

When do researchers trust AI? Trust as a mediator and experience as a boundary condition in researchers' intentions to use generative AI

Anne-Kathrin Kleine^{a,*}, Insa Schaffernak^b, Eva Lerner^{a,b}

^a LMU Munich, Germany

^b Technical University of Applied Sciences Augsburg, Germany

ARTICLE INFO

Keywords:

Generative AI
Stimulus-organism-response
Trust in technology
Perceived accuracy
Usage intention
User experience

ABSTRACT

Generative artificial intelligence (GenAI) is increasingly embedded in academic work, yet its opaque processes and the possibility of misleading outputs make reliance on these tools inherently risky. Building on trust theory and stimulus–organism–response (SOR) logic, we examine how perceived accuracy shapes researchers' intention to use GenAI, and whether this relationship operates through trust and varies by users' experience. Across two online studies with researchers (total $N = 402$), we test a moderated mediation model in which accuracy cues (stimulus) influence trust in GenAI (organism), which in turn predicts usage intention (response), with familiarity/experience moderating the trust–intention link. Study 1 used a three-wave longitudinal design with one-week time lags ($N = 232$), temporally separating perceived accuracy (T1), trust (T2), and usage intention (T3). Results were consistent with (a) a positive association between perceived accuracy and trust, (b) an indirect association between perceived accuracy and usage intention via trust, and (c) a negative interaction showing that the trust–intention relationship weakened as familiarity/experience increased. Study 2 conceptually replicated these findings using a vignette experiment manipulating accuracy of ChatGPT recommendations ($N = 170$): the accuracy manipulation increased trust, and trust predicted usage intention; the indirect effect via trust was moderated by experience. Together, these findings suggest that trust is a key psychological mechanism through which accuracy cues are linked to researchers' intentions to use GenAI in academic work, while highlighting experience as a boundary condition: trust functions more like a heuristic for novices and becomes less relevant as experience grows.

1. Introduction

The increasing availability of generative artificial intelligence (GenAI) has begun to reshape research and academic practice. GenAI tools are changing how researchers access information, draft manuscripts, and engage with scholarly content, supporting tasks such as idea generation, literature review, reference management, and data analysis (Burger et al., 2023; Imran & Almusharraf, 2023; Khlaif et al., 2023; Waduge et al., 2024). Recent work documents broad and growing influence of GenAI across many areas of research activity (Barros et al., 2023; Lee et al., 2024; Nguyen et al., 2024; Panda & Kaur, 2024; Van Quaquebeke et al., 2025). At the same time, GenAI outputs can be inaccurate or misleading (e.g., hallucinated citations, erroneous summaries), creating epistemic and practical risks that users must manage (Lorenz et al., 2024; Naik et al., 2024; Sun et al., 2024).

GenAI is not adopted uniformly across the scientific community:

some researchers resist its use, while others integrate it into everyday workflows (Panda & Kaur, 2024; Watermeyer et al., 2024). Because many GenAI processes are opaque and error-prone, decisions to rely on them often hinge on trust (Glikson & Woolley, 2020; Omrani et al., 2022). Indeed, prior work identifies trust as an important driver of AI adoption (Glikson & Woolley, 2020; Hoff & Bashir, 2015; Kaplan et al., 2023; Omrani et al., 2022), while trust itself is informed by cues about system performance and accuracy (Glikson & Woolley, 2020; Kaplan et al., 2023; Yu et al., 2018).

Theoretically, trust helps users cope with uncertainty and compensate for limited ability to predict outcomes (Chiu et al., 2012; Glikson & Woolley, 2020; Mayer et al., 1995). This suggests that trust may be particularly important for shaping usage intentions among less-experienced users who cannot reliably evaluate GenAI outputs themselves, whereas more experienced users may depend less on trust when deciding whether to use GenAI. Although investigation of

* Corresponding author. LMU Munich, Geschwister-Scholl-Platz 1, 80539, Munich, Germany.

E-mail address: anne-kathrin.kleine@psy.lmu.de (A.-K. Kleine).

<https://doi.org/10.1016/j.chbah.2026.100310>

Received 3 February 2026; Received in revised form 14 April 2026; Accepted 20 April 2026

Available online 28 April 2026

2949-8821/© 2026 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

experience as a moderator of the trust – usage intention link is sparse, prior findings point to reduced dependency on trusting beliefs as users gain experience (Chiu et al., 2012; Lankton et al., 2012).

In two studies (Study 1: a three-wave longitudinal observational design; Study 2: a conceptual replication that manipulates accuracy via vignettes), we examine a moderated mediation model in which perceived accuracy is linked to intention to use GenAI via trust, and in which user experience moderates the strength of the trust–usage intention path. This research contributes to theory and research on AI usage intentions in two ways. First, we theorize and test perceived accuracy as an antecedent of behavioral intention through its relationship with trust, thereby recentring trust as a mediator that links specific system qualities to intentions to use GenAI—consistent with the stimulus–organism–response (SOR) model (Mehrabian & Russell, 1974). We argue that accuracy cues are relevant predictors of trust and usage intention in research contexts because accuracy can be central to whether a tool is considered a reliable partner in knowledge production (Van Quaquebeke et al., 2025). Second, by modelling experience as a moderator of the trust–usage intention relationship, we propose that trust may function like a heuristic for novices and become less central as evaluative competence grows, which invites future research on over-reliance and unwarranted trust among inexperienced users.

Practically, the study examines accuracy feedback and trust as actionable points for intervention—measures that developers, universities, and research institutes could consider when designing deployment, governance, and training strategies for GenAI in research settings. Because we explicitly test whether the influence of trust on usage intention varies with prior experience, the project aims to provide an empirical basis for thinking about segmented interventions (e.g., taking measures to increase experience before trust; not emphasizing trust among experienced users) and motivate future work on when trust should be limited to protect research integrity.

2. Theoretical background

2.1. Opportunities and challenges of using GenAI in research

Generative Artificial Intelligence (GenAI) refers to advanced AI systems capable of producing varied forms of content—such as text, images, music, and other media—based on limited input data. These technologies leverage deep learning techniques, especially neural networks trained on vast datasets, to learn complex patterns and generate new outputs (McIntosh et al., 2025; Sengar et al., 2024). The most prominent GenAI applications in everyday and professional contexts are large language models (LLMs) such as ChatGPT, which are increasingly integrated into education, research, and professional practice (Bobula, 2024; Van Quaquebeke et al., 2025). While ethical and practical concerns about the use of GenAI persist, there is broad agreement that these systems will influence how knowledge is produced and used. Recent surveys highlight the role of LLMs in idea generation, hypothesis discovery, methodological planning, data analysis, scientific writing, publication decisions, and peer review, indicating support across the entire research cycle (Biyela et al., 2024; Bobula, 2024; Dadkhah et al., 2024; Harries et al., 2025; Van Noorden & Perkel, 2023; Van Quaquebeke et al., 2025; Waduge et al., 2024). Scholarly work has shifted from questioning *whether* GenAI will matter to examining *how* it is transforming existing practices and which safeguards are needed to guide its development and deployment (Bordas et al., 2024; Kansal, 2024; McIntosh et al., 2025; Sætra, 2023; Van Quaquebeke et al., 2025).

Despite this potential, the use of GenAI presents notable limitations and risks for researchers across fields, primarily due to the possibility of false, misleading, or low-quality outputs. Reviews in higher education and professional domains warn that LLMs can propagate misinformation, introduce subtle inaccuracies, and blur boundaries around authorship and accountability if used uncritically (Bobula, 2024; Van Quaquebeke et al., 2025). Users may encounter AI “hallucinations”,

where responses contain fabricated or erroneous information presented with unwarranted confidence—for example, non-existent citations or misattributed findings, inaccurate summaries that add content not present in the source texts, or errors in generated data-analysis code that can lead to incorrect interpretations (Lorenz et al., 2024; Naik et al., 2024; Sun et al., 2024; Thorne, 2024). These concerns align with broader calls in the GenAI literature to attend to ethical implications, transparency, and the potential for harmful or misleading content (Dowling & Lucey, 2023; McIntosh et al., 2025; Sengar et al., 2024; Van Quaquebeke et al., 2025). Consequently, researchers are advised to treat GenAI as an assistive tool whose outputs must be critically assessed, verified against primary sources or alternative methods, and embedded within robust research and educational practices (Dowling & Lucey, 2023; McIntosh et al., 2025; Sengar et al., 2024; Van Quaquebeke et al., 2025).

In conclusion, although GenAI is a powerful research tool with rapidly expanding applications, it remains at a stage where careful implementation, institutional guidance, and human oversight are essential to harness its benefits while maintaining academic integrity. Because GenAI outputs can be variable in quality, researchers must decide when and how far to rely on them. This decision, in turn, may depend on whether they trust the system's outputs and the conditions under which it is deployed. In the next section, we examine the factors that build trust in GenAI and its critical role in shaping researchers' intention to use GenAI in their work.

2.2. Understanding the relevance of trust in GenAI

Trust is commonly defined as a directional relationship between two parties in which one party (the trustor) is willing to accept vulnerability to the actions of another (the trustee), based on the expectation that the trustee will act in ways that are beneficial or at least not harmful. In this sense, trust facilitates coordination and collaboration but always entails risk and uncertainty for the trustor, who cannot fully control or predict the trustee's behavior (Das & Teng, 2001; Jacovi et al., 2021; Misztal, 2013). For example, when a researcher relies on text generated by a GenAI application, they expose themselves to the risk that the output may contain false information. Crucially, this possibility of undesirable consequences must be recognized by the researcher; without an awareness of potential harm, reliance would reflect blind faith rather than trust (Jacovi et al., 2021).

Because humans have limited cognitive and affective resources, they cannot scrutinize every decision or piece of information in depth. Under these constraints, trust often functions as a mental heuristic or shortcut that allows interaction with complex systems to feel manageable, predictable, and worthwhile (Lukyanenko et al., 2022). Trust helps individuals move from constant monitoring and evaluation to a state in which they can proceed “as if” the trustee will behave appropriately, reducing the perceived need for continuous checking and control.

Research on interpersonal and organizational trust typically distinguishes between cognitive and affective trust. Cognitive trust refers to confidence in another's competence, reliability, and predictability, developed through accumulated knowledge of their behavior and performance over time (McAllister, 1995; Rempel et al., 1985). Affective trust, in contrast, is grounded in emotional bonds, perceived benevolence, and a sense of mutual care. It reflects the belief that the other party is concerned about one's welfare, not just capable or reliable. Applied to GenAI, cognitive trust is built when the system appears tangible, transparent, and consistently reliable in its outputs, whereas affective trust may arise more indirectly—through anthropomorphic design, perceived “honesty” in how limitations are communicated, or the sense that the system follows rules that align with the user's values (Glikson & Woolley, 2020; Riley & Dixon, 2024).

2.3. The relationship between perceived accuracy and the intention to use GenAI

The research model investigated in the two studies is shown in Fig. 1. A central relationship in major technology-adoption models such as the Technology Acceptance Model (TAM) and the Unified Theory of Acceptance and Use of Technology (UTAUT) is the link between expected performance gains and usage intentions. In TAM, *perceived usefulness* refers to the belief that a system improves one's job performance (F. D. Davis, 1989). In UTAUT, *performance expectancy* similarly captures the belief that using the system helps users achieve better outcomes or higher productivity performance (Venkatesh et al., 2003). Both constructs are consistently among the strongest predictors of intention to use a technology, and this pattern has also been found for GenAI (Baig & Yadegaridehkordi, 2025; Mun & Hwang, 2025; Zhang & Wang, 2025).

At the same time, for research work, "better performance" is not only about speed or convenience, but also about the *epistemic quality* of outputs (Van Quaquebeke et al., 2025). Perceived accuracy is conceptually distinct from both performance expectancy/usefulness and trust. Performance expectancy or usefulness concerns whether GenAI helps users perform their work more effectively or efficiently. Perceived accuracy, by contrast, concerns whether the system's outputs are judged to be correct and dependable. A GenAI tool may save time or support productivity while still generating inaccurate summaries, misleading interpretations, or false citations. Thus, perceived accuracy is not simply a sub-facet of performance expectancy/usefulness, but a conceptually separable judgment that may be especially relevant in domains where correctness is central to the task. At the same time, perceived accuracy is also distinct from trust. Accuracy reflects a judgment about the quality of the system's outputs, whereas trust reflects a willingness to rely on those outputs under conditions of uncertainty (Afroogh et al., 2024). We therefore position perceived accuracy as a system-quality cue that informs trust.

In the current research, we frame perceived accuracy as users' belief that GenAI will produce correct results for research-relevant tasks (e.g., summarizing literature, generating explanations, proposing analyses) under typical use conditions. We intentionally frame this as a *task-contextual belief* (accuracy for the kinds of tasks researchers use GenAI for), rather than a purely abstract belief about the tool in general. Prior work supports treating accuracy-related evaluations as a distinct construct: in fitness wearables, perceived accuracy is empirically separable from performance expectancy and can mediate its effect on adoption intentions (Rana et al., 2024). Relatedly, human-AI research shows that communicating an algorithm's accuracy calibrates users' expectations and can shape acceptance (Do et al., 2023). For researchers, inaccurate GenAI outputs can directly threaten the validity of conclusions and damage credibility; therefore, when GenAI outputs are perceived as accurate, researchers have stronger reasons to integrate the tool into their work. Hence, we propose:

H1. Perceived accuracy is positively related to the intention to use GenAI.

2.4. The mediating role of trust in GenAI

We anticipate that trust in GenAI will play an essential mediating

role in the relationship between perceived accuracy and usage intention. The degree to which trust in GenAI is established depends on users' perceptions of the system's trustworthiness (Afroogh et al., 2024; Y. L. Cui et al., 2025; Glikson & Woolley, 2020). Because the actual trustworthiness of AI systems cannot be directly observed, it is inferred from cues. Such cues include interface design, transparency about model uncertainty, the apparent quality and coherence of individual outputs, and any rationale or explanation provided for recommendations (Afroogh et al., 2024; van der Werff et al., 2021). Accuracy-related cues may arise subjectively through users' own interactions with the technology or be provided externally through evaluations and feedback on output quality (Schlicker & Langer, 2021). We propose trust as the psychological mechanism through which perceptions of accuracy become a willingness to accept vulnerability to the system's outputs. Perceived accuracy reflects an assessment that the system's outputs are likely to be correct. Trust, in turn, reflects the willingness to accept vulnerability by depending on those outputs despite uncertainty. Accordingly, we theorize perceived accuracy as an antecedent system-level cue that contributes to trust. When users observe that a GenAI system produces accurate content, they may become confident in its competence and more willing to accept vulnerability in relying on its suggestions or decisions (Y. L. Cui et al., 2025; Jacovi et al., 2021; Kaplan et al., 2023; Xu et al., n.d.). Conversely, visible inaccuracies undermine cognitive trust and increase the perceived need for external verification. Thus, in the context of GenAI, accuracy-related cues are not merely desirable system attributes; they are key psychological inputs to the formation and calibration of users' trust.

Studies consistently identify trust as a central predictor of intention to use AI in general (Choung et al., 2023; Dingel et al., 2024; Reinhardt et al., 2025; Wissawaswaengsuk et al., 2025) and GenAI in particular (Cao & Peng, 2024; Mun & Hwang, 2025). When users trust GenAI, they are prepared to act on its suggestions and integrate its content into their work, thereby turning perceived accuracy into a concrete intention to use the technology. This perspective aligns with the SOR theory, where perceived accuracy (stimulus) influences internal states like trust (organism), which in turn drive behavioral outcomes such as the intention to use (response) (Mehrabian & Russell, 1974). Additionally, this logic closely mirrors foundational trust models in information systems, in which users' beliefs about a system's *ability* or *competence* are a primary basis for trust, and trust in turn drives trusting behaviors (McKnight et al., 2002). Building on this reasoning and previous research, we hypothesize:

H2. The relationship between perceived accuracy of GenAI and the intention to use GenAI is mediated through trust in GenAI.

2.5. The moderating role of experience

Experience may diminish the influence that trust has on the intention to use a technology for two reasons. First, trust may reflect an anchoring belief that loses relevance with increasing experience. This idea resonates with work on the TAM, which applies the anchoring and adjustment framework from human decision making to technology adoption. According to this view, individuals initially form perceptions of a system based on broad cognitive anchors derived from their general beliefs about technology. With increasing experience, users adjust these

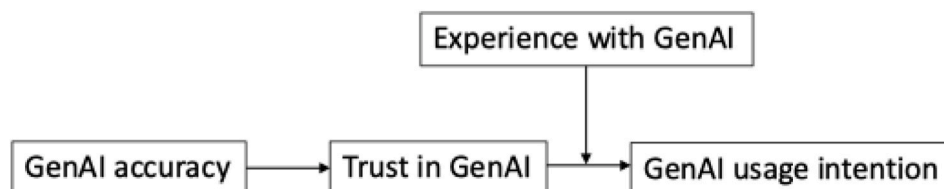


Fig. 1. The research model.

perceptions based on direct interaction, and the influence of such anchoring beliefs on key evaluations diminishes (Mussweiler & Strack, 2001; Venkatesh, 2000; Venkatesh & Bala, 2008; Wansink et al., 1998; Yadav, 1994). Applied to GenAI, this suggests that trust may be particularly influential when users have limited direct knowledge of the system and must act under uncertainty. As experience accumulates, users gain clearer feedback about the system's strengths and limitations, can rely more on their own evaluations of performance, and therefore depend less on trust as a substitute for unknowns (Chiu et al., 2012; Earle & Siegrist, 2006; Lankton et al., 2012; Siegrist & Cvetkovich, 2000).

Second, increasing experience may not simply reduce reliance on trust, but may also make trust more calibrated. In this case, experienced users may still trust GenAI, but in a more differentiated and conditional way that is tied to particular tasks, contexts, or levels of risk. A general trust judgment should become less predictive of overall usage intention as experience grows—not because trust disappears, but because it becomes more bounded and context-sensitive (McDermott & ten Brink, 2019). Thus, experience may weaken the trust–usage intention relationship either because users substitute direct evaluation for trust, or because they rely more on calibrated trust than overall generalized trust. In both cases, we expect that the predictive power of general cognitive trust for explaining behavioral intention will weaken as experience increases. To our knowledge, only one previous study has directly investigated this question, finding that the influence of trusting beliefs on continued Facebook use diminishes as users gain experience (Lankton et al., 2012).

Consistent with prior research that defines familiarity as accumulated experience (Alba & Hutchinson, 1987; Chen et al., 2011), we refer to ‘experience’ as capturing both subjective familiarity (Study 1) and subjective experience with a technology (Study 2). Based on TAM's anchoring and adjustment perspective and previous research, we propose:

H3. Experience with GenAI moderates the relationship between trust and the intention to use GenAI for research.

2.6. Overview, transparency, and declaration of generative AI use

To test our hypotheses, we conducted two online studies with a total sample size of $N = 402$ participants. In Study 1, we investigated the relationship between perceived accuracy of GenAI and the intention to use GenAI for research tasks, considering trust as a mediator. We also examined how experience with GenAI moderated the relationship between trust and usage intention. A longitudinal design with a one-week interval between assessments was employed to test the proposed moderated mediation model. Study 2 aimed to conceptually replicate the findings from Study 1. We manipulated accuracy by providing participants with vignettes indicating either low or high accuracy of recommendations generated by ChatGPT to further explore the link between accuracy and trust. Additionally, we used alternative measurements of trust and experience, following recommendations for conceptual replication studies (Fabrigar & Wegener, 2016). We again tested the full moderated mediation model, examining experience as a moderator in the relationship between trust in ChatGPT and the intention to use it for research tasks. The studies were not preregistered.

To support transparency, we share the survey materials, analysis code, and de-identified data at https://osf.io/mvcn5/overview?view_only=f476d962840f4b5a81bbba2962a6b34c, enabling reproduction of all results. Portions of the manuscript text were developed with assistance from an AI language model (Logically AI Assistant, based on OpenAI GPT-5.2). The tool was used to support writing and clarity (e.g., sentence-level rewriting, improving grammar). All AI-assisted text was reviewed, verified, and edited by the authors, who take full responsibility for the content.

3. Study 1

3.1. Methods study 1

3.1.1. Procedure and participants

After obtaining IRB approval (institution and reference number blinded for review), we conducted an online study using Qualtrics. Participants were recruited through Prolific, using specific custom screeners: highest education level completed (graduate degree such as MA, MSc, MPhil, or doctorate/PhD), work function (research), and industry (college, university, adult education, or research laboratories). We invited eligible participants to complete a brief screener survey to further identify those who have used AI-powered tools in their work. A total of 311 participants met the eligibility criteria and were invited to take part in the main study. We used a three-wave longitudinal design with a one-week interval between each wave, measuring perceived accuracy and experience (Wave 1), trust (Wave 2), and subsequent usage intentions (Wave 3) at separate time points to strengthen temporal ordering. A one-week lag between waves balances the need to capture relatively rapid changes in perceptions of accuracy and trust against participant burden and attrition (Dormann & Griffin, 2015).

Participants received £13.24 per hour as compensation for their participation in each survey. We received completed surveys from 232 participants at T1, 203 at T2, and 186 at T3. The average age of participants was 31.65 years ($SD = 6.64$), with 55.2% identifying as female. On average, participants had been working in research for 6.88 years ($SD = 5.41$).

The majority of participants worked at universities ($n = 179$, 77.2%), while the remainder ($n = 53$, 22.8%) were employed at private or governmental labs. Participants' roles included PhD students ($n = 113$, 48.7%), research scientists ($n = 48$, 20.7%), postdoctoral researchers ($n = 41$, 17.7%), professors ($n = 23$, 9.9%), and student research assistants ($n = 7$, 3.0%). Most participants conducted research in the Social Sciences ($n = 73$, 31.5%), followed by Natural Sciences ($n = 59$, 25.4%), Engineering and Technology ($n = 37$, 15.9%), Medical and Health Sciences ($n = 34$, 14.7%), Business and Management ($n = 17$, 7.3%), and other fields ($n = 12$, 5.2%). Nearly half of the sample dedicated over 80% of their working hours to research tasks ($n = 115$, 49.6%). We also inquired about the AI-powered tools used in their research (multiple answers possible). The most commonly mentioned tool was ChatGPT (cited 228 times, 56.6%), followed by Gemini ($n = 28$, 6.9%) and Copilot ($n = 23$, 5.7%). A comprehensive list of all AI-powered tools mentioned is available in the Online Appendix https://osf.io/mvcn5/overview?view_only=f476d962840f4b5a81bbba2962a6b34c (Table A1).

3.1.2. Measures

3.1.2.1. Perceived accuracy. We used the three-item accuracy subscale of the information quality scale to measure perceived accuracy of the recommendations generated by AI tools at T1 (Wixom & Todd, 2005). The items were rated on a seven-point Likert scale ranging from 1 (*strongly disagree*) to 7 (*strongly agree*). One example item is “The information provided by AI-powered tools is accurate.” ($\alpha = .74$).

3.1.2.2. Trust. Trust was measured at T2 with a three-item trust scale (Shin & Chan-Olmsted, 2023). The items were rated on a seven-point Likert scale ranging from 1 (*strongly disagree*) to 7 (*strongly agree*). One example item is “I would trust AI-powered tools for research-related tasks.” ($\alpha = .85$).

3.1.2.3. Behavioral intention to use GenAI. Behavioral intention to use GenAI for research tasks was measured at T3 with the three-item scale of the UTAUT (Venkatesh et al., 2003). The items were rated on a seven-point Likert scale ranging from 1 (*strongly disagree*) to 7 (*strongly agree*). One example item is “I intend to use AI-powered tools for my

research tasks in the next few months.” ($\alpha = .95$).

3.1.2.4. Experience. Experience was measured at T1 with one item: “Prior to participating in this study, how familiar were you with the use of AI-powered tools for research-related tasks?”. Answers were provided on a five-point Likert scale ranging from 1 (*very unfamiliar*) to 5 (*very familiar*). We used a single-item measure to capture participants’ overall perceived familiarity with AI-powered tools in a parsimonious way and to limit participant burden.

3.1.2.5. Control variables. We included age and gender (1 = female, 0 = other gender or prefer not to say) as demographic control variables. The percentage of research-related tasks (e.g., compared to teaching) could influence the interest in using GenAI for research. Thus, we considered the percentage of time participants dedicated to research as a control variable. Individuals in STEM fields may be more exposed to and familiar with advanced technologies, which might affect their trust and intention to use GenAI (Venkatesh & Davis, 2000). Additionally, the workplace, whether a university or a private/governmental lab, might impact access to resources and institutional support for using innovative technologies (Mahmood et al., 2001). Accordingly, we assessed the field of work (1 = STEM, 0 = other) and the workplace setting (1 = University, 0 = private or governmental labs) as potential control variables. Finally, since the acceptance of GenAI may be influenced by perceptions of general media coverage (D. Cui & Wu, 2021), we controlled for participants’ perceptions of media positivity concerning GenAI. Media perception was assessed based on a single item (Kleine et al., 2025): “Which of the following best describes your perception of media coverage on the use of AI-powered tools for research?”, with answer options ranging from 1 (*overwhelmingly negative*) to 5 (*overwhelmingly positive*). All control variables were measured at T1.

3.1.3. Data analysis

Statistical analyses were performed with R 4.4.3 (R Core Team, 2025) using the *lavaan* package (Rosseel, 2012) for structural equation modeling (SEM). First, a preliminary measurement model (including the multi-item scales accuracy, trust, and usage intention) was estimated with a confirmatory factor analysis (CFA) to assess the factor structure and psychometric adequacy of the measures. Standard errors were estimated using bootstrap resampling with 1000 iterations. This technique is effective for providing robust estimates, especially when data deviate from normality (Shrout & Bolger, 2002). Full Information Maximum Likelihood (FIML) was employed to handle missing data, leveraging all available data and avoiding listwise deletion, thus minimizing potential biases (Enders & Bandalos, 2001). Model fit was assessed using the chi-square test, comparative fit index (CFI ≥ 0.90),

Tucker and Lewis Index (TLI ≥ 0.90), root mean square error of approximation (RMSEA ≤ 0.08), and standardized root mean square residual (SRMR ≤ 0.08) (Hu & Bentler, 1999).

Second, for the main analyses, we fit the structural model, examining how accuracy predicted usage intention both directly and indirectly through trust, with experience moderating the path between trust and usage intention. We used 1000 bootstrap replication samples to estimate 95% bias-corrected confidence intervals (BootCIs). To estimate the interaction between the observed moderator (experience) and the latent variable (trust) in the moderated mediation model, the product indicator approach (Kenny & Judd, 1984) with the double-mean-centering strategy (Lin et al., 2010) was used. Effects were considered significant if the BootCIs excluded zero. Confirmation of moderated mediation was based on the index of moderated mediation (Hayes, 2015).

3.2. Results study 1

3.2.1. Descriptive statistics and preliminary results

Table 1 presents the means, standard deviations, and correlations among the study variables. The model variables were correlated to a small to moderate extent (e.g., $r = .09$ for the correlation between experience and accuracy; $r = .43$ for the correlation between usage intention and both trust and experience).

The results of the CFA indicated adequate fit to the data ($\chi^2(24) = 43.350, p = .009$, RMSEA = 0.06, CFI = 0.98, TLI = 0.98, SRMR = 0.05). The fit was better than the fit of an alternative model with trust and usage intention combined as one factor ($\chi^2(26) = 268.163, p < .001$, RMSEA = 0.20, CFI = 0.80, TLI = 0.72, SRMR = 0.14). We conclude that trust and usage intention are distinct from each other, indicating adequate discriminant validity. The substantially poorer fit of the one-factor alternative also reduces concern that the observed relationships were driven primarily by a single common method source. Furthermore, all standardized factor loadings exceeded 0.4, indicating adequate convergent validity (Brown, 2015).

3.2.2. Hypotheses tests

As the results did not vary when control variables were included or excluded, we report findings without them (Bernerth & Aguinis, 2016). Results with control variables are available in the Online Appendix (Table A2). Table 2 presents the results of the moderated mediation model. The overall model fit was adequate ($\chi^2(51) = 116.00, p < .001$, RMSEA = 0.07, CFI = 0.96, TLI = 0.94, SRMR = 0.06). H1 predicted a positive relationship between perceived accuracy and usage intention. Although perceived accuracy was positively related to usage intention at the zero-order level (Table 1; $r = .35, p < .05$), in the moderated mediation model, the direct path from accuracy to usage intention (c')

Table 1
Descriptive statistics and correlations (Study 1).

	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.
1. T1 Accuracy										
2. T2 Trust	.33									
3. T3 Usage	.35	.43								
4. T1 Experience	.09	.17	.43							
5. Age	.03	.11	.06	-.02						
6. Female gender ^a	.07	-.16	-.04	-.01	-.16					
7. Perc. research ^b	-.07	-.05	-.07	.06	-.13	.07				
8. Field ^c	-.01	-.03	-.02	-.06	.05	-.20	.18			
9. Workplace ^d	-.04	-.01	-.14	.02	.04	-.02	-.12	-.19		
10. Media	.08	.05	.03	.10	.01	-.06	-.07	.04	-.02	
<i>M</i>	4.39	4.37	5.94	4.06	31.65	0.55	4.04	0.56	0.77	3.30
<i>SD</i>	1.08	1.17	1.02	0.80	6.64	0.50	1.20	0.50	0.42	1.03

Note. $N = 186$ – 232 . All correlations $\geq |0.13|$ are significant at $p < .05$. All correlations are based on uncentered variables. Usage = usage intention.

^a Code for gender: 0 = male, non-binary, or prefer not to say, 1 = female.

^b Code for percentage research: 1 = 0–20%, 2 = 21–40%, 3 = 41–60%, 4 = 61–80%, 5 = 81–100%.

^c Code for field: 0 = non-STEM, 1 = STEM.

^d Code for workplace: 0 = private/governmental lab, 1 = university.

Table 2

Results of the structural equation model (SEM) (Study 1).

	β	SE	z	p	95%BootCI
Path a: Accuracy → Trust	.373	.089	4.212	<.001	[.199, .547]
Path b: Trust → Usage	.353	.074	4.772	<.001	[.208, .498]
Path c' (direct effect): Accuracy → Usage	.166	.097	1.714	0.087	[-.024, .356]
Indirect: Accuracy → Trust → Usage	.132	.042	3.122	0.002	[.049, .215]
Interaction: Trust × Experience → Usage	-.281	.097	-2.886	.004	[-.472, -.090]
Path b: Trust → Usage; Experience:					
Low (one SD below mean = -0.801)	.578	.111	5.198	<.001	[.360, .796]
Medium (mean = 0)	.353	.074	4.772	<.001	[.208, .498]
High (one SD above mean = 0.801)	.128	.104	1.240	.215	[-.075, .332]
Index of Moderated Mediation	-.105	.042	-2.476	.029	[-.179, -.017]

Note. $N = 232$. Usage = usage intention; β = Standardized Beta; SE = Standard Error; BootCI = Bootstrap Confidence Interval.

was non-significant (95% BootCI [-.024, .356]) once trust and the trust × experience interaction were included. This finding indicates that perceived accuracy did not explain additional variance in usage intention beyond the indirect pathway through trust. Accordingly, H1 was supported at the zero-order level but not as a direct effect in the full model. A significant indirect effect of accuracy on usage intention via trust was observed (95% BootCI [.049, .215]), supporting H2. H3 was supported as the Trust × Experience interaction predicting usage intention was significant (95% BootCI [-.472, -.090]), indicating that the trust–intention relationship weakened as experience increased. Consistent with this pattern, the conditional indirect effect of accuracy on usage intention via trust was significant at low and medium, but not high levels of experience, and the index of moderated mediation was significant (95% BootCI [-.179, -.017]).

3.3. Discussion study 1

The findings of Study 1 offer partial support for our hypotheses concerning the relationships between perceived accuracy, trust, and the intention to use GenAI for research tasks. H1 predicted a positive relationship between perceived accuracy and usage intention. This prediction was supported at the zero-order level, indicating that researchers who perceived GenAI outputs as more accurate also reported stronger intentions to use such tools. However, in the full moderated mediation model, the direct path from accuracy to usage intention was no longer significant once trust (and its interaction with experience) was included, suggesting that perceived accuracy did not explain additional variance in usage intention beyond the trust-based mechanism.

The results support H2 by showing a mediation pattern in which trust accounts for the relationship between perceived accuracy and usage intention. This pattern is consistent with the view that trust functions as an important psychological process linking accuracy perceptions to willingness to integrate GenAI into research workflows. The mediating role of trust aligns with theoretical expectations that trust helps users manage uncertainty associated with opaque or “black box” AI systems.

In support of H3, the moderated mediation model revealed a significant trust × experience interaction predicting usage intention. As experience with GenAI increased, the influence of trust on usage intention weakened. This finding resonates with prior research (e.g., Lankton et al., 2012) and suggests that trust may be particularly influential in earlier adoption phases, whereas more experienced users rely less on trust when forming intentions to use GenAI.

Despite these insights, several limitations in Study 1 underscore the

need for continued investigation. First, the study employed a general measure of perceived accuracy without accounting for contextual nuances across specific GenAI tools. Although the observed associations were consistent with the proposed model, variability in system capabilities and output quality across tools warrants further study. Second, because Study 1 was observational, causal inferences about the proposed process remain limited. Although the longitudinal design strengthens temporal ordering by measuring accuracy (T1), trust (T2), and intention (T3) at separate time points, it does not establish a causal process and cannot rule out omitted-variable explanations or reciprocal relationships. Study 2 addresses this limitation by manipulating accuracy via vignettes. Finally, the reliance on self-reported perceptions of accuracy may introduce bias (e.g., driven by prior attitudes or recent experiences). Study 2 mitigates this concern by providing standardized accuracy cues in the vignettes (experimental manipulation). Together, Study 2 was designed to replicate and extend Study 1, enabling a more comprehensive understanding of the processes shaping researchers' intentions to use GenAI.

4. Study 2

4.1. Methods study 2

4.1.1. Procedure and participants

Study 2 recruited participants from Prolific's prescreened population of users who met the same eligibility criteria as Study 1 (graduate-level education; work function: research; industry: higher education or research laboratories). To ensure an independent sample, we excluded all users who had participated in Study 1. In total, 179 invited users completed the survey. Consistent with the study's a priori focus on ChatGPT usage, we included only participants who indicated that they had previously used ChatGPT. Nine participants did not meet this criterion and were excluded, resulting in a final sample of 170 participants. Participants were compensated at a rate of £9.60 per hour. The average age of participants was 32.64 years ($SD = 7.68$), with $n = 88$ (51.8%) identifying as female. On average, participants had 7.62 years of research experience ($SD = 6.69$). Most participants were employed at universities ($n = 140$, 82.4%), while the rest worked in private or governmental labs ($n = 30$, 17.6%). Roles included PhD students ($n = 73$, 42.9%), research scientists ($n = 42$, 24.7%), postdoctoral researchers ($n = 38$, 22.4%), professors ($n = 15$, 8.8%), and student research assistants ($n = 2$, 1.2%). Research areas included Social Sciences ($n = 61$, 35.9%), Natural Sciences ($n = 41$, 24.1%), Medical and Health Sciences ($n = 29$, 17.1%), Engineering and Technology ($n = 22$, 12.9%), Business and Management ($n = 8$, 4.7%), and other fields ($n = 9$, 5.3%). Nearly half of the sample dedicated over 80% of their work hours to research tasks ($n = 82$, 48.2%).

The final analytic sample comprised $N = 170$ participants. Of these, $n = 83$ were included in the analyses in the low-accuracy condition and $n = 87$ in the high-accuracy condition. In these conditions, participants viewed an interaction with ChatGPT along with feedback regarding the accuracy of its responses. Scripts and prompts are available in the Online Appendix (Fig. A1). Participants were then asked to assess the accuracy of the recommendations (for manipulation checks), their trust in ChatGPT, their intention to use ChatGPT, and to provide demographic information.

4.1.2. Experimental manipulation and measures

4.1.2.1. Vignettes. To reinforce perceptions of accuracy, the vignettes depicted interactions between a user and ChatGPT in three contexts: advising on journal selection, checking text consistency, and selecting references for a scientific claim. Following each interaction, an evaluation of the response was presented. The order in which the three contexts were presented was randomized. In the high accuracy condition,

Table 3

Descriptive statistics and correlations (Study 2).

	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.
1. Accuracy ^a										
2. Trust	.43									
3. Usage	.12	.43								
5. 4. Experience	-.01	.28	.53							
5. Age	.04	-.07	.07	-.09						
6. Female gender ^b	.05	-.01	-.08	-.23	-.03					
7. Perc. research ^c	-.06	-.05	-.07	-.03	-.20	.01				
8. Field ^d	-.07	.11	-.06	.02	-.12	-.04	.17			
9. Workplace ^e	.15	.08	.12	-.02	.04	.03	-.07	-.14		
10. Media	.11	.24	.07	.01	.11	-.11	-.04	-.02	.05	
<i>M</i>	.51	2.86	4.97	2.65	32.64	.52	4.05	.53	.82	2.87
<i>SD</i>	.50	1.19	1.68	1.08	7.68	.50	1.16	.50	.39	1.00

Note. $N = 170$. All correlations $\geq |0.15|$ are significant at $p < .05$. All correlations are based on uncentered variables. Usage = usage intention.

^a Code for experimental condition: 0 = low accuracy, 1 = high accuracy.

^b Code for gender: 0 = male, non-binary, or prefer not to say, 1 = female.

^c Code for percentage research: 1 = 0-20%, 2 = 21-40%, 3 = 41-60%, 4 = 61-80%, 5 = 81-100%.

^d Code for field: 0 = non-STEM, 1 = STEM.

^e Code for workplace: 0 = private/governmental lab, 1 = university.

responses were evaluated positively; in the low accuracy condition, they were evaluated negatively. More detailed content of the vignettes is available in the Online Appendix (Fig. A1). As a manipulation check, participants rated the accuracy of the recommendations by answering, "Based on the evaluation of the output, how well did ChatGPT perform?" Responses were provided using a slider scale ranging from 0 (*poor accuracy*) to 100 (*perfect accuracy*).

4.1.2.2. Trust. Trust was measured after all three vignettes were presented (assessed for overall impression). It was measured with an adapted four-item trust in AI scale (Gillath et al., 2021). The items were rated on a seven-point Likert scale ranging from 1 (*strongly disagree*) to 7 (*strongly agree*). One example item is "How likely are you to trust ChatGPT's suggestions?" ($\alpha = .93$).

4.1.2.3. Behavioral intention to use GenAI. As in Study 1, behavioral intention to use GenAI for research tasks was measured with the adapted three-item scale of the UTAUT (Venkatesh et al., 2003). One example item is "I intend to use ChatGPT for my research tasks in the next few months." The items were rated on a seven-point Likert scale ranging from 1 (*strongly disagree*) to 7 (*strongly agree*; $\alpha = .99$).

4.1.2.4. Experience. Experience was measured with one item: "How experienced would you say you are using ChatGPT in your work?". Answers were provided on a five-point Likert scale ranging from 1 (*not experienced at all*) to 5 (*extremely experienced*).

4.1.2.5. Control variables. We included the same control variables as in Study 1.

Table 4

Results of the structural equation models (SEM) (Study 2).

	β	SE	z	p	95%BootCI
Path a: Accuracy ^a → Trust	.447	.065	6.920	<.001	[.320, .573]
Path b: Trust → Usage	.335	.085	3.950	<.001	[.169, .501]
Path c' (direct effect): Accuracy ^a → Usage	-.043	.076	-.563	.573	[-.193, .107]
Indirect: Accuracy ^a → Trust → Usage	.150	.044	3.381	.001	[.063, .236]
Interaction: Trust × Experience → Usage	-.211	.078	-2.718	.007	[-.364, -.059]
Path b: Trust → Usage; Experience:					
Low (one SD below mean = -1.078)	.563	.123	4.575	<.001	[.322, .804]
Medium (mean = 0)	.335	.085	3.950	<.001	[.169, .501]
High (one SD above mean = 1.078)	.107	.115	.927	.354	[-.119, .333]
Index of Moderated Mediation	-.094	.036	-2.613	.009	[-.165, -.024]

Note. $N = 170$. Usage = usage intention; β = Standardized Beta; SE = Standard Error; BootCI = Bootstrap Confidence Interval.

^a Code for experimental condition: 0 = low accuracy, 1 = high accuracy.

4.1.3. Data analysis

We used the same data analysis approach as in Study 1, with accuracy entered as a dummy-coded independent variable.

4.2. Results study 2

4.2.1. Descriptive statistics and preliminary results

Table 3 provides the means, standard deviations, and correlations among the study variables. The model variables showed small to moderate correlations. Interestingly, the experimental condition was unrelated to usage intention. For the manipulation check, we compared the mean perceived accuracy across all three contexts between the two experimental groups. A statistically significant difference was found in perceived accuracy, $t(166.97) = -30.532$, $p < .001$, with means of $M = 16.9$ ($SD = 14.2$) for the low accuracy and $M = 85.3$ ($SD = 14.9$) for the high accuracy group.

The results of the preliminary CFA (measurement model) indicated a good fit to the data ($\chi^2(18) = 17.921$, $p = .161$, RMSEA = 0.047, CFI = 1.00, TLI = 1.00, SRMR = 0.02). Importantly with respect to common method concerns, this fit was substantially better than that of an alternative one-factor model in which trust and usage intention were combined into a single factor ($\chi^2(14) = 848.212$, $p < .001$, RMSEA = 0.59, CFI = 0.44, TLI = 0.15, SRMR = 0.21). This pattern suggests that the association between trust and usage intention is unlikely to be attributable solely to a single common response tendency. Furthermore, all standardized factor loadings exceeded 0.7, indicating good convergent validity (Brown, 2015).

4.2.2. Hypotheses tests

We fit the moderated mediation model to the data. The results are

shown in Table 4. Because the results did not differ when control variables were included versus excluded, we report results without control variables (Bernerth & Aguinis, 2016). Results with control variables are available in the Online Appendix (Table A3). The overall model fit was adequate ($\chi^2(48) = 61.876, p = .086, RMSEA = 0.04, CFI = 0.99, TLI = 0.99, SRMR = 0.076$). H1 predicted a positive relationship between accuracy and usage intention; however, this was not supported in Study 2, as the zero-order association between accuracy condition and usage intention was non-significant (Table 3; $r = .12, p > .05$). Consistent with this, the direct path from accuracy to usage intention (c') was also non-significant in the moderated mediation model (95% BootCI [-.193, .107]) once trust and the trust $\times$ experience interaction were included. In support of H2, the indirect effect of accuracy on usage intention via trust was significant (95% BootCI [.063, .236]). Moreover, the interaction effect of Trust $\times$ Experience was significant (95% BootCI [-.364, -.059]), supporting H3. The conditional indirect effect was significant at low and medium, but not high levels of experience, and the index of moderated mediation was significant (95% BootCI [-.165, -.024]), indicating that the indirect effect through trust depends on experience.

5. Overall discussion

The current research set out to examine why researchers intend to use GenAI for their work, focusing on perceived accuracy as a key antecedent, trust as a mediator, and experience as a moderator. Across a three-wave longitudinal survey (Study 1) and a vignette experiment (Study 2), we found converging evidence that higher perceived accuracy was associated with greater trust in GenAI in Study 1, and that experimentally induced accuracy increased trust in Study 2. In both the longitudinal model (accuracy at T1 $\rightarrow$ trust at T2 $\rightarrow$ usage intention at T3) and the experiment (accuracy manipulation $\rightarrow$ trust $\rightarrow$ usage intention), trust emerged as the central psychological mechanism linking accuracy to usage intentions, as reflected in significant indirect effects. With regard to the overall accuracy–intention association, Study 1 showed a positive zero-order relationship, whereas Study 2 did not show a significant bivariate association, suggesting that accuracy cues do not uniformly translate into higher intended use at the overall level. Importantly, however, in both studies the direct path from accuracy to usage intention was non-significant once trust (and its interaction with experience) was included, consistent with the interpretation that accuracy primarily influences usage intention via trust rather than through a direct effect. The absence of a direct zero-order effect in Study 2 deserves attention. One plausible explanation is that the experimental sample consisted exclusively of prior ChatGPT users. Under these conditions, experimentally varying accuracy may have had limited leverage to further increase reported usage intention directly. More importantly, however, this pattern strengthens the mediation logic of the present model. It suggests that accuracy information may not be behaviorally consequential in a simple, immediate way. Rather, the findings are consistent with the interpretation that researchers translate accuracy cues into intentions to use GenAI after these cues shape a trust judgment about whether the system can be relied upon despite uncertainty. In this sense, trust is not merely an additional predictor of usage intention, but the key mechanism through which performance-related cues become behaviorally relevant.

We also found support for experience as a moderator of the relationship between trust and usage intention: as experience increases, the effect of trust on usage intention diminishes, suggesting that more experienced users base their usage intentions less on trust. Together, these findings position perceived accuracy as a central stimulus shaping trust in GenAI and show that the extent to which trust translates into usage intention depends on users' experience.

5.1. Theoretical implications

The moderated mediation model is conceptually consistent with SOR

logic, which emphasizes that external stimuli affect behavior primarily through internal psychological states (T. R. V. Davis & Luthans, 1980). In this framework, *perceived accuracy* functions as a stimulus; *trust* is the organism state; and *usage intention* is the behavioral consequence. The consistent finding across both studies that the direct accuracy–intention path becomes non-significant once trust is included, while the indirect path through trust is significant, supports a mediation structure in which accuracy matters for usage intention *because* it informs trust. Theoretically, this suggests that models of GenAI usage intention in research should explicitly position trust as a mechanistic link between specific system qualities (e.g., accuracy) and behavioral intentions, rather than treating trust as a predictor alongside performance expectancy or perceived usefulness (cf. Choudhury & Shamszare, 2024).

Our findings align with other research based on SOR theory in the context of AI technologies. For instance, it has been shown that trust fully mediates the relationship between AI-enabled personalization (stimulus) and consumer engagement (response) in social media marketing (Teepapal, 2025). Similar to the current research, the authors found no direct relationship between AI-enabled personalization and consumer engagement; however, there was an indirect effect through trust (Teepapal, 2025). Another study found that trust mediated the relationship between perceptions of change due to the introduction of AI in human resource management and the adoption of AI (Arora & Mittal, 2024). The current studies complement previous research by examining the mediation link in the context of researchers using GenAI. Moreover, by using longitudinal data and an experimental setting, we provide further evidence for trust as a mediator between accuracy and usage intention.

The distinction between general performance beliefs and perceived accuracy is particularly relevant in the context of scientific work. Prior work on AI-generated outputs has shown that even seemingly coherent outputs can contain subtle but consequential inaccuracies, such as non-existent citations or incorrect study details (Lorenz et al., 2024; Naik et al., 2024; Sun et al., 2024; Thorne, 2024). Our results suggest that, for researchers, accuracy is a *primary epistemic criterion* for deciding whether a GenAI tool is a viable partner in knowledge production. From a theoretical standpoint, the findings imply that technology acceptance models applied to research settings should (a) explicitly separate perceived accuracy from general usefulness or performance expectancy and (b) model accuracy primarily as an input to trust rather than directly to usage intention.

A major contribution of the current work is the demonstration that the predictive power of trust decreases as users gain experience with GenAI. This finding advances theorizing on trust in human–technology interaction by suggesting two possible interpretations of why the trust–intention link weakens with experience. First, the results are consistent with a trust-as-anchor view: for novices, trust operates as an initial heuristic that reduces the cognitive load of evaluating a complex, opaque system. As users accumulate experience, they can draw more on direct knowledge of the system's capabilities and limitations, reducing their reliance on trust as a general basis for deciding whether to use GenAI. This aligns with broader accounts in technology research emphasizing that initial global evaluations give way to more differentiated assessments as users gain experience or insight into how recommendations are generated (Karahanna et al., 1999; Lee & Xia, 2011). The weaker trust–intention relationship among more experienced users may indicate that, with repeated use, researchers become less dependent on global trust as a heuristic and more likely to rely on selective, task-sensitive judgments about when GenAI should and should not be used.

Second, the results may also reflect a calibrated-trust interpretation. Increasing experience does not necessarily make trust less important; rather, it may make trust more differentiated, conditional, and task-specific. Experienced users may still rely on trust, but less in the form of a broad, global judgment that predicts overall usage intention across situations. If so, the weaker trust–intention relationship at higher levels

of experience would indicate not simply substitution of direct evaluation for trust, but a shift from general heuristic trust to more bounded and calibrated trust. In our research, this is reflected in the moderated mediation pattern: the accuracy → trust → intention pathway is strong for low- and medium-experience researchers but negligible for highly experienced ones. Such a pattern is consistent with the broader literature on calibrated trust and appropriate reliance, which emphasizes that effective human–AI interaction depends on matching reliance to actual system capability rather than defaulting to acceptance or rejection (Lee & See, 2004). It is also consistent with emerging work on critical thinking in AI use showing that users higher in critical thinking are more likely to verify AI-generated statements, use broader corroboration strategies, and seek corroboration before trusting AI-generated information (Lau et al., 2025). Our findings suggest that experience may reduce the influence of generalized trust on usage intention because experienced users are better positioned to supplement trust with verification, corroboration, and reflective oversight.

Theoretically, this suggests that in the context of researchers adopting GenAI, the SOR should be applied as a conditional process model, rather than a "one-size-fits-all" mediation model. Specifically, the SOR mechanism (accuracy → trust → intention) is stage-dependent. As users gain experience, the link between trust and intention weakens, indicating that trust (organism) affects usage intention (response) differently for novices compared to experts. This aligns with other SOR research that incorporates moderators, acknowledging variable rather than uniform effects. For example, a study on GenAI adoption in the workplace found that personality traits, particularly conscientiousness, moderate the trust and adoption intention relationship, highlighting the role of individual differences (Dalvi-Esfahani et al., 2025). Another study revealed a negative moderation effect of perceived performance risk on the relationship between attitude toward and intention to use a Clinical Decision Support System (Dalvi-Esfahani et al., 2023). The current research is the first to explore the moderating role of experience within an SOR framework in relation to GenAI usage intentions. The findings suggest that models of AI acceptance and usage intention should view trust not as a universal resource but as a context-dependent mechanism whose importance varies with user experience. This has implications for how we think about "calibrated trust": for novices, the challenge is to avoid both unwarranted distrust (disuse) and overtrust (misuse). For experienced users, the focus may shift to preventing routine overreliance even when errors are detected. More research on calibrated trust and user experience is needed to explore these issues further.

5.2. Limitations and directions for future research

Despite its contributions, the present research has several limitations that point to recommendations for future work. First, although the three-wave design improved temporal ordering by measuring perceived accuracy, trust, and usage intention at separate time points, Study 1 remained observational. As such, it cannot establish a causal process. The Study 1 findings may be interpreted as supporting temporal ordering and a pattern of associations consistent with the proposed model, rather than definitive evidence that perceived accuracy causally shapes trust and subsequent usage intention. Study 2 provides a stronger basis for causal inference with respect to the effect of the accuracy manipulation on trust, but future research should use stronger longitudinal or panel designs, ideally with repeated measures of trust and usage intention, to test reciprocal and change-based processes more directly. Relatedly, Study 1 used a three-wave design with one-week lags, and Study 2 was a cross-sectional experiment. While these designs capture short-term dynamics between accuracy perceptions, trust, and intentions, they do not capture long-term adaptation processes or the evolution of GenAI systems themselves. As GenAI systems are updated and integrated into more platforms, both perceived and actual accuracy may change, potentially reshaping trust and usage patterns over time.

Longitudinal panel studies spanning months or years could examine how repeated exposure, institutional policies, and evolving capabilities interact to shape trust and usage intention.

Second, most key constructs—perceived accuracy (Study 1), trust, experience, and usage intention—were assessed via self-report. Future work could integrate behavioral metrics such as log data from AI tools, eye-tracking for verification behavior, or experimentally controlled tasks where researchers complete real literature searches or data analyses with GenAI assistance, subsequently examining the effects of these interactions on trust and usage intention. Additionally, it would strengthen the conclusions to examine whether the observed mechanism also predicts actual GenAI use, given that prior work on ChatGPT adoption explicitly distinguishes trust, intention, and actual use (Choudhury & Shamszadeh, 2023). An additional consideration is common method variance. Although the studies incorporated several safeguards, this concern cannot be fully excluded. In Study 1, predictor, mediator, and outcome were temporally separated across three waves, which reduces same-session reporting effects. In Study 2, the focal predictor was experimentally manipulated and randomly assigned, and the measurement model showed clear separation between trust and usage intention. However, because trust and usage intention in Study 2 were both assessed by self-report in the same session, some shared method influence remains possible. Future research should further address this issue by temporally separating mediator and outcome measures, using alternative response formats or sources, and incorporating behavioral indicators of GenAI use.

Third, an additional limitation concerns construct coverage. Although we conceptually distinguished perceived accuracy from usefulness or performance expectancy, we did not measure those established constructs directly. As a result, the present studies cannot determine whether perceived accuracy explains unique variance in trust or usage intention beyond these established constructs. Future research should test this more directly by including perceived accuracy alongside perceived usefulness and/or performance expectancy in the same model and comparing their relative and incremental effects on trust and usage intention. Relatedly, we used a general measure of trust, not differentiating between cognitive and affective trust. While our data clearly show that trust mediates the accuracy–intention pathway, they do not address whether a specific kind of trust matters most. For instance, researchers might be willing to use a system they see as highly competent but normatively misaligned (e.g., one that encourages questionable research practices) only under strong oversight, whereas they might show more autonomous reliance on systems seen as both competent and ethically aligned. Future studies could separate cognitive and affective dimensions of trust and test whether they differentially predict GenAI use in research.

Fourth, experience was measured with a single self-report item capturing subjective familiarity with AI-powered tools (Study 1) and experience with ChatGPT (Study 2). Although single-item measures of experience are common (Komiak & Benbasat, 2006; O'Brien et al., 2015), this operationalization remains limited because it may blur conceptually related but not identical aspects of prior interaction with the technology. In particular, rather than isolating concrete accumulated use, our measure relied on respondents' subjective sense of how familiar or experienced they were with the system. As a result, the measure may conflate distinct dimensions of experience, including depth of functional knowledge (Schulenberg & Melton, 2008), frequency of use, and length of exposure (Potosky & Bobko, 1998). This means that the moderation results should be interpreted as reflecting the role of perceived experience or familiarity rather than demonstrated competence or objectively accumulated use. Future research should therefore move beyond a single-item operationalization and capture experience multi-dimensionally, by using scales that distinguish breadth, intensity, duration, task-specific experience, incorporating objective indicators such as usage logs where feasible. Additionally, future research should more directly disentangle why the predictive

strength of trust declines with experience. More generally, the present studies did not directly measure calibrated trust or appropriate reliance. Although our findings are consistent with the idea that experienced users may show more bounded and better-calibrated reliance, we did not assess whether participants' trust tracked actual system reliability, nor did we measure verification, corroboration, or other forms of critical oversight. Future research should therefore examine calibration more directly by combining perceived trust with objective indicators of system performance, reliance behavior, and verification behavior. Such work would help determine whether weaker dependence on general trust among experienced users reflects greater critical thinking in AI use, more appropriate reliance, or both (Lau et al., 2025; Lee & See, 2004).

Fifth, an additional limitation concerns the scope of generalizability. Although we refer throughout to “researchers,” our samples were recruited through Prolific and drawn specifically from higher-education and research-laboratory contexts. In both studies, most participants worked at universities, and Study 1 in particular was weighted toward early-career researchers, especially PhD students, research scientists, and postdoctoral researchers. Accordingly, the present findings should be interpreted as most directly generalizing to researchers in academic or adjacent research settings, rather than equally to all career stages or institutional environments. It remains possible that the role of accuracy, trust, and experience differs for senior faculty, principal investigators, industry researchers, or other professional groups whose incentives, autonomy, and exposure to GenAI differ from those represented here. Future research should therefore test the model across more diverse career stages, organizational settings, and non-academic research environments.

Finally, it is also important to consider the study context: researchers typically self-determine whether and how to use GenAI in their work. The proposed model may therefore generalize less well to settings in which AI use is mandatory (e.g., required by the organization or a direct supervisor). In such mandated contexts, it may be more informative to test whether trust mediates relationships between perceived accuracy/performance and other consequential outcomes—such as work engagement or counterproductive behaviors.

5.3. Practical recommendations

Building on our theoretical model and empirical findings, several practical recommendations emerge. Most importantly, our results suggest that institutions and developers should aim to support calibrated trust in GenAI rather than simply trying to maximize trust. Although the predictive power of general trust decreases as users gain experience, this does not mean that trust becomes unimportant for experienced users. Rather, trust may become more differentiated, conditional, and task-specific. Practical interventions should therefore be designed to promote warranted trust.

In this sense, segmented interventions may be useful. For less experienced researchers, training should focus on building appropriately cautious trust by combining transparency about GenAI's strengths with explicit instruction in verification, source checking, and recognition of hallucinations or other inaccuracies. For more experienced researchers, interventions should focus less on boosting general trust and more on maintaining well-calibrated reliance, for example by reinforcing task-specific evaluation, encouraging continued verification in high-stakes contexts, and reducing the risk of routine overreliance (Liebherr et al., 2026; Okamura & Yamada, 2020).

Given that trust plays a larger role in novice intentions, institutions should avoid leaving early-career researchers to experiment with GenAI entirely on their own. Instead, they could encourage supervised use—for example, joint sessions where supervisors and trainees work through real tasks with GenAI while explicitly discussing accuracy, verification, and ethical boundaries. At the same time, because experienced users may rely on GenAI in a more routine and task-dependent way, institutional guidance should also help them periodically reassess when

reliance is appropriate and when closer scrutiny is required.

To mitigate risks of overreliance, especially among less experienced users, GenAI tools used in research could implement conservative defaults: warnings when answers are not grounded in verifiable sources, prompts encouraging users to cross-check key outputs, or optional “verification modes” that automatically scan references against bibliographic databases. This is consistent with the broader goal of supporting calibrated trust—trust that is warranted by the system's demonstrated performance and bounded by appropriate human oversight (Liebherr et al., 2026).

6. Conclusion

Across a longitudinal field study and an experimental vignette study with researchers, this work suggests that perceived accuracy of GenAI is linked to intentions to use such tools primarily through trust, and that this mediating effect of trust is moderated by users' experience. For novices and intermediate users, trust is the psychological hinge that translates accuracy cues into willingness to rely on GenAI; for highly experienced users, intentions become less dependent on trust.

Theoretically, these findings refine technology adoption models by positioning trust as a central mediating mechanism linking specific system qualities (accuracy) to behavioral intentions, and by demonstrating that the function of trust is contingent on experience. Practically, they underscore the need for calibrated—not maximal—trust in GenAI, supported by transparent system design and targeted training. As AI continues to reshape how research is conducted, the challenge is not merely to increase researchers' willingness to use these tools, but to ensure that any use is informed, critical, and compatible with scientific standards.

CRedit authorship contribution statement

Anne-Kathrin Kleine: Writing – review & editing, Writing – original draft, Visualization, Validation, Project administration, Methodology, Formal analysis, Data curation, Conceptualization. **Insa Schaffernak:** Writing – review & editing, Writing – original draft, Validation, Methodology, Conceptualization. **Eva Lerner:** Writing & editing, Resources, Supervision, Project administration, Funding acquisition.

Informed consent statement

The privacy rights of all human participants were observed throughout the study, and informed consent was obtained from all participants prior to their involvement in the research. Participants were provided with appropriate information about the study and were free to decline participation or withdraw at any time without penalty.

Ethical approval statement

All procedures performed in this study involving human participants were conducted in accordance with relevant laws, institutional guidelines, and ethical standards, and were approved by the ethical committee “Gemeinsame Ethikkommission der Hochschulen Bayern” under reference number **GEHBA-202401-V-159-R**.

Funding statement

This work was supported by the Volkswagen Foundation [grant number 98 525].

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Anne-Kathrin Kleine reports financial support was provided by

Volkswagen Foundation. Eva Lerner reports financial support was provided by Volkswagen Foundation. Anne-Kathrin Kleine reports a relationship with Volkswagen Foundation that includes: funding grants. Eva Lerner reports a relationship with Volkswagen Foundation that includes: funding grants. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The research was funded by the VW Foundation (“Human-AI-Interaction in Healthcare: Identifying Factors Contributing to Clinical Utility”; Grant Number: 98 525). The sponsor was not involved in the choice of the study design, data collection, data analysis, and interpretation of results.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.chbah.2026.100310>.

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), 1568. <https://doi.org/10.1057/s41599-024-04044-8>
- Alba, J. W., & Hutchinson, J. W. (1987). Dimensions of consumer expertise. *Journal of Consumer Research*, 13(4), 411–454. <https://doi.org/10.1086/209080>
- Arora, M., & Mittal, A. (2024). Employees' change in perception when artificial intelligence integrates with human resource management: A mediating role of AI tech trust. *Benchmarking: An International Journal*, 32(6), 1884–1908. <https://doi.org/10.1108/BLJ-11-2023-0795>
- Baig, M. I., & Yadegaridehkordi, E. (2025). Factors influencing academic staff satisfaction and continuous usage of generative artificial intelligence (GenAI) in higher education. *International Journal of Educational Technology in Higher Education*, 22(1), 5. <https://doi.org/10.1186/s41239-025-00506-4>
- Barros, A., Prasad, A., & Šliwa, M. (2023). Generative artificial intelligence and academia: Implication for research, teaching and service. *Management Learning*, 54(5), 597–604. <https://doi.org/10.1177/13505076231201445>
- Bernerth, J. B., & Aguinis, H. (2016). A critical review and best practice recommendations for control variable usage. *Personnel Psychology*, 69(1), 229–283. <https://doi.org/10.1111/peps.12103>
- Biyela, S., Dihal, K., Gero, K. I., Ippolito, D., Menczer, F., Schäfer, M. S., & Yokoyama, H. M. (2024). Generative AI and science communication in the physical sciences. *Nature Reviews Physics*, 6(3), 162–165. <https://doi.org/10.1038/s42254-024-00691-7>
- Bobula, M. (2024). Generative artificial intelligence (AI) in higher education: A comprehensive review of challenges, opportunities, and implications. *Journal of Learning Development in Higher Education*, 30. <https://doi.org/10.47408/jldhe.vi30.1137>
- Bordas, A., Le Masson, P., Thomas, M., & Weil, B. (2024). What is generative in generative artificial intelligence? A design based perspective. *Research in Engineering Design*, 35(4), 427–443. <https://doi.org/10.1007/s00163-024-00441-x>
- Brown, T. A. (2015). *Confirmatory factor analysis for applied research* (2nd ed.). Guilford Publications.
- Burger, B., Kanbach, D. K., Kraus, S., Breier, M., & Corvello, V. (2023). On the use of AI based tools like ChatGPT to support management research. *European Journal of Innovation Management*, 26(7), 233–241. <https://doi.org/10.1108/EJIM-02-2023-0156>
- Cao, Z., & Peng, L. (2024). An empirical study of factors influencing usage intention for generative artificial intelligence products: A case study of China. *Journal of Information Science*. <https://doi.org/10.1177/01655515241297329>. Advance online publication.
- Chen, J. V., Lin, C., Yen, D. C., & Linn, K.-P. (2011). The interaction effects of familiarity, breadth and media usage on web browsing experience. *Computers in Human Behavior*, 27(6), 2141–2152. <https://doi.org/10.1016/j.chb.2011.06.008>
- Chiu, C.-M., Hsu, M.-H., Lai, H., & Chang, C.-M. (2012). Re examining the influence of trust on online repeat purchase intention: The moderating role of habit and its antecedents. *Decision Support Systems*, 53(4), 835–845. <https://doi.org/10.1016/j.dss.2012.05.021>
- Choudhury, A., & Shamszade, H. (2023). Investigating the impact of user trust on the adoption and use of ChatGPT: Survey analysis. *Journal of Medical Internet Research*, 25, Article e47184. <https://doi.org/10.2196/47184>
- Choudhury, A., & Shamszade, H. (2024). The impact of performance expectancy, workload, risk, and satisfaction on trust in ChatGPT: Cross sectional survey analysis. *JMIR Human Factors*, 11(1), Article e55399. <https://doi.org/10.2196/55399>
- Choung, H., David, P., & Ross, A. (2023). Trust in AI and its role in the acceptance of AI technologies. *International Journal of Human-Computer Interaction*, 39(9), 1727–1739. <https://doi.org/10.1080/10447318.2022.2050543>
- Cui, D., & Wu, F. (2021). The influence of media use on public perceptions of artificial intelligence in China: Evidence from an online survey. *Information Development*, 37(1), 45–57. <https://doi.org/10.1177/0266666919893411>
- Cui, Y. L., Zeng, L. M., Ke, D. X., & Miao, H. W. (2025). What shapes learners' trust in AI? A meta analytic review of its antecedents and consequences. *IEEE Access*, 13, 164008–164025. <https://doi.org/10.1109/ACCESS.2025.3611367>
- Dadkhah, M., Oermann, M. H., Hegedüs, M., Raman, R., & Dávid, L. D. (2024). Diagnosis unreliability of ChatGPT for journal evaluation. *Advanced Pharmaceutical Bulletin*, 14(1), 1–4. <https://doi.org/10.34172/apb.2024.020>
- Dalvi-Esfahani, M., Barati Ahmadabadi, H., Ramayah, T., Turner, J. J., Iahad, N. A., & Azar, N. (2025). Stimulus–organism–response framework of decision makers' intention to adopt generative AI to replace entry level jobs: The moderating impact of personality traits. *Technological Forecasting and Social Change*, 219, Article 124291. <https://doi.org/10.1016/j.techfore.2025.124291>
- Dalvi-Esfahani, M., Mosharaf Dehkordi, M., Leong, L. W., Ramayah, T., & Jamal Kanaan Jebna, A. M. (2023). Exploring the drivers of XAI enhanced clinical decision support systems adoption: Insights from a stimulus–organism–response perspective. *Technological Forecasting and Social Change*, 195, Article 122768. <https://doi.org/10.1016/j.techfore.2023.122768>
- Das, T. K., & Teng, B.-S. (2001). Trust, control, and risk in strategic alliances: An integrated framework. *Organization Studies*, 22(2), 251–283. <https://doi.org/10.1177/0170840601222004>
- Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, 13(3), 319–340. <https://doi.org/10.2307/249008>
- Davis, T. R. V., & Luthans, F. (1980). A social learning approach to organizational behavior. *Academy of Management Review*, 5(2), 281–290. <https://doi.org/10.5465/amr.1980.4288758>
- Dingel, J., Kleine, A.-K., Cecil, J., Sigl, A. L., Lerner, E., & Gaube, S. (2024). Predictors of health care practitioners' intention to use AI enabled clinical decision support systems: Meta analysis based on the unified theory of acceptance and use of technology. *Journal of Medical Internet Research*, 26(1), Article e57224. <https://doi.org/10.2196/57224>
- Do, H. J., Kong, H.-K., Tetali, P., Karahalios, K., & Bailey, B. P. (2023). Inform, explain, or control: Techniques to adjust end user performance expectations for a conversational agent facilitating group chat discussions. *Proceedings of the ACM on Human-Computer Interaction*, 7(CSCW2), 1–26. <https://doi.org/10.1145/3610192>
- Dormann, C., & Griffin, M. A. (2015). Optimal time lags in panel studies. *Psychological Methods*, 20(4), 489–505. <https://doi.org/10.1037/met0000041>
- Dowling, M., & Lucey, B. (2023). ChatGPT for (Finance) research: The Bananarama conjecture. *Finance Research Letters*, 53, Article 103662. <https://doi.org/10.1016/j.frl.2023.103662>
- Earle, T. C., & Siegrist, M. (2006). Morality information, performance information, and the distinction between trust and confidence. *Journal of Applied Social Psychology*, 36(2), 383–416. <https://doi.org/10.1111/j.0021-9029.2006.00012.x>
- Enders, C. K., & Bandalos, D. L. (2001). The relative performance of full information maximum likelihood estimation for missing data in structural equation models. *Structural Equation Modeling: A Multidisciplinary Journal*, 8(3), 430–457. https://doi.org/10.1207/S15328007SEM0803_5
- Fabrigar, L. R., & Wegener, D. T. (2016). Conceptualizing and evaluating the replication of research results. *Journal of Experimental Social Psychology*, 66, 68–80. <https://doi.org/10.1016/j.jesp.2015.07.009>
- Gillath, O., Ai, T., Branicky, M. S., Keshmiri, S., Davison, R. B., & Spaulding, R. (2021). Attachment and trust in artificial intelligence. *Computers in Human Behavior*, 115, Article 106607. <https://doi.org/10.1016/j.chb.2020.106607>
- Glikson, E., & Woolley, A. W. (2020). Human trust in artificial intelligence: Review of empirical research. *The Academy of Management Annals*, 14(2), 627–660. <https://doi.org/10.5465/annals.2018.0057>
- Harries, R., Lawson, C., & Shapira, P. (2025). Generative AI in science: Applications, challenges, and emerging questions. *arXiv [Preprint]* <https://arxiv.org/abs/2507.08310>
- Hayes, A. F. (2015). An index and test of linear moderated mediation. *Multivariate Behavioral Research*, 50(1), 1–22. <https://doi.org/10.1080/00273171.2014.962683>
- 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>
- Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, 6(1), 1–55. <https://doi.org/10.1080/10705519909540118>
- Imran, M., & Almusharraf, N. (2023). Analyzing the role of ChatGPT as a writing assistant at higher education level: A systematic review of the literature. *Contemporary Educational Technology*, 15(4), ep464. <https://doi.org/10.30935/cedtech/13605>
- Jacovi, A., Marasović, A., Miller, T., & Goldberg, Y. (2021). Formalizing trust in artificial intelligence: Prerequisites, causes and goals of human trust in AI. *arXiv*. <https://doi.org/10.48550/arXiv.2010.07487> [Preprint].
- Kansal, A. (2024). Introduction to generative AI. In A. Kansal (Ed.), *Building generative AI powered apps: A hands on guide for developers* (pp. 1–15). Apress. https://doi.org/10.1007/978-8-8688-0205-8_1
- Kaplan, A. D., Kessler, T. T., Brill, J. C., & Hancock, P. A. (2023). Trust in artificial intelligence: Meta analytic findings. *Human Factors*, 65(2), 337–359. <https://doi.org/10.1177/00187208211013988>

- Karahanna, E., Straub, D. W., & Chervany, N. L. (1999). Information technology adoption across time: A cross-sectional comparison of pre-adoption and post-adoption beliefs. *Management Information Systems Quarterly*, 23(2), 183–213. <https://doi.org/10.2307/249751>
- Kenny, D. A., & Judd, C. M. (1984). Estimating the nonlinear and interactive effects of latent variables. *Psychological Bulletin*, 96(1), 201–210. <https://doi.org/10.1037/0033-2909.96.1.201>
- Khlaf, Z. N., Mousa, A., Hattab, M. K., Itmazi, J., Hassan, A. A., Sanmugam, M., & Ayyoub, A. (2023). The potential and concerns of using AI in scientific research: ChatGPT performance evaluation. *JMIR Medical Education*, 9, Article e47049. <https://doi.org/10.2196/47049>
- Kleine, A.-K., Schaffernak, I., & Lerner, E. (2025). Exploring predictors of AI chatbot usage intensity among students: Within and between person relationships based on the technology acceptance model. *Computers in Human Behavior: Artificial Humans*, 3, Article 100113. <https://doi.org/10.1016/j.chbah.2024.100113>
- Komiak, S. Y. X., & Benbasat, I. (2006). The effects of personalization and familiarity on trust and adoption of recommendation agents. *MIS Quarterly*, 30(4), 941–960. <https://doi.org/10.2307/25148760>
- Lankton, N. K., McKnight, D. H., & Thatcher, J. B. (2012). The moderating effects of privacy restrictiveness and experience on trusting beliefs and habit: An empirical test of intention to continue using a social networking website. *IEEE Transactions on Engineering Management*, 59(4), 654–665. <https://doi.org/10.1109/TEM.2011.2179048>
- 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*. <https://doi.org/10.48550/arXiv.2512.12413> [Preprint].
- Lee, D., Arnold, M., Srivastava, A., Plastow, K., Strelan, P., Ploekel, F., Lekkas, D., & Palmer, E. (2024). The impact of generative AI on higher education learning and teaching: A study of educators' perspectives. *Computers and Education: Artificial Intelligence*, 6, Article 100221. <https://doi.org/10.1016/j.caeai.2024.100221>
- Lee, J. D., & See, K. A. (2004). Trust in automation: Designing for appropriate reliance. *Human Factors*, 46(1), 50–80. <https://doi.org/10.1518/hfes.46.1.50.30392>
- Lee, G., & Xia, W. (2011). A longitudinal experimental study on the interaction effects of persuasion quality, user training, and first-hand use on user perceptions of new information technology. *Information & Management*, 48(7), 288–295. <https://doi.org/10.1016/j.im.2011.09.003>
- Liebherr, M., Enkel, E., Law, E. L.-C., Mousavi, M. R., Sammartino, M., & Sieberg, P. (2026). Dynamic calibration of trust and trustworthiness in AI-enabled systems. *International Journal on Software Tools for Technology Transfer*. <https://doi.org/10.1007/s10009-026-00840-6>
- Lin, G.-C., Wen, Z., Marsh, H. W., & Lin, H.-S. (2010). Structural equation models of latent interactions: Clarification of orthogonalizing and double mean centering strategies. *Structural Equation Modeling: A Multidisciplinary Journal*, 17(3), 374–391. <https://doi.org/10.1080/10705511.2010.488999>
- Lorenz, F., Lorenzen, S., Franco, M., Velz, J., & Clauß, T. (2024). Generative artificial intelligence in management research: A practical guide on mistakes to avoid. *Management Review Quarterly*. <https://doi.org/10.1007/s11301-024-00469-2>
- Lukyanenko, R., Maass, W., & Storey, V. C. (2022). Trust in artificial intelligence: From a foundational trust framework to emerging research opportunities. *Electronic Markets*, 32(4), 1993–2020. <https://doi.org/10.1007/s12525-022-00605-4>
- Mahmood, M. A., Hall, L., & Swanberg, D. L. (2001). Factors affecting information technology usage: A meta analysis of the empirical literature. *Journal of Organizational Computing and Electronic Commerce*, 11(2), 107–130. https://doi.org/10.1207/S15327744JOCE1102_02
- Mayer, R. C., Davis, J. H., & Schoorman, F. D. (1995). An integrative model of organizational trust. *Academy of Management Review*, 20(3), 709–734. <https://doi.org/10.5465/amr.1995.9508080335>
- McAllister, D. J. (1995). Affect and cognition based trust as foundations for interpersonal cooperation in organizations. *Academy of Management Journal*, 38(1), 24–59. <https://doi.org/10.5465/256727>
- McDermott, P. L., & ten Brink, R. N. (2019). Practical guidance for evaluating calibrated trust. *Proceedings of the Human Factors and Ergonomics Society - Annual Meeting*, 63(1), 362–366. <https://doi.org/10.1177/1071181319631379>
- McIntosh, T. R., Susnjak, T., Liu, T., Watters, P., Xu, D., Liu, D., & Halgamuge, M. N. (2025). From google gemini to OpenAI Q* (Q star): A survey on reshaping the generative artificial intelligence (AI) research landscape. *Technologies*, 13(2), 51. <https://doi.org/10.3390/technologies13020051>
- McKnight, H. D., Choudhury, V., & Kacmar, C. (2002). The impact of initial consumer trust on intentions to transact with a web site: A trust building model. *The Journal of Strategic Information Systems*, 11(3), 297–323. [https://doi.org/10.1016/S0963-8687\(02\)00020-3](https://doi.org/10.1016/S0963-8687(02)00020-3)
- Mehrabian, A., & Russell, J. A. (1974). *An approach to environmental psychology*. MIT Press.
- Misztal, B. (2013). *Trust in modern societies: The search for the bases of social order*. John Wiley & Sons.
- Mun, I. B., & Hwang, K.-H. (2025). Understanding ChatGPT continuous usage intention: The role of information quality, information usefulness, and source trust. *Information Development*, 41(3), 675–691. <https://doi.org/10.1177/02666669241307595>
- Mussweiler, T., & Strack, F. (2001). The semantics of anchoring. *Organizational Behavior and Human Decision Processes*, 86(2), 234–255. <https://doi.org/10.1006/obhd.2001.2954>
- Naik, D., Naik, I., & Naik, N. (2024). Imperfectly perfect AI chatbots: Limitations of generative AI, large language models and large multimodal models. In N. Naik, P. Jenkins, S. Prajapat, & P. Grace (Eds.), *Contributions presented at the international conference on computing, communication, cybersecurity and AI*, July 3–4, 2024 (pp. 43–66). London, UK: Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-74443-3_3
- Nguyen, A., Hong, Y., Dang, B., & Huang, X. (2024). Human AI collaboration patterns in AI assisted academic writing. *Studies in Higher Education*, 49(5), 847–864. <https://doi.org/10.1080/03075079.2024.2323593>
- O'Brien, C., Kelly, J., Lehan, E. A., Livingstone, V., Cotter, B., Butt, A., Kelly, L., & Corrigan, M. A. (2015). Validation and assessment of a technology familiarity score in patients attending a symptomatic breast clinic. *World Journal of Surgery*, 39(10), 2441–2449. <https://doi.org/10.1007/s00268-015-3134-1>
- Okamura, K., & Yamada, S. (2020). Adaptive trust calibration for human-AI collaboration. *PLoS One*, 15(2), Article e0229132. <https://doi.org/10.1371/journal.pone.0229132>
- Omrani, N., Rivieccio, G., Fiore, U., Schiavone, F., & Agreda, S. G. (2022). To trust or not to trust? An assessment of trust in AI based systems: Concerns, ethics and contexts. *Technological Forecasting and Social Change*, 181, Article 121763. <https://doi.org/10.1016/j.techfore.2022.121763>
- Panda, S., & Kaur, D. N. (2024). *Exploring the role of generative AI in academia: Opportunities and challenges* (SSRN scholarly Paper No. 4915738). *Social Science Research Network*. <https://papers.ssrn.com/abstract=4915738>
- Potosky, D., & Bobko, P. (1998). The computer understanding and experience scale: A self-report measure of computer experience. *Computers in Human Behavior*, 14(2), 337–348. [https://doi.org/10.1016/S0747-5632\(98\)00011-9](https://doi.org/10.1016/S0747-5632(98)00011-9)
- R Core Team. (2025). *R: A language and environment for statistical computing [Manual]*. R Foundation for Statistical Computing. <https://www.R-project.org/>
- Rana, A., Soodan, V., Kumar, A., Jayawardena, N. S., Thaichon, P., Gupta, K., & Rana, N. P. (2024). Identifying the influence of obsolescence risk and health beliefs in fitness wearable healthcare technology. *Journal of Global Scholars of Marketing Science*, 34(3), 366–392. <https://doi.org/10.1080/21639159.2024.2326496>
- Reinhardt, A., Matthes, J., Bojic, L., Maindal, H. T., Paraschiv, C., & Rym, K. (2025). Help me, doctor AI? A cross national experiment on the effects of disease threat and stigma on AI health information seeking intentions. *Computers in Human Behavior*, 172, Article 108718. <https://doi.org/10.1016/j.chb.2025.108718>
- Rempel, J. K., Holmes, J. G., & Zanna, M. P. (1985). Trust in close relationships. *Journal of Personality and Social Psychology*, 49(1), 95–112. <https://doi.org/10.1037/0022-3514.49.1.95>
- Riley, B. K., & Dixon, A. (2024). Emotional and cognitive trust in artificial intelligence: A framework for identifying research opportunities. *Current Opinion in Psychology*, 58, Article 101833. <https://doi.org/10.1016/j.copsyc.2024.101833>
- Rosseel, Y. (2012). Lavaan: An R package for structural equation modeling. *Journal of Statistical Software*, 48(2), 1–36. <https://doi.org/10.18637/jss.v048.i02>
- Sætra, H. S. (2023). Generative AI: Here to stay, but for good? *Technology in Society*, 75, Article 102372. <https://doi.org/10.1016/j.techsoc.2023.102372>
- Schlicker, N., & Langer, M. (2021). Towards warranted trust: A model on the relation between actual and perceived system trustworthiness. In *Proceedings of Mensch und Computer 2021* (pp. 325–329). <https://doi.org/10.1145/3473856.3474018>
- Schulenberg, S. E., & Melton, A. M. A. (2008). The computer aversion, attitudes, and familiarity index (CAAFI): A validity study. *Computers in Human Behavior*, 24(6), 2620–2638. <https://doi.org/10.1016/j.chb.2008.03.002>
- Sengar, S. S., Hasan, A. B., Kumar, S., & Carroll, F. (2024). Generative artificial intelligence: A systematic review and applications. *Multimedia Tools and Applications*, 84(21), 23661–23700. <https://doi.org/10.1007/s11042-024-20016-1>
- Shin, J., & Chan-Olmsted, S. (2023). User perceptions and trust of explainable machine learning fake news detectors. *International Journal of Communication*, 17, 23–23. <https://ijoc.org/index.php/ijoc/article/view/19534>
- Shrout, P. E., & Bolger, N. (2002). Mediation in experimental and nonexperimental studies: New procedures and recommendations. *Psychological Methods*, 7(4), 422–445. <https://doi.org/10.1037/1082-989X.7.4.422>
- Siegrist, M., & Cvetkovich, G. (2000). Perception of hazards: The role of social trust and knowledge. *Risk Analysis*, 20(5), 713–719. <https://doi.org/10.1111/0272-4332.205064>
- Sun, Y., Sheng, D., Zhou, Z., & Wu, Y. (2024). AI hallucination: Towards a comprehensive classification of distorted information in artificial intelligence-generated content. *Humanities and Social Sciences Communications*, 11(1), 1278. <https://doi.org/10.1057/s41599-024-03811-x>
- Teepapal, T. (2025). AI-driven personalization: Unraveling consumer perceptions in social media engagement. *Computers in Human Behavior*, 165, Article 108549. <https://doi.org/10.1016/j.chb.2024.108549>
- Thorne, S. (2024). Understanding the interplay between trust, reliability, and human factors in the age of generative AI. *International Journal of Simulation, Systems, Science and Technology*. <https://doi.org/10.5013/IJSSST.a.25.01.10>
- van der Werff, L., Blomqvist, K., & Koskinen, S. (2021). Trust cues in artificial intelligence. In *Understanding trust in organizations: A multilevel perspective* (pp. 307–333). Routledge.
- Van Noorden, R., & Perkel, J. M. (2023). AI and science: What 1,600 researchers think. *Nature*, 621(7980), 672–675. <https://doi.org/10.1038/d41586-023-02980-0>
- Van Quaquebeke, N., Tonidandel, S., & Banks, G. C. (2025). Beyond efficiency: How artificial intelligence (AI) will reshape scientific inquiry and the publication process. *The Leadership Quarterly*, 36(4), Article 101895. <https://doi.org/10.1016/j.leaqua.2025.101895>
- Venkatesh, V. (2000). Determinants of perceived ease of use: Integrating control, intrinsic motivation, and emotion into the technology acceptance model. *Information Systems Research*, 11(4), 342–365. <https://doi.org/10.1287/isre.11.4.342.11872>
- Venkatesh, V., & Bala, H. (2008). Technology acceptance model 3 and a research agenda on interventions. *Decision Sciences*, 39(2), 273–315. <https://doi.org/10.1111/j.1540-5915.2008.00192.x>

- Venkatesh, V., & Davis, F. D. (2000). A theoretical extension of the technology acceptance model: Four longitudinal field studies. *Management Science*, 46(2), 186–204. <https://doi.org/10.1287/mnsc.46.2.186.11926>
- Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. *MIS Quarterly*, 27(3), 425–478. <https://doi.org/10.2307/30036540>
- Waduge, A. O., Kulasooriya, W. K. V. J. B., Ranasinghe, R. S. S., Ekanayake, I., Rathnayake, U., & Meddage, D. P. P. (2024). Navigating the ethical landscape of ChatGPT integration in scientific research: Review of challenges and recommendations. *Journal of Computational and Cognitive Engineering*, 3(4), 360–372. <https://doi.org/10.47852/bonviewjce42023238>
- Wansink, B., Kent, R. J., & Hoch, S. J. (1998). An anchoring and adjustment model of purchase quantity decisions. *Journal of Marketing Research*, 35(1), 71–81. <https://doi.org/10.1177/002224379803500108>
- Watermeyer, R., Phipps, L., Lanclos, D., & Knight, C. (2024). Generative AI and the automating of academia. *Postdigital Science and Education*, 6(2), 446–466. <https://doi.org/10.1007/s42438-023-00440-6>
- Wissawaswaengsuk, P., Kumar, P., Frank, B., & Badir, Y. F. (2025). The role of trust as the facilitator and contingency factor in the adoption of digital healthcare services: A telemedicine context. *Computers in Human Behavior*, 172, Article 108722. <https://doi.org/10.1016/j.chb.2025.108722>
- Wixom, B. H., & Todd, P. A. (2005). A theoretical integration of user satisfaction and technology acceptance. *Information Systems Research*, 16(1), 85–102. <https://doi.org/10.1287/isre.1050.0042>
- Xu, S., Gong, Z., Li, Y., Zhao, Y., & Huang, S. (n.d.). The trust cost of AI errors: Examining cognitive and affective trust dynamics in human–AI collaboration. *International Journal of Human-Computer Interaction*. <https://doi.org/10.1080/10447318.2025.2560522>
- Yadav, M. S. (1994). How buyers evaluate product bundles: A model of anchoring and adjustment. *Journal of Consumer Research*, 21(2), 342–353. <https://www.jstor.org/stable/2489825>
- Yu, K., Berkovsky, S., Conway, D., Taib, R., Zhou, J., & Chen, F. (2018). Do I trust a machine? Differences in user trust based on system performance. In J. Zhou, & F. Chen (Eds.), *Human and machine learning: Visible, explainable, trustworthy and transparent* (pp. 245–264). Springer. https://doi.org/10.1007/978-3-319-90403-0_12
- Zhang, R., & Wang, J. (2025). Perceptions, adoption intentions, and impacts of generative AI among Chinese university students. *Current Psychology*, 44(11), 11276–11295. <https://doi.org/10.1007/s12144-025-07928-3>