Computers in Human Behavior: Artificial Humans 8 (2026) 100295

Contents lists available at ScienceDirect

Computers in Human Behavior: Artificial Humans

journal homepage: www.journals.elsevier.com/computers-in-human-behavior-artificial-humans

Trusting the black box: Adapting a multidimensional measure of trust in generative AI

Lilach Alon *, Inbar Levkovich

Faculty of Education and Teaching, Tel-Hai Academic College, Kiryat Shemona, Israel

ARTICLE INFO

Keywords:
Trust in GenAI
Trust in automation
Human GenAI interaction
Scale adaptation
Calibrated reliance
Trust dimensions

ABSTRACT

As Generative AI (GenAI) becomes integrated into routine workflows, the critical question shifts from adoption to reliance. However, research is hindered by a lack of an adapted measurement tools, as studies often depend on unverified adaptations of traditional automation scales. This study addresses this gap by developing and validating a multidimensional trust scale explicitly tailored for GenAI. We adapted Körber's Trust in Automation (TiA) questionnaire to the GenAI context and administered it to a large adult sample (N = 2169). We first examined whether the original TiA structure transferred to the GenAI context. Because it did not, we used exploratory factor analysis to identify the dimensional structure that emerged in this sample. The analysis revealed that the original factor structure did not replicate. Instead, items consolidated into three distinct dimensions: (1) competence and reliability, (2) familiarity, and (3) caution and uncertainty awareness. A key finding was that negatively worded items formed a coherent, verification-oriented component rather than merely reflecting distrust, supported by higher reliability when items were retained in their original direction. Moreover, distinct patterns of association with GenAI experience further validated the structure: while experience predicted competence and familiarity, it showed weaker associations with caution, suggesting that caution can coexist with confidence in GenAI capabilities. We introduce the TiAI scale, an adapted instrument designed to measure calibrated reliance. These findings support a nuanced understanding of trust in GenAI, distinguishing between uncritical acceptance and a prudent, uncertainty-aware approach to generative outputs.

1. Introduction

As generative AI (GenAI) tools such as large language models become integrated into everyday work and learning, the central question is no longer whether people will try them, but when they will rely on them for real tasks (Göndöcs et al., 2025). This widespread use makes trust a population-level determinant of everyday reliance and verification, with implications for how errors spread and how decisions are made across domains such as learning, work, and civic information (Alon & Malinoff, 2025). Yet GenAI differs from traditional automation in ways that can shift what "trust" means and how standard trust items are interpreted. GenAI is often used for open-ended tasks where correctness is ambiguous (Venter et al., 2025), outputs can vary with small prompt changes or repeated runs, affecting perceived reliability and predictability (Funk et al., 2024), and users typically cannot inspect the underlying logic, relying instead on surface cues such as coherence, consistency, and available explanations (Spitzer et al., 2025). These features create a timely need for an adapted measure that captures how people form trust judgments specifically about GenAI tools.

Because GenAI systems are opaque and their outputs are difficult to evaluate fully, users must make reliance decisions under uncertainty, often with limited time or expertise (Wang & Ding, 2024). In practice, they decide whether to accept an output, treat it as provisional, verify it, or avoid using it (Chen, 2025; Li, 2025). We define trust in GenAI as a willingness to rely on a generative system under such uncertainty and vulnerability, where internal reasoning is not observable and errors may carry consequences (Li, 2025). Trust shapes oversight and how reliance is recalibrated after successes, errors, and inconsistencies (Freel et al., 2025), and it is informed not only by perceived usefulness or accuracy but also by concerns about transparency, fairness, privacy, and accountability (Afroogh et al., 2024; Cheong, 2024; Gunasekara et al., 2025; Nastoska et al., 2025).

Despite its centrality, progress is constrained by measurement. In GenAI research, trust is often assessed with brief, tool-specific items that do not support cumulative comparison across systems, samples, and contexts (Li, 2025). Other studies import measures developed for earlier

* Corresponding author.
E-mail addresses: alonlil@telhai.ac.il (L. Alon), levkovinb@telhai.ac.il (I. Levkovich).

https://doi.org/10.1016/j.chbah.2026.100295
Received 19 January 2026; Received in revised form 18 March 2026; Accepted 18 March 2026
Available online 20 March 2026

automation settings or general AI systems with minimal adaptation, without establishing that item meanings and dimensional structure transfer to GenAI tools (Kohn et al., 2021; McGrath et al., 2025). This is a substantive limitation because trust in GenAI is unlikely to be unitary: users may view a tool as capable yet unpredictable, or feel familiar with it while remaining reluctant to rely on it in high-consequence situations (Körber, 2019).

To address these gaps, we adapt a multidimensional scale that explicitly measures trust in GenAI tools. Building on Körber's Trust in Automation (TiA) questionnaire (Körber, 2019), we adapt the items to contemporary GenAI use and validate the resulting instrument in a large adult sample. We examine internal consistency and the factor structure that emerges when respondents report trust-related judgments about GenAI tools in everyday use, and we test how these trust dimensions relate to demographic characteristics and indicators of GenAI experience, given that experience may shape different trust components in different ways (Scholz et al., 2025).

The contribution of this study is therefore exploratory and context-specific: rather than treating trust in AI as a generic attitude or relying on tool-specific items, we examine whether a widely used automation scale retains a meaningful structure when adapted to GenAI.

In doing so, we introduce TiAI as an initial, exploratory adaptation of TiA for generative AI tools. TiAI supports cumulative research by enabling clearer comparisons across studies, populations, and settings, and it has practical value for organizations, public services, and educational programs seeking to assess trust as part of responsible GenAI integration and to identify where support is needed to promote appropriate reliance and critical use.

2. Theoretical background

2.1. Conceptualizing trust in automation

Trust in automation describes how users judge whether a system can and should be relied on (McGrath et al., 2025). At its core, trust is a psychological state involving positive expectations about another agent's behavior under conditions of uncertainty and vulnerability (Lee & See, 2004). Classic models view trust not as blind acceptance but as calibrated reliance that must be proportionate to the system's actual capabilities (Kelly et al., 2023). When trust exceeds capability, misuse or overtrust occurs; when it falls below capability, distrust or disuse results (Bach et al., 2024).

Early conceptualizations by Jian et al. (2000) and Lee and See (2004) identify three recurring bases of trust judgments: purpose, process, and performance. Purpose concerns what the system is for and whether its goals align with user values. Process concerns how the system works: its transparency, predictability, and controllability. Lastly, performance concerns how well it functions over time in terms of reliability, accuracy, and consistency. When the purpose is clear, the process is understandable, and the performance is reliable, perceived ambiguity decreases and trust strengthens. However, when any of these foundations are unclear, perceived ambiguity increases and trust weakens (Hoff & Bashir, 2015; Körber, 2019). These three bases remain visible in most contemporary trust in automation instruments and adaptation studies (Kohn et al., 2021).

Lee and See's (2004) closed-loop model situates trust within a dynamic feedback system linking observation, belief, and reliance. According to their model, users form trust beliefs based on perceived system performance, which then shape their decisions to rely on automation. Subsequent interaction outcomes update trust over time (Sheridan, 2019). Therefore, trust evolves through phases of formation, dissolution, and restoration, shifting with observed reliability and contextual risk (Poornikoo et al., 2025). This perspective emphasizes that trust is history-dependent and sensitive to expectation versus performance alignment.

Building on these foundations, Hoff and Bashir (2015) proposed a three-layer model distinguishing dispositional, situational, and learned trust. Dispositional trust reflects a stable propensity to trust automation across contexts; situational trust depends on task demands and environmental risk; and learned trust develops through experience with system performance over time. This account is consistent with three-category frameworks that locate trust antecedents in the human trustor, the automated system, and the interaction context (Hancock et al., 2011; Li, 2025).

In line with this framing, a measurement-focused review of self-report trust measures highlights that trust is most commonly operationalized through performance-based cues (e.g., reliability/competence), process cues (e.g., predictability/understandability), and experience-related cues (e.g., familiarity), which together support calibrated reliance (Kohn et al., 2021). These same core determinants are also emphasized in conceptual and integrative discussions of trust in automation and its extension to AI (Körber, 2019).

Based on the literature, trust in automation can be understood as a shifting attitude that develops through use and is updated as users gain experience. It reflects both cognitive judgments about system performance and affective reactions to success and failure (Lee & See, 2004). Simulation and modelling studies further suggest that trust evolves through feedback processes over time, and that negative events (e.g., malfunctions) often have stronger and more enduring effects than positive events (Mahmud et al., 2022; Poornikoo et al., 2025). Individual differences, including trust propensity, personality, and domain expertise, also help explain variation in initial trust and in how quickly trust changes with experience (Kraus et al., 2021; Scholz et al., 2025). Overall, trust in automation rests on judgments about purpose, process, and performance, and these judgments are refined through experience as users observe successes, failures, and recovery over time. Table 1 summarizes these dimensions as a bridge to the next section on trust in GenAI.

2.2. Trust measurement in AI contexts

In GenAI settings, the classic bases of trust in automation (e.g., perceived competence and reliability) remain central. However, GenAI introduces distinctive features that can shift how trust is formed and what trust measures must capture. GenAI systems are typically used for open-ended, creative, or advisory tasks where a single correct answer is often unclear, and outputs can vary across prompts and even across repeated runs, making consistency and predictability less stable than in many traditional automation settings (Afroogh et al., 2024; Gunasekara et al., 2025).

In addition, the generative process is largely opaque to users, who cannot inspect the underlying reasoning and therefore lean on surface cues such as fluency, coherence, and conversational presentation when judging outputs. These cues can inflate perceived credibility and reliance even when responses are partially inaccurate (Anderl et al., 2024; Kim et al., 2025), which is especially problematic because LLMs can generate fluent, coherent text while still producing unfactual content

Table 1
Core dimensions of trust in automation.

Dimension	What it captures	Key sources
Purpose	Whether the system's role aligns with user goals and values	Jian et al., 2000; Lee & See, 2004
Process	Whether operation seems understandable, predictable, and controllable	
Performance	Whether the system is reliable, competent, and accurate over time	
Experience-based learning	How trust is updated through use and interaction history	Lee & See, 2004; Hoff & Bashir, 2015
Calibration	Match between trust and actual capability, enabling appropriate reliance	Lee and See (2004)

(Chang & Bergen, 2024).

These characteristics heighten concerns that become especially salient when system logic is difficult to assess and downstream risks extend beyond the immediate task, including transparency and explainability, fairness, privacy, and accountability (Agur Cohen et al., 2025; Huynh & Aichner, 2025). In practice, trust is most often assessed using self-report scales, sometimes complemented by behavioral indicators such as reliance and compliance (Mehrotra et al., 2024).

The trust in automation literature offers several established instruments that have been applied in AI research. Widely used examples include the checklist for trust between people and automation, which operationalizes trust as a single continuum and treats distrust as reverse-coded trust (Jian et al., 2000), and the human-computer trust questionnaire, which captures multiple trust-related beliefs about intelligent decision aids (Hadar-Shoval et al., 2024; Madsen & Gregor, 2000). Other tools prioritize brevity and focus on core drivers such as perceived understanding and perceived performance (Wojton et al., 2020). In addition, many studies assess relatively stable individual differences that shape initial trust, including dispositional trust and propensity to trust automation (Hoff & Bashir, 2015; Kohn et al., 2021), and dedicated propensity measures have been validated for automated technology (Scholz et al., 2025). Table 2 summarizes common self-report instruments and highlights examples of their use in GenAI settings.

As shown in Table 2, recent AI studies often draw on established automation instruments either to capture a global trust attitude or to focus on a small set of drivers (e.g., perceived performance and understanding). However, many studies have used these scales in AI or GenAI settings without re-examining whether item meaning and dimensional structure transfer to contemporary tools. For example, Lelli et al. (2025) used Körber's TiA questionnaire as an outcome measure in an AI-supported design framework, without validating whether the scale functions similarly when interpreted as trust in AI. More recent validation work has begun to evaluate trust measures directly for AI, including evidence for the TiA Scale and its short form across AI applications (McGrath et al., 2025). Yet this line of work has largely emphasized global trust scores, leaving open the question of how multidimensional instruments perform when adapted to GenAI tools.

Körber's TiA questionnaire (2019) provides a strong starting point for GenAI measurement. TiA was developed to assess trust in automated systems as separable trust-related judgments that inform willingness to rely on the system. This approach aligns with models that conceptualize trust as distinct evaluations (e.g., competence, predictability, and experience-based learning) that may not move together in practice (Hoff & Bashir, 2015; Lee & See, 2004). TiA has been widely used across automation settings, supporting cumulative comparison and synthesis across studies (Gutzwiller et al., 2025; Kohn et al., 2021). Its multidimensional structure is particularly relevant for GenAI, where perceived capability can coexist with uncertainty about predictability and where familiarity can increase even when verification remains salient (Li, 2025).

At the same time, three empirical gaps limit current measurement in GenAI contexts. First, the factor structure of automation trust instruments cannot be assumed to hold once item wording targets general purpose GenAI tools, and therefore requires validation. Second, it remains unclear how negatively framed or caution oriented items function in GenAI use. They may behave as simple reverse indicators of trust, or they may capture a distinct verification oriented component that coexists with perceived competence. Third, evidence is limited on whether the resulting trust dimensions show systematic variation with GenAI experience and demographic characteristics in broad non expert populations, where GenAI use is widespread and reliance decisions are frequent.

Accordingly, the contribution of the current study is not a wording update but a context specific validation of whether TiA's items retain their intended structure when applied to general purpose GenAI tools in everyday use. For this reason, we retained the original TiA item content as closely as possible rather than introducing new GenAI-specific items. This design allowed us to test whether an established multidimensional trust framework transfers to GenAI use before extending it to additional concerns such as factual reliability, value alignment, or ethics.

2.3. Research questions

Building on classic conceptualizations of trust in automation as evolving judgments about purpose, process, and performance, this study examines how trust is structured and measured in contemporary GenAI tool use. We adapt the TiA questionnaire (Körber, 2019) to refer explicitly to GenAI tools while retaining its original item content as closely as possible, and evaluate its psychometric performance in this new context. We assess internal consistency and factor structure, and we test how the resulting trust dimensions vary by demographic characteristics and indicators of GenAI experience. Specifically, we address the following research questions:

RQ1. What factor structure characterizes trust in GenAI tools when measured using an adapted version of the TiA questionnaire, and to what extent do the resulting components demonstrate acceptable internal consistency?
RQ2. How do the resulting trust components vary across demographic characteristics (age, gender, education) and across levels of perceived GenAI knowledge and GenAI use?

By addressing these questions, the study provides an adapted measure for assessing trust in GenAI tools across settings. Such a measure can support needs assessment and evaluation in organizations and public services adopting GenAI, inform higher education and professional development initiatives that aim to promote appropriate reliance and critical use, and enable large-scale surveys that track how trust and use vary across social groups and everyday contexts.

Table 2
Self-report tools used to assess trust in automation.

Instrument	Authors	What it captures	Structure (brief)	Use in GenAI contexts
Checklist for trust between people and automation	Jian et al. (2000)	Trust and distrust as opposing poles	12 items	Used and evaluated as a trust in AI measure in validation work (McGrath et al., 2025).
Human-computer trust questionnaire	Madsen and Gregor (2000)	Multi-facet trust in intelligent aids	25 items	Applied in AI enabled decision support and used as a foundation for subsequent measures in human AI interaction (Katzburg et al., 2024; Wang & Ding, 2024).
Trust in automation (TiA) questionnaire	Körber (2019)	Multi-dimensional trust plus related beliefs	19 items across six dimensions	Used as an outcome measure in AI settings and sometimes adapted to refer to AI tools (Lelli et al., 2025).
Trust in automated systems test	Wojton et al. (2020)	Core trust drivers	9 items; performance and understanding	Used as a short trust measure focused on understanding and performance (Wojton et al., 2020).
Propensity to trust in automated technology (TOAST)	Scholz et al. (2025)	Dispositional tendency to trust automation	14 items	Used as a covariate or moderator in studies published after its validation (Scholz et al., 2025).

3. Methods

3.1. Sample and procedure

The sample in this study included 2169 Israeli adults recruited through online snowball sampling. The survey link was initially distributed via social networks and mailing lists and participants were encouraged to share it with others in their networks. Inclusion criteria required that participants be at least 18 years old, reside in Israel, and be able to read Hebrew well enough to complete the questionnaire.

The sample included 734 men (33.9%), 1432 women (66.0%), and three participants who selected another gender category (0.1%). Ages ranged from 18 to 84 years ($M = 37.51$, $SD = 14.68$). In terms of education, 1040 participants reported high school education a professional certificate (48%), 746 reported a bachelor's degree (34.3%), and 383 reported a master's degree or above (17.7%). For occupation, responses were aggregated into four groups: most participants were employed (N = 1657; 76.4%), followed by students (N = 336; 15.5%), unemployed (N = 135; 6.2%), and retired (N = 41; 1.9%).

To characterize experience with GenAI tools, participants reported how frequently they use these tools on a 1 to 5 scale ($M = 3.17$, $SD = 1.23$), and how knowledgeable they perceive themselves to be about GenAI tools on a 1 to 5 scale (M = 3.29, SD = 1.12), indicating moderate levels on average. Participants also reported the number of different GenAI tools they currently use ($M = 1.49$, $SD = 1.07$; range 0-8).

The survey was administered online via a questionnaire hosted on Google Forms. Data were collected between October and December 2025. Participants accessed the survey through a link distributed online, completed the questionnaire in Hebrew, and provided informed consent prior to participation. Participation was voluntary, respondents could discontinue at any time, and responses were analyzed and reported in aggregate form to protect participants' privacy. The study was conducted in accordance with institutional and national ethical guidelines for research with human participants, and the research protocol was reviewed and approved by the ethics committee (IRB).

3.2. Scale development and validation

Trust in GenAI tools (TiAI) was measured using an adapted Hebrew version of Körber's TiA questionnaire (Körber, 2019). The original TiA is a 19-item instrument developed through psychometric refinement to capture six trust-related dimensions: competence/reliability (six items), predictability/understanding (four items), familiarity (two items), intention of developers (two items), propensity to trust (three items), and trust in automation (two items). The scale includes both positively and negatively worded items, with the latter typically reverse-scored in automation research.

For the present study, items were adapted to refer explicitly to GenAI tools while preserving the original meaning as closely as possible (e.g., replacing references to "the system" with references to GenAI tools). We retained the original 19 items, but treated GenAI as a distinct measurement context and therefore revalidated the scale's reliability and factor structure rather than assuming TiA would transfer unchanged.

All items were administered in Hebrew and rated on a five-point Likert scale from 1 (strongly disagree) to 5 (strongly agree). Negatively worded items were retained in their original direction, and primary analyses report results without reverse-scoring. Because reverse-scoring is common in TiA based work, we also repeated the analyses using the conventional reverse-scoring approach as a robustness check; the substantive structure and conclusions were unchanged.

To support content validity and clarity, the adapted Hebrew version underwent expert review. Ten reviewers, including three academic experts in AI and seven individuals meeting the study inclusion criteria, evaluated item clarity, readability, and suitability for assessing trust in GenAI use. Based on their feedback, minor wording revisions were made to three items while preserving their intended meaning. The original response format included a "no response" option; following reviewer recommendations, this option was removed to reduce ambiguity and improve interpretability.

In the current sample, the adapted 19-item scale showed good internal consistency. Reliability was higher when negatively worded items were kept in their original direction (Cronbach's $\alpha = .877$) than when they were reverse-scored (Cronbach's $\alpha = .789$), a pattern addressed in the Discussion. The Results section further evaluates the scale's dimensional structure and provides additional validity evidence.

3.3. Data analysis

Data were first screened for completeness and data quality. Overall, 2182 survey submissions were collected. Seven incomplete responses and six redundant submissions were identified and removed, resulting in a sample of 2169 participants. Data were then screened for accuracy, outliers, and univariate normality. Descriptive statistics were computed for all trust items, and inter-item correlations were examined to assess their suitability for factor analysis.

To address RQ1, we first examined whether the adapted items reproduced Körber's (2019) original dimensional structure in the GenAI context. Because the original structure did not transfer cleanly to the present dataset, we then conducted exploratory analyses using principal components analysis (PCA) with oblimin rotation to identify the structure that best characterised trust in GenAI in this sample (Fabrigar et al., 1999). Given the unexpected change in Cronbach's α after reverse-scoring, we repeated the PCA under both scoring conventions (prior to recoding and after recoding) to verify that the extracted structure was not an artifact of item direction. This analytic sequence is consistent with established guidance for scale development and evaluation, which distinguishes between examinations of dimensionality, reliability, and validity (Boateng et al., 2018).

To address RQ2, we examined how the trust dimensions varied across participant characteristics (age, gender, and education level) and AI experience indicators (perceived AI knowledge and frequency of GenAI use). These relationships were examined using correlations and group comparisons, and by estimating multiple regression models to identify the unique contribution of each predictor when considered simultaneously.

4. Findings

4.1. Factor structure and internal consistency of the TiAI scale

We first examined whether the adapted items reproduced Körber's (2019) original dimensional structure in the GenAI context. Because the original structure did not transfer cleanly to the present dataset, we proceeded with exploratory analysis to identify the structure that best characterised trust in GenAI in this sample. A principal components analysis (PCA) with oblimin rotation was conducted to allow for correlated dimensions. Given the unexpected change in Cronbach's α after reverse-scoring, we repeated the PCA under both scoring conventions (prior to recoding and after recoding). The extracted component pattern was identical, indicating that the dimensional structure was not an artifact of item direction; results are reported using the non-reversed item scoring.

Factor interpretation was based on a combination of statistical and substantive criteria, including eigenvalues, interpretability of the component structure, and the suitability of components for scoring. We retained items based on their strongest substantive loading pattern shown in Table 3. Where items showed secondary loadings, assignment was based on their primary loading together with conceptual fit. Components that did not form a coherent multi-item dimension were not retained as standalone subscales.

As shown in Table 3, four components had eigenvalues above 1. However, the fourth component did not form a stable multi-item

Table 3
PCA of the adapted TiAI items (oblimin rotation)[a].

Statements	Factor 1	Factor 2	Factor 3	Factor 4
1. I can rely on GenAI tools	0.792			
2. I trust GenAI tools.	0.779			0.402
3. I feel confident in GenAI's capabilities.	0.777			0.484
4. I tend to trust GenAI rather than mistrust it.	0.762			
5. GenAI systems operate reliably.	0.743			
6. GenAI systems generally work well.	0.691			0.411
7. The developers of GenAI are trustworthy people.	0.655			
8. GenAI tools are capable of interpreting situations correctly.	0.622			
9. The developers of GenAI tools take my personal well-being into account.	0.614		−0.409	
10. When using GenAI, I can understand why things happened the way they did.	0.602			0.542
11. GenAI is capable of performing complex tasks.	0.569	−0.482		
12. GenAI may make occasional errors.		0.789		
13. One should be cautious with unfamiliar GenAI systems.		0.735		
14. It is difficult to predict what GenAI will do next.		0.727		
15. A GenAI system malfunction is likely.		0.690		
16. GenAI systems react unpredictably.		0.400	0.604	
17. I am familiar with GenAI tools.				0.850
18. I understand what GenAI tools are doing at any given moment.	0.414			0.721
19. I have used GenAI tools before.		−0.484	0.408	0.655
Eigenvalues	**6.42**	**2.44**	**1.81**	**1.14**
% Variance	**33.80**	**12.86**	**6.22**	**6.00**

[a] Loadings <.40 have been omitted.

dimension and was defined primarily by a single item ('I have used GenAI tools before'). Because our goal was to derive interpretable and scoreable trust dimensions, we did not retain this component as a standalone subscale. Instead, we retained a three-dimensional scoring solution. The first component reflected competence and reliability. The second reflected caution and uncertainty awareness, with strong loadings for negatively worded items referring to errors, unpredictability, malfunction, and caution. Although these items shared negative phrasing, their content consistently pointed to evaluative concern and monitoring in GenAI use. The third reflected familiarity, with high loadings for items referring to familiarity, understanding, and prior use.

Several items showed secondary loadings across factors, but we retained them because their primary loading pattern remained interpretable and conceptually coherent in the GenAI context. In particular, items reflecting trust, confidence, and perceived understanding (e.g., items 2, 3, 6, and 10) showed overlap between competence/reliability and familiarity, which is theoretically plausible given that repeated use and subjective understanding may shape confidence in GenAI performance. Similarly, items related to unpredictability and prior use (e.g., items 9 and 11) showed overlap with caution or familiarity dimensions, but their strongest and most interpretable placement remained within the retained three-dimensional solution. Overall, because these secondary loadings did not obscure the overall structure and the resulting dimensions remained internally coherent, we retained these items in the final scale rather than removing them and re-estimating the solution.

Based on these results, we retained a three-dimensional scoring solution for TiAI, comprising competence and reliability, caution and uncertainty awareness, and familiarity. For TiAI scoring, we therefore recommend not reverse-scoring the negatively worded items; instead, they comprise the caution and uncertainty awareness subscale, such that higher scores indicate stronger verification-oriented caution (see Appendix A for the final TiAI items).

Internal consistency was then examined for each subscale. Competence and reliability showed high internal consistency (11 items; Cronbach's $\alpha = .894$). Caution and uncertainty awareness showed acceptable internal consistency (5 items; Cronbach's $\alpha = .730$). Familiarity showed modest internal consistency (3 items; Cronbach's $\alpha = .650$), consistent with the small number of items.

The intercorrelations among the three TiAI dimensions provided additional evidence that the subscales are related but not redundant (Table 4). Reliability and competence was moderately correlated with familiarity ($r = .559$, $p < .001$), whereas its association with caution and uncertainty awareness was smaller ($r = .248$, $p < .001$). Familiarity was also positively correlated with caution and uncertainty awareness ($r = .308$, $p < .001$). All correlations were statistically significant, but their magnitudes indicate that the three dimensions capture distinguishable aspects of trust in GenAI rather than a single undifferentiated construct.

4.2. Associations between TiAI dimensions, demographic and AI experience

First, we examined descriptive statistics for the three TiAI dimensions to characterize overall levels of trust in GenAI in the sample (Fig. 1).

As shown, participants reported mid-range trust judgments on the 1-5 scale. Reliability/competence and familiarity showed similar mean levels and were slightly below the midpoint ($M = 2.90$, $SD = 0.70$; and $M = 2.90$, $SD = 0.90$, respectively). By contrast, caution/uncertainty awareness was higher and above the midpoint ($M = 3.46$, $SD = 0.75$), indicating that participants tended to endorse verification-oriented caution alongside their evaluations of AI capability and experience.

We then examined how the three trust dimensions were associated with demographic characteristics (age, gender, and education) and with AI experience (perceived AI knowledge, frequency of AI use, and number of AI tools used). Table 5 presents the correlations among the trust dimensions, AI experience indicators, and age.

Reliability/competence was positively associated with perceived GenAI knowledge ($r = .333$), frequency of GenAI use ($r = .399$), and the number of GenAI tools used ($r = .202$), indicating that participants who reported greater engagement with AI also tended to report higher confidence in AI's competence and reliability. Familiarity showed stronger associations with GenAI experience, including correlations with perceived AI knowledge ($r = .588$), frequency of AI use ($r = .561$), and the number of GenAI tools used ($r = .418$). Caution/uncertainty awareness (non-reversed items) showed small positive correlations with perceived GenAI knowledge ($r = .176$), frequency of GenAI use ($r = .155$), and number of tools used ($r = .170$). This suggests that more experienced users not only reported greater familiarity and confidence in GenAI, but also expressed slightly greater uncertainty awareness and a stronger verification-oriented stance.

Trust dimensions were also related to one another. Reliability/competence correlated positively with familiarity ($r = .559$) and with caution/uncertainty awareness ($r = .248$), and familiarity also correlated positively with caution/uncertainty awareness ($r = .308$), indicating that confidence and experience co-occurred with uncertainty-

Table 4
Intercorrelations among the TiAI dimensions.

Variable	1	2	3
1. Reliability/competence	-		
2. Caution/uncertainty awareness	.248**	-	
3. Familiarity	.559**	.308**	-

**$p < .001$.

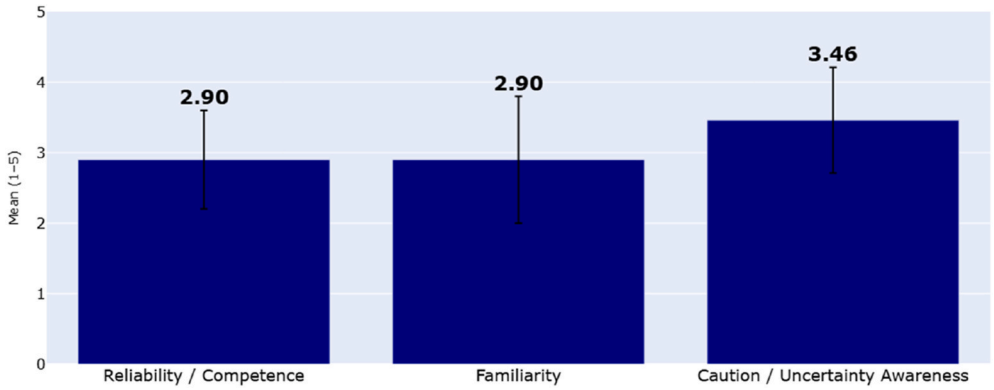

Fig. 1. Mean scores and SDs for TiAI dimensions on a 5-point scale (N = 2169).

Table 5
Pearson correlations among TiAI dimensions, AI experience, and age (N = 2169).

Variable	1	2	3	4	5	6	7
1. Reliability/competence	—						
2. Caution/uncertainty awareness	.248**	—					
3. Familiarity	.559**	.308**	—				
4. Perceived GenAI knowledge	.333**	.176**	.588**	—			
5. GenAI use frequency	.399**	.155**	.561**	.727**	—		
6. Number of GenAI tools used	.202**	.170**	.418**	.373**	.402**	—	
7. Age	−.100**	−.094**	−.228**	−.255**	−.277**	−.041	—

**$p < .001$.

aware judgments.

Age showed the opposite pattern for the trust and experience indicators: it was weakly negatively associated with reliability/competence ($r = -.100$) and moderately negatively associated with familiarity ($r = -.228$), suggesting that older participants tended to report lower familiarity with AI and somewhat lower confidence in AI's capabilities. Age was also weakly negatively associated with caution/uncertainty awareness ($r = -.094$), indicating slightly lower endorsement of uncertainty-aware caution among older participants.

To examine the unique contribution of each predictor, we estimated three multiple regression models with age, gender, education level, perceived AI knowledge, and AI use frequency entered simultaneously (Table 6). The full regression tables are provided in the supplementary materials. Because perceived GenAI knowledge and GenAI use frequency were moderately to strongly associated, we examined multicollinearity diagnostics. These did not indicate problematic collinearity among the predictors (tolerance = .460-.971; VIF = 1.030-2.172), suggesting that the regression estimates were not adversely affected by predictor overlap.

For reliability and competence, the model explained 16.5% of the variance. GenAI use frequency was the strongest predictor ($\beta = .331$), and perceived GenAI knowledge also contributed positively ($\beta = .088$). Age and gender were not significant predictors, and education showed only a small, borderline effect.

For caution/uncertainty awareness, the model explained 4.4% of the variance. Age was a small negative predictor ($\beta = -.081$), whereas education ($\beta = .104$) and perceived AI knowledge ($\beta = .114$) predicted higher scores on this dimension. GenAI use frequency and gender were not significant predictors in this model. This pattern suggests that uncertainty awareness tends to be slightly higher among participants with more perceived GenAI knowledge and higher education.

For familiarity, the model explained 40.4% of the variance. Perceived GenAI knowledge ($\beta = .356$) and AI use frequency ($\beta = .265$) were the strongest predictors, indicating that familiarity is closely tied to perceived AI experience. In addition, familiarity was lower with age ($\beta = -.080$) and showed smaller but significant associations with education ($\beta = .096$). In addition, women reported lower familiarity with GenAI tools than men, controlling for age, education, perceived GenAI knowledge, and GenAI use ($\beta = - .212$).

Table 6
Model fit statistics for multiple regression models predicting TiAI dimensions (N = 2169).

Model	R^2	F (df1, df2)	p
Reliability/competence	.165	85.25 (5, 2163)	<.001
Caution/uncertainty awareness	.044	20.00 (5, 2163)	<.001
Familiarity	.404	293.18 (5, 2163)	<.001

5. Discussion

In this study, we adapted the TiA questionnaire (Körber, 2019) to the context of generative AI and examined its dimensional structure in a large adult sample. Grounded in trust in automation theory (Hoff & Bashir, 2015; Lee & See, 2004), we examined how respondents interpret trust items when the target is a general purpose, prompt sensitive GenAI tool rather than a bounded, rule based system. The findings make a measurement focused contribution. TiAI provides a multidimensional instrument for assessing trust in GenAI in the general population, while also showing that the structure of a widely used automation scale shifts in meaningful ways when applied to GenAI tools. Importantly, we retained the original 19 item content and re-examined its psychometric structure for GenAI use.

A central result is that the original multidimensional configuration did not transfer intact. Instead of reproducing the more differentiated structure typical of automation questionnaires, the adapted items consolidated into three interpretable dimensions: competence/reliability, caution/uncertainty awareness, and familiarity. This pattern is novel in the GenAI measurement context because it demonstrates, empirically, that trust in GenAI is not well captured as a single global attitude yet also does not necessarily preserve the full set of TiA subdimensions when the target is a generative tool. At the same time, the factors align with recurrent components in the trust literature,

supporting continuity in what is being measured while clarifying how these components cohere in GenAI use (Kohn et al., 2021; Körber, 2019). More broadly, this finding suggests that scales developed for earlier technology contexts cannot be assumed to transfer unchanged to GenAI use. Recent work on adjacent AI-related constructs similarly points to the need to re-examine inherited measures in light of the distinctive forms of interaction, verification, and reflection involved in GenAI use (Lau et al., 2025).

The competence/reliability factor maps onto the performance basis of trust, reflecting judgments about dependability and accuracy (Jian et al., 2000; Lee & See, 2004). This emphasis is consistent with common practice in AI trust measurement, where performance perceptions often dominate global trust ratings and short-form instruments foreground perceived performance (Madsen & Gregor, 2000; McGrath et al., 2025; Wojton et al., 2020). The familiarity factor captures the experience-based learning component emphasized in conceptual accounts, reflecting accumulated interaction history and comfort that shapes how users interpret system behavior over time (Kohn et al., 2021). The third factor, caution/uncertainty awareness, is especially diagnostic for GenAI: it reflects endorsement of possible errors, variability, and the need to verify outputs. This component is theoretically consistent with process-based trust in contexts where internal logic is difficult to inspect, but in GenAI it is expressed as an explicit verification stance that supports calibrated reliance (Hoff & Bashir, 2015; Lee & See, 2004).

This structure also clarifies a persistent measurement challenge in GenAI research. Recent studies continue to operationalize "trust" using different instruments, shortened versions, or single global scores (e.g., Afroogh et al., 2024; Gutzwiller et al., 2025; McGrath et al., 2025), making it difficult to accumulate findings when the same label reflects different underlying judgments. Our results support reporting a small, interpretable profile of trust components in GenAI contexts, rather than relying solely on total scores that can mask meaningful differences across tasks, risk levels, and user experience.

The behavior of negatively framed items provides a second novel contribution with direct implications for GenAI measurement. In many automation questionnaires, negatively worded items are treated as simple inverses and are reverse-scored to align with favorable trust. In our data, these items cohered as a distinct component regardless of whether they were reverse-scored for comparability or retained in their original direction. This indicates that their clustering is not a scoring artifact but reflects a substantive response tendency. When interpreted in their original direction, the items capture uncertainty awareness and verification-oriented caution, suggesting that in GenAI contexts, endorsing uncertainty can coexist with perceived competence and may reflect calibrated reliance rather than global distrust. This is particularly relevant for GenAI systems whose outputs can be fluent yet variable and whose reasoning process is not directly observable to users.

This interpretation is consistent with accounts of GenAI as a black box for end users, where outputs are visible but the basis for them is difficult to inspect and failure conditions can be hard to anticipate (Alon & Krtalić, 2025; Mohamed et al., 2025; Picard et al., 2023). It also aligns with evidence that explainability features can shape trust without necessarily improving users' ability to evaluate correctness, sometimes functioning as reassurance cues rather than supporting meaningful inspection (Park & Young Yoon, 2025; Rosenbacke et al., 2024). In our data, competence/reliability and familiarity were strongly related to GenAI experience indicators, whereas caution/uncertainty awareness showed weaker associations and was less explained by standard predictors. This pattern suggests that greater engagement with GenAI is associated with increased confidence and comfort while only modestly increasing uncertainty awareness and verification orientation. Put differently, experience may strengthen both confidence and calibration, rather than producing uniformly higher perceived predictability or transparency. Demographic effects were comparatively small, consistent with findings that demographic differences in trust can vary by context and measurement approach (Afroogh, 2024).

Overall, the findings support conceptualizing trust in GenAI as a differentiated profile that combines confidence in competence with a distinct uncertainty-aware component that supports verification and appropriate oversight. This has both theoretical and practical implications: it strengthens cumulative measurement by offering a multidimensional instrument tailored to GenAI, and it supports applied assessment by distinguishing between trust that reflects uncritical reliance and trust that is calibrated to uncertainty and risk in GenAI use.

5.1. Implications for research and practice

These findings have clear implications for both theory and practice. Theoretically, they provide measurement-based support for accounts that conceptualize trust as multidimensional and emphasize calibrated reliance rather than trust maximization (Hoff & Bashir, 2015; Lee & See, 2004). In GenAI contexts, trust is not a single continuum: confidence in competence and reliability can coexist with caution/uncertainty awareness, reflected in endorsement of possible error, variability, and the need to verify outputs.

This pattern cautions against treating higher caution/uncertainty awareness as a simple deficit in trust. In GenAI use, it may instead reflect an adaptive, verification-oriented stance that is responsive to black-box properties, limited inspectability, and output variability (Park & Young Yoon, 2025; Rosenbacke et al., 2024). Methodologically, the results also underscore the limitations of global trust scores, which can mask meaningful profiles (e.g., high confidence alongside high caution) and contribute to inconsistent associations between trust and user characteristics or outcomes across studies. Practically, a multidimensional profile can help organizations and educators distinguish uncritical reliance from calibrated trust and target interventions accordingly. Table 7 summarizes how the three adapted TiAI dimensions map onto core components of trust in automation.

For practice, the findings support using TiAI as a component-level profile rather than a single trust score when assessing GenAI use in everyday settings. Lower competence/reliability confidence points to where users may hesitate to rely on GenAI, suggesting a need for clearer organizational guidance about what the tool does well, where it commonly fails, and which routine tasks it is appropriate for. Elevated caution/uncertainty awareness signals that users anticipate variability and error, underscoring the importance of practical safeguards even in day to day use.

Support should therefore focus on building calibrated reliance. This can include AI literacy that explains why outputs vary across prompts and contexts, what error types are common, and how to detect warning signs, paired with concrete verification routines (e.g., cross-checking against trusted sources, triangulating with a second method, or

Table 7
Mapping classic trust in automation to TiAI dimensions.

Dimension	What it captures	How it appeared in TiAI
Purpose	Whether the system's role aligns with user goals and values	Not retained as a distinct factor in TiAI structure
Process	Whether operation seems understandable, predictable, and controllable	Expressed as *caution/ uncertainty awareness*, rather than as perceived understanding.
Performance	Whether the system is reliable, competent, and accurate over time	Retained as *reliability/ competence*
Experience-based learning	How trust is updated through use and interaction history	Retained as *familiarity*
Calibration	Match between trust and actual capability, enabling appropriate reliance	Reflected in the *trust profile* across dimensions

documenting checks when stakes are meaningful) (Alon & Levkovich, 2026). Organizations can also provide risk-stratified guidance that distinguishes low-stakes uses (e.g., drafting, brainstorming, or summarizing for internal use) from higher-stakes uses where corroboration, traceability, or nonuse is expected. Across contexts, calibrated reliance means using GenAI when it fits the task, treating outputs as provisional by default, and verifying before acting when consequences matter.

5.2. Limitations and future studies

Several limitations should be noted. First, the present study does not constitute a full scale-validation program. Although the sample was large, the psychometric evidence reported here is exploratory and based on a single dataset. Future research should test the resulting structure using confirmatory methods in independent samples and examine measurement invariance across relevant subgroups. Second, the study relied on a self-report survey, so the findings are correlational and cannot speak to causality or directly capture reliance and verification behavior. In addition, because the caution and uncertainty awareness factor was composed largely of negatively worded items, future research should test whether this pattern reflects a substantive dimension of trust in GenAI, a wording effect, or a combination of both.

Third, the sample was recruited through online snowball sampling in a single national context and was characterized by a substantial gender imbalance. These features limit the generalizability of the findings and mean that the present results should not be interpreted as establishing full population-level applicability. In addition, because measurement invariance across gender, age groups, and experience levels was not tested, group comparisons should be interpreted cautiously. The sample also skewed younger, which may limit generalizability to older adults and to settings where GenAI exposure, goals, and reliance patterns differ by age. In addition, the items referred to GenAI tools in general rather than a specific system or task. Relatedly, the scale was designed to test whether a well-established trust framework transfers to GenAI use, rather than to capture all GenAI-specific concerns. This broad framing improves portability across everyday contexts, but it may encourage more global responding and reduce sensitivity to context-specific trust judgments. Future work may therefore extend the measure with items targeting issues such as factual reliability, value alignment, and ethical oversight more directly. Finally, the study was conducted in Israel, and further research in a wider range of countries is needed to examine cross-cultural generalizability.

Future work should extend the key implication of the caution and uncertainty awareness dimension. In GenAI use, endorsing uncertainty may reflect calibrated reliance rather than generalized distrust. A priority is to connect TiAI profiles to objective indicators of understanding and calibration, such as knowledge tests about GenAI limitations, performance in detecting hallucinations or unsupported claims in controlled tasks, and observed verification behaviors such as source checking, cross validation, and selective reliance. Longitudinal studies could examine how trust profiles change with experience over time, including whether competence and reliability confidence increases alongside caution and uncertainty awareness, and whether this combination predicts safer use. Experimental research can test which support features shape calibration, such as provenance indicators, uncertainty cues, or different explanation designs, and whether they improve appropriate reliance without inflating overconfidence.

CRediT authorship contribution statement

Lilach Alon: Writing – original draft, Validation, Methodology, Formal analysis, Data curation, Conceptualization. **Inbar Levkovich:** Writing – review & editing, Visualization, Methodology, Data curation, Conceptualization.

Declaration on informed consent

All participants were adults and provided informed consent before participation. Participation was voluntary, respondents could discontinue at any time without penalty, and no identifying personal information was collected beyond what was necessary for the study. The privacy rights of all human participants were observed.

Funding statement

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Declaration of competing interest

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

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.chbah.2026.100295.

Appendix A. TiAI item set and scoring instructions*

Statements	Strongly disagree	Rather disagree	Neither disagree nor agree	Rather agree	Strongly agree
Reliability/competence					
1. I can rely on Generative AI tools	1	2	3	4	5
2. I trust Generative AI tools.	1	2	3	4	5
3. I feel confident in Generative AI's capabilities.	1	2	3	4	5
4. I tend to trust Generative AI rather than mistrust it.	1	2	3	4	5
5. Generative AI systems operate reliably.	1	2	3	4	5
6. Generative AI systems generally work well.	1	2	3	4	5
7. The developers of Generative AI are trustworthy people.	1	2	3	4	5
8. Generative AI tools are capable of interpreting situations correctly.	1	2	3	4	5
9. The developers of Generative AI tools take my personal well-being into account.	1	2	3	4	5
10. When using Generative AI, I can understand why things happened the way they did.	1	2	3	4	5
11. Generative AI is capable of performing complex tasks.	1	2	3	4	5

(continued on next page)

(*continued*)

Statements	Strongly disagree	Rather disagree	Neither disagree nor agree	Rather agree	Strongly agree
Caution/uncertainty awareness					
12. Generative AI may make occasional errors.	1	2	3	4	5
13. One should be cautious with unfamiliar Generative AI systems.	1	2	3	4	5
14. It is difficult to predict what Generative AI will do next.	1	2	3	4	5
15. A Generative AI system malfunction is likely.	1	2	3	4	5
16. Generative AI systems react unpredictably.	1	2	3	4	5
Familiarity					
17. I am familiar with Generative AI tools.	1	2	3	4	5
18. I understand what Generative AI tools are doing at any given moment.	1	2	3	4	5
19. I have used Generative AI tools before.	1	2	3	4	5

* TiAI is scored as three subscales computed as item means: competence and reliability (11 items), caution and uncertainty awareness (5 items), and familiarity (3 items). Negatively worded items should not be reverse-scored; instead, they comprise the caution and uncertainty awareness subscale, such that higher scores indicate stronger verification-oriented caution.

References

Afroogh, S., Akbari, A., Malone, E., Kargar, M., & Alambeigi, H. (2024). Trust in AI: Progress, challenges, and future directions. *Humanities and Social Sciences Communications, 11*(1), 1–30. https://doi.org/10.1057/s41599-024-04044-8

Agur Cohen, D., Heymann, A., & Levkovich, I. (2025). Partners in practice: Primary care physicians define the role of artificial intelligence. *Healthcare, 13*(16). https://doi.org/10.3390/healthcare13161972, 1972–1972.

Alon, L., & Krtalić, M. (2025). Designing for older users: A theoretical framework for information seeking and evaluation in AI systems. *Proceedings of the Association for Information Science and Technology, 62*, 871–876. https://doi.org/10.1002/pra2.1305

Alon, L., & Levkovich, I. (2026). Information literacy in the age of generative tools: Development and validation of the AI information literacy scale (AILIS). *Computers in Human Behavior: Artificial Humans, 7*, Article 100254. https://doi.org/10.1016/j.chbah.2026.100254

Alon, L., & Malinoff, T. (2025). Understanding barriers to AI use among public sector knowledge workers. *Proceedings of the Association for Information Science and Technology, 62*, 1346–1348. https://doi.org/10.1002/pra2.1399

Anderl, C., Klein, S. H., Sarigül, B., Schneider, F. M., Han, J., Fiedler, P. L., & Utz, S. (2024). Conversational presentation mode increases credibility judgements during information search with ChatGPT. *Scientific Reports, 14*(1), Article 17127. https://doi.org/10.1038/s41598-024-67829-6

Bach, T. A., Khan, A., Hallock, H., Beltrão, G., & Sousa, S. (2024). A systematic literature review of user trust in AI-enabled systems: An HCI perspective. *International Journal of Human-Computer Interaction, 40*(5), 1251–1266. https://doi.org/10.1080/10447318.2022.2138826

Boateng, G. O., Neilands, T. B., Frongillo, E. A., Melgar-Quiñonez, H. R., & Young, S. L. (2018). Best practices for developing and validating scales for health, social, and behavioral research: A primer. *Frontiers in Public Health, 6*, 149. https://doi.org/10.3389/fpubh.2018.00149

Chang, T. A., & Bergen, B. K. (2024). Language model behavior: A comprehensive survey. *Computational Linguistics, 50*(1), 293–350. https://doi.org/10.1162/coli_a_00492

Chen, Z. (2025). From traditional to technological: Integrating AI tools into leadership development programs. *The Leadership & Organization Development Journal, 46*(4), 543–558. https://doi.org/10.1108/LODJ-12-2024-0810

Cheong, B. C. (2024). Transparency and accountability in AI systems: Safeguarding wellbeing in the age of algorithmic decision-making. *Frontiers in Human Dynamics, 6*, Article 1421273. https://doi.org/10.3389/fhumd.2024.1421273

Fabrigar, L. R., Wegener, D. T., MacCallum, R. C., & Strahan, E. J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods, 4*(3), 272–299. https://doi.org/10.1037/1082-989X.4.3.272

Freel, A., Pias, S. B. H., Šabanović, S., & Kapadia, A. (2025). How misclassification severity and timing influence user trust in AI image classification: User perceptions of high-and low-stakes contexts. In *Proceedings of the 2025 ACM conference on fairness, accountability, and transparency* (pp. 2906–2923). https://doi.org/10.1145/3715275.3732187

Funk, P. F., Hoch, C. C., Knoedler, S., Knoedler, L., Cotofana, S., Sofo, G., & Alfertshofer, M. (2024). ChatGPT's response consistency: A study on repeated queries of medical examination questions. *European Journal of Investigation in Health, Psychology and Education, 14*(3), 657–668. https://doi.org/10.3390/ejihpe14030043

Göndöcs, D., Horváth, S., & Dörfler, V. (2025). Uncovering the dynamics of human-AI hybrid performance: A qualitative meta-analysis of empirical studies. *International Journal of Human-Computer Studies.* , Article 103622. https://doi.org/10.1016/j.ijhcs.2025.103622

Gunasekara, L., El-Haber, N., Nagpal, S., Moraliyage, H., Issadeen, Z., Manic, M., & De Silva, D. (2025). A systematic review of responsible artificial intelligence principles and practice. *Applied System Innovation, 8*(4), 97. https://doi.org/10.3390/asi8040097

Gutzwiller, R. S., Yousefi, R., Larson-Calcano, T., Lee, J. R., Verma, A., Tenhundfeld, N. L., & Maknojia, I. (2025). Systematic review of the use and modification of the "trust in automated systems scale". *Proceedings of the Human Factors and Ergonomics Society - Annual Meeting, 69*(1), 2176–2181. https://doi.org/10.1177/10711813251357911

Hadar-Shoval, D., Asraf, K., Shinan-Altman, S., Elyoseph, Z., & Levkovich, I. (2024). Embedded values-like shape ethical reasoning of large language models on primary care ethical dilemmas. *Heliyon, 10*(18), Article e38056. https://doi.org/10.1016/j.heliyon.2024.e38056

Hancock, P. A., Billings, D. R., Schaefer, K. E., Chen, J. Y., 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

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

Huynh, M. T., & Aichner, T. (2025). In generative artificial intelligence we trust: Unpacking determinants and outcomes for cognitive trust. *AI & Society, 40*, 5849–5869. https://doi.org/10.1007/s00146-025-02378-8

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. https://doi.org/10.1207/S15327566IJCE0401_04

Katzburg, O., Roimi, M., Frenkel, A., Ilan, R., & Bitan, Y. (2024). The impact of information relevancy and interactivity on intensivists' trust in a machine learning–based prediction system: Simulation study. JMIR Human Factors, 11(1), e56924. doi: doi.org/10.2196/56924.

Kelly, S., Kaye, S. A., & Oviedo-Trespalacios, O. (2023). What factors contribute to the acceptance of artificial intelligence? A systematic review. *Telematics and Informatics, 77*, Article 101925. https://doi.org/10.1016/j.tele.2022.101925

Kim, D. H., Jeong, D., Yadgarova, S., Shin, H., Son, J., Subramonyam, H., & Kim, J. (2025). PlanTogether: Facilitating AI application planning using information graphs and large language models. In *Proceedings of the 2025 CHI conference on human factors in computing systems* (pp. 1–23). https://doi.org/10.1145/3706598.3714044

Kohn, S. C., De Visser, E. J., Wiese, E., Lee, Y. C., & Shaw, T. H. (2021). Measurement of trust in automation: A narrative review and reference guide. *Frontiers in Psychology, 12*, Article 604977. https://doi.org/10.3389/fpsyg.2021.604977

Körber, M. (2019). Theoretical considerations and development of a questionnaire to measure trust in automation. In S. Bagnara, R. Tartaglia, S. Albolino, T. Alexander, & Y. Fujita (Eds.), *Proceedings of the 20th congress of the international ergonomics association (IEA 2018)* (pp. 13–30). Cham: Springer. https://doi.org/10.1007/978-3-319-96074-6_2.

Kraus, J., Scholz, D., & Baumann, M. (2021). What's driving me? Exploration and validation of a hierarchical personality model for trust in automated driving. *Human Factors, 63*(6), 1076–1105. https://doi.org/10.1177/0018720820922653

Lau, G. R., Low, W. Y., Tay, L., Guevarra, Y., Gašević, D., & Hartanto, A. (2025). Understanding critical thinking in generative artificial intelligence use: Development, validation, and correlates of the critical thinking in AI use scale. *(arXiv:2512.12413). arXiv*. https://doi.org/10.48550/arXiv.2512.12413

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

Lelli, C., Chiarello, F., & Giordano, V. (2025). Co-intelligence in design: The importance of trust in artificial intelligence. *Proceedings of the Design Society, 5*, 971–980. https://doi.org/10.1017/pds.2025.10111

Li, W. (2025). A study on factors influencing designers' behavioral intention in using AI-generated content for assisted design: Perceived anxiety, perceived risk, and UTAUT. *International Journal of Human-Computer Interaction, 41*(2), 1064–1077. https://doi.org/10.1080/10447318.2024.2310354

Madsen, M., & Gregor, S. (2000). Measuring human-computer trust. In *11th australasian conference on information systems* (Vol. 53, pp. 6–8).

Mahmud, H., Islam, A. N., Ahmed, S. I., & Smolander, K. (2022). What influences algorithmic decision-making? A systematic literature review on algorithm aversion. *Technological Forecasting and Social Change, 175*, Article 121390. https://doi.org/10.1016/j.techfore.2021.121390

McGrath, M. J., Duenser, A., Lacey, J., & Paris, C. (2025). Collaborative human-AI trust (CHAI-T): A process framework for active management of trust in human-AI collaboration. *Computers in Human Behavior: Artificial Humans.* , Article 100200. https://doi.org/10.1016/j.chbah.2025.100200

Mehrotra, S., Degachi, C., Vereschak, O., Jonker, C. M., & Tielman, M. L. (2024). A systematic review on fostering appropriate trust in Human-AI interaction: Trends,

opportunities and challenges. *ACM Journal on Responsible Computing, 1*(4), 1–45. https://doi.org/10.1145/3696449

Mohamed, Y. A., Khoo, B. E., Asaari, M. S. M., Aziz, M. E., & Ghazali, F. R. (2025). Decoding the Black box: Explainable AI (XAI) for cancer diagnosis, prognosis, and treatment planning-A state-of-the art systematic review. *International Journal of Medical Informatics, 193*, Article 105689. https://doi.org/10.1016/j.ijmedinf.2024.105689

Nastoska, A., Jancheska, B., Rizinski, M., & Trajanov, D. (2025). Evaluating trustworthiness in AI: Risks, metrics, and applications across industries. *Electronics, 14*(13), 2717. https://doi.org/10.3390/electronics14132717

Park, K., & Young Yoon, H. (2025). AI algorithm transparency, pipelines for trust not prisms: Mitigating general negative attitudes and enhancing trust toward AI. *Humanities and Social Sciences Communications, 12*(1), 1–13. https://doi.org/10.1057/s41599-025-05116-z

Picard, A., Mualla, Y., Gechter, F., & Galland, S. (2023). Human-computer interaction and explainability: Intersection and terminology. In L. Longo (Ed.), *Explainable artificial intelligence, communications in computer and information science*. Springer. https://doi.org/10.1007/978-3-031-44067-0_12.

Poornikoo, M., Gyldensten, W., Vesin, B., & Øvergård, K. I. (2025). Trust in automation (TiA): Simulation model, and empirical findings in supervisory control of maritime autonomous surface ships (MASS). *International Journal of Human-Computer Interaction, 41*(12), 7521–7548. https://doi.org/10.1080/10447318.2024.2399439

Rosenbacke, R., Melhus, Å., McKee, M., & Stuckler, D. (2024). How explainable artificial intelligence can increase or decrease clinicians' trust in AI applications in health care: Systematic review. *Jmir AI, 3*, Article e53207. https://doi.org/10.2196/53207

Scholz, D. D., Kraus, J., & Miller, L. (2025). Measuring the propensity to trust in automated technology: Examining similarities to dispositional trust in other humans and validation of the PTT-A scale. *International Journal of Human-Computer Interaction, 41*(2), 970–993. https://doi.org/10.1080/10447318.2024.2307691

Sheridan, T. B. (2019). Individual differences in attributes of trust in automation: Measurement and application to system design. *Frontiers in Psychology, 10*, 1117. https://doi.org/10.3389/fpsyg.2019.01117

Spitzer, P., Holstein, J., Morrison, K., Holstein, K., Satzger, G., & Kühl, N. (2025). Don't be fooled: The misinformation effect of explanations in Human–AI collaboration. *International Journal of Human-Computer Interaction*, 1–29. https://doi.org/10.1080/10447318.2025.2574511

Venter, J., Coetzee, S. A., & Schmulian, A. (2025). Exploring the use of artificial intelligence (AI) in the delivery of effective feedback. *Assessment & Evaluation in Higher Education, 50*(4), 516–536. https://doi.org/10.1080/02602938.2024.2415649

Wang, P., & Ding, H. (2024). The rationality of explanation or human capacity? Understanding the impact of explainable artificial intelligence on human-AI trust and decision performance. *Information Processing & Management, 61*(4), Article 103732. https://doi.org/10.1016/j.ipm.2024.103732

Wojton, H. M., Porter, D., Lane, S. T., Bieber, C., & Madhavan, P. (2020). Initial validation of the trust of automated systems test (TOAST). *The Journal of Social Psychology, 160*(6), 735–750. https://doi.org/10.1080/00224545.2020.1749020