust after moment-to-moment interaction with automation. *Human Factors, 65*(5), 862-878.

Young, T., & Winstead, M. (2025, January 7). Human-machine integration: tactical-level employment and the EXFOR Ras Platoon. *U.S. Army*. https://www.army.mil/article/282372/human_machine_integration_tactical_level_employment_and_the_exfor_ras_platoon

Young, J. E., Hawkins, R., Sharlin, E., & Igarashi, T. (2009). Toward acceptable domestic robots: applying insights from social psychology. *International Journal of Social Robotics, 1*, 95-108.

Zhao, F., Henrichs, C., & Mutlu, B. (2020). Task interdependence in human-robot teaming. In *2020 29th IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)* (pp. 1143-1149). IEEE.

Zimmermann, S., Poranne, R., & Coros, S. (2021). Go fetch! Dynamic grasps using Boston Dynamics Spot with external robotic arm. In *2021 IEEE International Conference on Robotics and Automation (ICRA)* (pp. 4488-4494). IEEE.

Złotowski, J., Sumioka, H., Eyssel, F., Nishio, S., Bartneck, C., & Ishiguro, H. (2018). Model of dual anthropomorphism: the relationship between the media equation effect and implicit anthropomorphism. *International Journal of Social Robotics, 10*, 701-714.