

When humans stop thinking: tackling the silent threat of AI complacency in service operations

Khanh Bao Quang Le

*Department of Marketing and International Business,
Auckland University of Technology, Auckland, New Zealand, and*

Werner H. Kunz

*College of Management, University of Massachusetts Boston,
Boston, Massachusetts, USA*

Abstract

Purpose – This research investigates the phenomenon of AI complacency – The employee’s tendency to intentionally neglect validating AI-generated output even in the presence of systematic errors. It identifies the lack of monitoring accountability as the underlying driver of this phenomenon, assesses its consequences, and offers strategies for mitigation.

Design/methodology/approach – Grounded in contingency theory, this research employs six experimental studies ($N = 1,370$ participants), including one study with service employees across industries ($N = 160$ participants), to examine how insufficient monitoring accountability facilitates the emergence of AI complacency. The research explores both the causal mechanisms and the boundary conditions that modulate this effect.

Findings – The results show that the primary driver of AI complacency is the lack of accountability in monitoring AI-generated outputs. This complacency leads to detrimental work-related outcomes, such as increased commission errors and a diminished willingness to evaluate AI-generated outputs critically. The research also identifies situational factors that exacerbate and buffer these effects.

Practical implications – The findings highlight the critical need for organizations to implement systemic accountability frameworks that ensure employees actively engage with and oversee AI-generated output.

Originality/value – This research is among the first to examine AI complacency in the context of service provision empirically. It provides a theoretical framework, robust empirical evidence, and practical recommendations for improving Employee-AI collaboration in service provision, contributing to both academic discourse and managerial practice.

Keywords Artificial intelligence, AI monitoring accountability, AI complacency, Contingency theory, Confabulated information

Paper type Research article

Introduction

Organizations across service sectors increasingly rely on artificial intelligence (AI) to support complex knowledge-intensive tasks, including customer support, content creation, and decision-making (Lin, 2023). As service processes become more technology-mediated, AI agents are no longer peripheral tools but active participants in value creation (Noble *et al.*, 2022), and AI has attracted particular attention for enhancing efficiency and productivity (Cillo and Rubera, 2024). For example, the Boston Consulting Group found that consultants using AI technology achieved 12.2% higher completion, 25.1% faster execution, and 40% higher quality (Acar, 2023). Globally, AI could add \$4.4 trillion to annual productivity, especially in marketing and sales (Chui *et al.*, 2023).



Recent advances in Generative Artificial Intelligence (GenAI), such as ChatGPT, Claude, Copilot, and Gemini, represent a qualitative shift from earlier rule-based or predictive systems. Unlike conventional AI used for structured analytical tasks (Davenport and High, 2024), GenAI produces probabilistic, open-ended outputs that often appear fluent and human-like (Grewal *et al.*, 2024a, b). Such outputs can sometimes include confabulated information – outputs that sound plausible but are factually incorrect (Karakas *et al.*, 2023). Thus, despite the benefits, AI integration introduces a distinct and under-examined risk for service organizations (Kunz and Wirtz, 2024; Sigala *et al.*, 2024; Wirtz *et al.*, 2023). Although such AI-generated output may contain inaccurate information, a primary challenge in service contexts arises when employees fail to actively check the AI-generated outputs. In such circumstances, GenAI's seemingly fluent and logical outputs can create an illusion of reliability, reducing employees' propensity to scrutinize such outputs. As a result, errors from probabilistic generation can (unintentionally) spread into service decisions. Ultimately, such decisions can directly influence customer outcomes.

Recent incidents vividly illustrate this pressing challenge in service provision. For example, two lawyers were fined for submitting fabricated court citations generated by AI (Bohannon, 2023). Similarly, a university teaching staff mistakenly issued failing grades to an entire class because the AI agent falsely flags submissions as AI-generated (Ede-Osifo, 2023). Further, The Men's Journal published an AI-generated article with factual errors, which required damage control after criticism from a medical expert (Christian, 2023). These examples reflect a systemic, non-vigilant tendency to monitor AI-generated outputs and, subsequently, inadvertently accept confabulated information provided by the system.

In the context of this research, we refer to this phenomenon as "AI complacency" and define it as the employee's tendency to intentionally neglect validating generated content produced by AI, even in the presence of factual errors. AI complacency reflects a shift in employee vigilance rather than a lack of technical competence (Banker and Khetani, 2019; Parasuraman and Manzey, 2010). As AI takes on more complex tasks in services, workers might stop paying close attention to the generated outputs, resulting in "human-out-of-the-loop" service systems (e.g. Endsley, 2023). From a service provision perspective, understanding this phenomenon is essential for organizations aiming to integrate AI technologies effectively.

Accordingly, the current research aims to achieve two objectives. First, it investigates the drivers of AI complacency and examines its downstream consequences. The second objective is to identify the boundary conditions of AI complacency and provide insights to help service organizations mitigate its negative impacts. Hence, this research is guided by three research questions:

- RQ1. What is the driver of the AI complacency among service employees when working with AI agents?
- RQ2. How does AI complacency affect service employees' decision quality and effectiveness in AI-enabled service processes?
- RQ3. How do task-, team-, and organizational-level contingencies mitigate or exacerbate AI complacency in service processes?

To address these research questions, the current research integrates two theoretical streams. First, drawing on the attentional bias in information processing (Lerner and Tetlock, 1999), we argue that insufficient accountability in monitoring AI-generated outputs undermines employees' cognitive vigilance and serves as a primary driver of AI complacency. Second, building on organizational design literature, specifically contingency theory (Hult, 2011), we posit that the effects of AI monitoring on complacency depend on contextual factors. Since AI can be used at different levels and for various tasks within an organization, accountability is handled differently depending on the specific situation. These circumstances decide when poor AI oversight leads to AI-related issues. Taken together, these perspectives offer a new

explanation for AI complacency by connecting psychological attention processes with organizational design conditions. Thus, we emphasize that successful AI integration in customer service processes depends not just on capable AI systems but also on the implementation of accountability frameworks that keep employees alert when working with AI agents across different tasks and organizational settings.

This research makes three contributions to the literature on Employee-AI collaboration in service contexts (e.g. [Alkire et al., 2024](#); [Le et al., 2025](#); [van Riel et al., 2025](#)). First, although recent studies document growing difficulties faced by service employees working with AI agents ([Cillo and Rubera, 2024](#); [Grewal et al., 2024a, b](#); [Hermann and Puntoni, 2025](#)), prior work has primarily investigated these challenges through the lens of acceptance and adoption readiness (e.g. [Blaurock et al., 2025](#); [Kunz et al., 2025](#); [Wirtz et al., 2023](#)). As a result, human biases in AI-enabled service work (e.g. [Banker and Khetani, 2019](#)), particularly employees' susceptibility to confabulated or misleading AI-generated outputs, remain underexamined ([Sigala et al., 2024](#)). This research addresses this gap by bridging attentional bias research (e.g. [Lerner and Tetlock, 1999](#)) into service technology engagement ([Alkire et al., 2024](#)) and by investigating AI complacency as a consequence of a lack of monitoring accountability when interacting with AI agents and their downstream consequences. In doing so, this research explains how human cognitive biases can create risks in value co-creation when working with AI agents in service delivery. Addressing this phenomenon is critical for customer service and experience. If AI outputs are not carefully checked, errors in service provision can reach customers, harming the quality of Human-AI service teams and disrupting the overall experience (e.g. [Le et al., 2025](#)). Therefore, this research directly addresses calls for a more nuanced understanding of how AI is integrated into service practices ([Sigala et al., 2024](#)) and for a deeper exploration of human biases in the adoption of emerging technologies in service frontlines ([Dowling et al., 2020](#)).

Second, while prior studies have highlighted psychological and contextual barriers to AI integration (e.g. [Blaurock et al., 2025](#); [Cillo and Rubera, 2024](#); [Yang et al., 2025](#)), only recently has research begun to consider how service climate design supports frontline AI adoption ([van Riel et al., 2025](#)). Nevertheless, the role of contextual factors in shaping AI complacency and its downstream effects remains underexplored ([Alkire et al., 2024](#); [Sigala et al., 2024](#)). This research advances this nascent investigation by conceptualizing service frontline environments as deliberate design systems that facilitate employee monitoring accountability in AI-mediated interactions. Drawing on contingency theory, it theorizes how task-, team-, and organizational-level factors interact with accountability gaps to influence the onset and consequences of AI complacency. In doing so, this research reframes frontline design not only as a facilitator of AI adoption but also as a safeguard for employee oversight, ensuring the reliability and quality of service delivered by hybrid Employee-AI teams.

Third, this research offers actionable guidance for managers and practitioners seeking to mitigate AI complacency in service settings. Grounded in empirical findings, these recommendations provide concrete strategies to reinforce employee accountability and maintain vigilance when working with AI agents. This is critical for service management, where frontline decisions often directly affect customer outcomes. Without adequate safeguards, reduced monitoring of AI-generated outputs can compromise service quality. The research highlights the need to continue managing how employees and AI work together in service delivery. It shows that service organizations should do more than deploy technology. They need to create settings, processes, structures, and rewards that keep employees involved and in control when using AI to provide customer services.

Theoretical background

To grasp the challenges of AI complacency, it is essential to consider the extant literature on human biases toward emergent technologies. [Table 1](#) summarizes studies examining human biases in interactions with AI agents. A key stream of prior research on biases toward emergent technologies centers on AI aversion – the tendency to avoid using automated systems even

Table 1. An overview of related research

Author (s)	Type of bias			Focus of work processing with GenAI		AI complacency as a mechanism of interest?		Main IVs	Main DVs	Interventions	Summary of findings
	AB	AV	CB	M	DM	Yes	No				
This research	×			×		×		AI monitoring accountability	<ul style="list-style-type: none"> - False confirmation decision - Commission error tendency - Willingness to oversee AI output 	<ul style="list-style-type: none"> - Task-specific contingency: Task complexity, consequentiality, task load - Team-specific contingency: Task interdependence - Organization-specific contingency: Locus of performance accountability 	<ul style="list-style-type: none"> ● Low monitoring accountability leads to false confirmation decisions mediated by AI complacency ● High task complexity and task load amplify the impact of low monitoring accountability on AI complacency ● High task interdependence exacerbates the impact of complacency on willingness to oversee AI-generated output ● Performance outcome accountability weakens complacency's effect on commission error propensity

(continued)

Table 1. Continued

Author (s)	Type of bias			Focus of work processing with GenAI		AI complacency as a mechanism of interest?		Main IVs	Main DVs	Interventions	Summary of findings
	AB	AV	CB	M	DM	Yes	No				
Banker and Khetani (2019)	×				×		×	Perceived algorithmic expertise	- Purchase decision	- Educating users	<ul style="list-style-type: none"> ● Users followed generated recommendations even when inferior ● Educating users can reduce their susceptibility to inferior suggestions
Bonezzi and Osinelli (2021)			×		×		×	- Decision maker (Human vs. Algorithm)	- Perceived decision bias	- N/A	<ul style="list-style-type: none"> ● Algorithmic decisions are perceived as less biased ● Blindness to unique characteristics creates a perception of equal treatment, mediating the bias potential
Castelo et al. (2019)		×			×		×	- Task objectivity	- Dependence on algorithm	- Affective human-likeness	<ul style="list-style-type: none"> ● Algorithms are relied upon for subjective tasks due to perceived effectiveness and reduced discomfort ● Increasing human-like effect can enhance algorithm effectiveness in subjective tasks

(continued)

Table 1. Continued

Author (s)	Type of bias			Focus of work processing with GenAI		AI complacency as a mechanism of interest?		Main IVs	Main DVs	Interventions	Summary of findings
	AB	AV	CB	M	DM	Yes	No				
Castelo (2024)		×	×		×		×	- Country group (more corrupt vs. less corrupt)	- Preference for algorithm	- Task type	<ul style="list-style-type: none"> ● Participants from countries with high levels of corruption find algorithmic decision-making more appealing ● Individuals from corrupt countries also prefer algorithms for resource allocation tasks
Longoni <i>et al.</i> (2019)		×			×		×	- Service provider (Human vs. Algorithm)	- Resistance to the use of the algorithm	- Personalization	<ul style="list-style-type: none"> ● Users are reluctant to use algorithmic service providers due to perceived neglect of uniqueness ● Framing the service as personalized can reduce resistance

(continued)

Table 1. Continued

Author (s)	Type of bias			Focus of work processing with GenAI		AI complacency as a mechanism of interest?		Main IVs	Main DVs	Interventions	Summary of findings
	AB	AV	CB	M	DM	Yes	No				
<i>AB = AI Abuse; AV = AI Avoidance/Aversion; CB = Contextual Bias; M = Monitoring Activity; DM = Decision-making Activity</i>											
Longoni and Cian (2022)			×		×		×	- Hedonic vs. Utilitarian consumption contexts	- Preference for algorithmic recommendations	- Algorithm role	<ul style="list-style-type: none"> ● Algorithmic recommenders are preferred in utilitarian contexts and resisted in hedonic contexts due to perceived competence ● A hybrid human-AI decision-making approach can eliminate contextual bias
Longoni et al. (2023)			×		×		×	- Failing agent (Human vs. Algorithm)	- Inferential judgment	- Human supervision	<ul style="list-style-type: none"> ● Perceived algorithmic failures can transfer to other algorithms due to group homogeneity, undermining consumers' use of public services ● Human oversight neutralizes this transference effect
Lopez and Garza (2023)			×		×		×	- Evaluator (Human vs. Algorithm)	- Perceive fairness	Type of evaluation (Positive vs. Negative)	<ul style="list-style-type: none"> ● Users perceive evaluations by algorithms as less fair due to heightened transparency concerns

(continued)

Table 1. Continued

Author (s)	Type of bias			Focus of work processing with GenAI		AI complacency as a mechanism of interest?		Main IVs	Main DVs	Interventions	Summary of findings
	AB	AV	CB	M	DM	Yes	No				
Gaczek et al. (2023)		×			×		×	Advising agent (Human vs. Algorithm)	- Willingness to follow the recommendation	- Social proof	<ul style="list-style-type: none"> ● Consumers are less likely to accept AI medical recommendations for favorable diagnoses, mediated by trust in the diagnosis and intensified by health anxiety ● Social proof reduces the negative impact of health anxiety on trustworthiness
Garvey et al. (2023)		×			×		×	- Offer type (worse than expected, better than expected)	- Offer acceptance likelihood	- Offer administrator (Human vs. Algorithm) - Anthropomorphism	<ul style="list-style-type: none"> ● Consumers are more likely to accept a worse-than-expected offer from AI due to inferred intentions ● Decreasing AI anthropomorphism increases the likelihood of accepting worse-than-expected offers

AB = AI Abuse; AV = AI Avoidance/Aversion; CB = Contextual Bias; M = Monitoring Activity; DM = Decision-making Activity

(continued)

Table 1. Continued

Author (s)	Type of bias			Focus of work processing with GenAI		AI complacency as a mechanism of interest?		Main IVs	Main DVs	Interventions	Summary of findings
	AB	AV	CB	M	DM	Yes	No				
Srinivasan and Sarial-Abi (2021)			×		×		×	- Error maker (Human vs. Algorithm)	- Various measurements on brands	- Anthropomorphism - Task subjectivity and interactivity - Human supervision	<ul style="list-style-type: none"> ● Consumers respond less negatively to errors made by algorithms, mediated by perceived agency ● Anthropomorphizing the AI agent increases negative reactions ● Errors during interactive or subjective tasks lead to stronger negative reactions toward the brand ● Human supervision intensifies negative responses to AI errors
Yalcin <i>et al.</i> (2022)			×		×		×	- Decision maker (Human vs. Algorithm) - Outcome (Favorable vs. Unfavorable)	- Attitude toward the company	- Anthropomorphism	<ul style="list-style-type: none"> ● Consumers respond less positively to favorable algorithmic decisions due to internal attribution ● Humanizing the AI agent improves attitudes toward the company when decisions come from AI.

Note(s): AB = AI Abuse; AV = AI Avoidance/Aversion; CB = Contextual Bias; M = AI Monitoring Activity; DM = Decision-making Activity with AI

Source(s): Authors' own work

when they outperform human decision-making (Castelo *et al.*, 2019). This work addresses the broader phenomenon of AI adoption resistance (Longoni *et al.*, 2019) and explores downstream consequences of avoidance. Prior studies have identified factors that fuel aversion and proposed strategies to reduce resistance and build trust in AI systems (see Table 1). A second, related line of inquiry examines contextual bias, investigating how situational factors, such as the consumption context (Longoni and Cian, 2022), shape user preferences for the AI agent (vs. human). Both streams share a common goal – increasing the adoption of automated systems in decision-making. While aversion research emphasizes strategies to reduce resistance, contextual bias research highlights the role of situational factors in influencing preferences for AI.

Three significant gaps can be observed from Table 1. First, most research focuses on biases in decision-making with automated systems, with little attention to monitoring activity and, more specifically, to monitoring AI-generated output when confabulated information is presented (cf. Table 1). This has led to an incomplete understanding of how practitioners can effectively address AI complacency when using generated content as inputs for decision-making or other activities. Additionally, the literature does not sufficiently address the conditions under which AI complacency is amplified or mitigated. Second, most studies in the literature focus on increasing users' willingness to adopt AI. There is limited research on how to adopt it responsibly, especially amid the emergence of AI agents, and there is a high risk of becoming complacent with such agents. Finally, there is limited research on strategies to mitigate complacency in the context of Employee-AI collaboration.

In contrast to previous studies, this research examines an emerging stream of inquiry – AI abuse, which we refer to as the excessive reliance on AI-generated content for taskwork processing (Bahner *et al.*, 2008). To date, AI abuse has received little empirical attention in the service marketing domain (cf. Table 1). However, it is increasingly recognized as a critical challenge in the adoption of AI agents in service work contexts (Sigala *et al.*, 2024). AI avoidance reflects under-utilization of AI agents (Castelo *et al.*, 2019; Longoni *et al.*, 2019) and contextual bias highlights the role of situational factors in shaping adoption decisions (Bonezzi and Osinelli, 2021; Lopez and Garza, 2023). In contrast, AI abuse emphasizes the over-utilization of AI agents for task processing.

Additionally, unlike avoidance or contextual bias, which are most salient during the initial adoption stage (Castelo *et al.*, 2019; Srinivasan and Sarial-Abi, 2021), AI abuse typically emerges after adoption. At this stage, the focus shifts from whether to adopt AI to the quality of interaction with AI agents (Grote *et al.*, 2024). The present research, therefore, focuses on this underexplored stream.

Within the emergent research stream of AI abuse, we distinguish AI complacency from AI overreliance – two distinct phenomena often conflated in the literature. Overreliance stems from a lay belief that the AI system possesses superior domain expertise, leading users to defer to its outputs during decision-making – even in contexts where they have sufficient knowledge to exercise independent judgment (Banker and Khetani, 2019). As such, overreliance is not a consequence of monitoring failure, but rather a manifestation of trust miscalibration.

In contrast, AI complacency emerges during monitoring activities and is driven not by trust *per se*, but by attentional bias and situational neglect. Complacency reflects users' conscious (or semi-conscious) decision to deprioritize supervision, assuming the system is functioning correctly, particularly when cognitive resources must be allocated across multiple concurrent tasks (Parasuraman and Manzey, 2010). This strategic reallocation of attention becomes more likely under high workload or multitasking demands, where system oversight is sacrificed in favor of more pressing tasks. Lowered metacognitive capacity suggests that even when users know there could be issues, they may be less able to monitor or reflect critically, particularly under load (Helal *et al.*, 2025). Thus, AI complacency is not driven by ignorance or lack of knowledge, but by situational constraints that induce users to make attentional tradeoffs.

Main research framework

Overview

This research draws on the attentional bias literature in human-machine interaction (Metzger *et al.*, 2000; Parasuraman and Manzey, 2010; Sarter and Woods, 1995). It also builds on contingency theory, which posits that individuals' actions are influenced by the expected reactions and feedback from the surrounding environment (Tetlock, 1992). Together, these perspectives are used to explore the drivers and boundary conditions of AI complacency. The core focus of this research is perceived accountability, defined as the belief that one may need to explain or justify their thoughts, feelings, and actions to others (Lerner and Tetlock, 1999).

Accountability heightens the need for self-justification and self-scrutiny (Simonson and Staw, 1992), reduces inattentiveness during judgmental processes, and enhances the use of external cues in information processing (Hagfors and Brehmer, 1983). Accountability imposes a rational obligation to justify one's actions, creating psychological pressure to become self-critical and to exert greater effort to identify appropriate responses and detect errors (Skitka *et al.*, 2000). When employees are held responsible for the outcomes of their interactions with AI agents, they are more likely to engage with the agents. They may allocate greater attention to scrutinizing its outputs. This occurs due to the psychological pressure to justify their actions in the future. Thus, in the context of Employee-AI interaction in service provision, we define AI monitoring accountability as "the employee perception of being held accountable to justify their decisions or actions when checking AI output before integrating it into their task processing". When preemptive criticism is diminished, employees may be less inclined to engage in cognitively complex and critical thinking when interacting with AI agents (Grote *et al.*, 2024), which could reduce their attentiveness to the potential outcomes of their actions.

On the other hand, contingency theory posits that employee performance is shaped by various contextual factors that collectively influence how employees behave in their work environment (Donaldson, 2001; Hult, 2011). Hence, the application of technological advancements to enhance employee performance depends on the organizational context in which these technologies are embedded (Keller, 1994). The integration of AI agents into organizational workflows is likely influenced by the strategic approach to their deployment (Cillo and Rubera, 2024). AI agents can be applied at different levels within a service organization (Huang and Rust, 2024; Le *et al.*, 2025; Grewal *et al.*, 2024a, b). Hence, we propose that factors contingent on taskwork processing at different organizational levels influence the extent of insufficient AI monitoring accountability. These factors, in turn, affect the extent to which AI complacency emerges.

Task-based contingencies describe how the nature of tasks shapes humans' interactions with AI agents (Castelo *et al.*, 2019). Previous research suggests that task-related processing characteristics – subjective task complexity and task load could shape how employees collaborate with robotic agents (Le *et al.*, 2023). We refer to task load as the subjective perception of a task's demands and the effort required to accomplish it (Meijman and Mulder, 2013). Additionally, we define task complexity as the perception of the cognitive demands of processing tasks with multiple steps and numerous informational cues (Funke, 2010).

Although both concepts are directly related to the taskwork processing aspect (Karakikes and Nathanael, 2023; Le *et al.*, 2023), they are distinctive. Task complexity is inherently derived from perceived task characteristics (Xu *et al.*, 2020), where perceived complexity stems from the amount of information cues and their interdependence, which are tied to task requirements (Liu and Li, 2012). Task complexity thus remains consistent across individuals, although individual ability and experience affect how they perceive and address it (Wood, 1986). On the other hand, task load is shaped by the individual's perception of a task's cognitive demands within specific constraints (Hart and Staveland, 1988). High task complexity does not always translate into a high task load and vice versa. Experienced individuals may efficiently manage complex tasks due to developed skills and knowledge, thus perceiving lower task loads (Sweller, 1988). Conversely, simple tasks performed under severe time pressure or in high-stress environments may generate significant workload despite low

inherent complexity (Endsley, 1995). Thus, both are stressors in distinct ways. Task complexity demands more intensive cognitive processing for each task, while task load requires allocating cognitive resources across multiple competing tasks (Ekin et al., 2025).

Team-related contingency focuses on the group-level influence that can shape an individual’s behavior in a team taskwork setting (Wageman, 1995). One pivotal aspect of this contingency is task interdependence, defined as “the degree to which members depend on each other’s knowledge and skills to perform their tasks effectively” (Menguc et al., 2013, p. 22). In organizational settings, tasks are often interdependent and require collective effort (Courtright et al., 2015) – a dynamic that shapes individual behavior within teams (Wageman, 1995).

Organization-related contingency emphasizes the impact of organization-wide initiatives on employee behavior (Hult, 2011). In this sense, how accountability is implemented within an organization affects how employees use AI agents in their work. Prior research has identified several aspects of accountability design (Lerner and Tetlock, 1999). The locus of performance accountability addresses the “accountable for what” notion (Simonson and Staw, 1992). It refers to the specific focus of accountability in an organizational workflow. Accountability may be directed at the taskwork process (how tasks are performed). Alternatively, it may be directed at the taskwork outcome (the results achieved, regardless of how they were produced). In the context of this research, it is essential to further distinguish between two forms of accountability – AI monitoring accountability and employee performance accountability. The former refers specifically to the responsibility for overseeing and verifying the output generated by AI (the primary focus of the present research). In contrast, employee performance accountability encompasses the broader obligation to fulfill job responsibilities, irrespective of whether AI is used.

Hypotheses

The central proposition is that a perceived lack of accountability in monitoring AI-generated output is a primary driver of AI complacency. The extent of these effects is influenced by task-, team-, and organization-level contingencies. We first develop the central hypothesis, explaining how low AI monitoring accountability leads to complacency, followed by hypotheses on the moderating roles of task-, team-, and organization-related contingencies. Figure 1 illustrates the overarching research framework.

Accountability often implies that providing an unsatisfactory justification for one’s actions or decisions may result in negative repercussions (Lerner and Tetlock, 1999). For example,

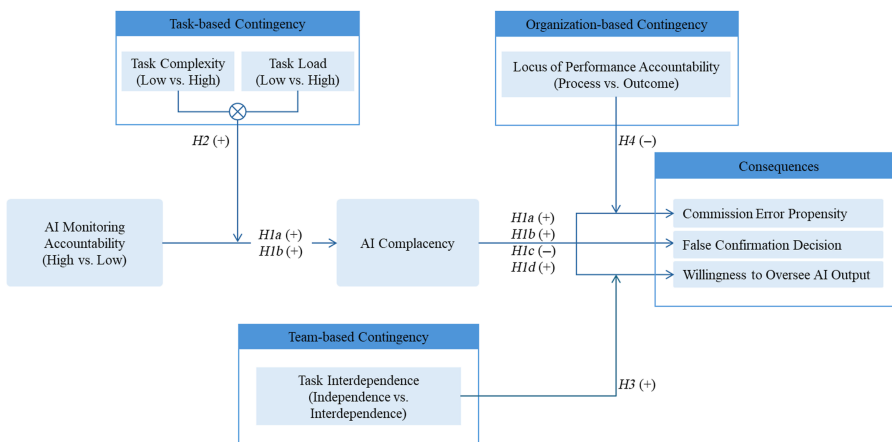


Figure 1. Main research framework. Source: Authors’ own work

many academic journals emphasize that authors are responsible for the accuracy of their articles and that Large Language Models (LLMs) are not qualified for authorship. This preemptive strategy can heighten awareness of one's actions and foster self-criticism, thereby helping avoid potential negative consequences (Tetlock, 1992).

In the service context, accountability mechanisms such as clear oversight roles, auditable processes, and governance frameworks are emphasized as essential for responsible AI adoption (Alkire *et al.*, 2024). Previous research shows that accountability reduces inattentiveness during judgment processes (Hagafors and Brehmer, 1983) and curbs overconfidence (Siegel-Jacobs and Yates, 1996). Conversely, low accountability could foster complacency by lowering the perceived need to evaluate AI outputs to gain performance efficiency critically (Spatola, 2024). Subsequently, it could lead to false confirmation decisions, in which we refer to the acceptance of AI-generated outputs containing errors as valid (Banker and Khetani, 2019), mediated by AI complacency.

Complacency carries important downstream consequences for how individuals interact with AI agents. It could influence willingness to oversee AI output, defined as the extent to which employees are willing to validate AI-generated content before using it actively, and commission error propensity, defined as the tendency to trust AI-generated output even when it is factually incorrect. When complacency is high, users are more likely to bypass monitoring to conserve attentional resources (Parasuraman and Manzey, 2010) and reduce critical scrutiny. For example, perceptions of AI as highly capable or reliable, particularly when systems are framed as "expert" decision aids, can lead users to minimize oversight, assuming little need to question outputs (Banker and Khetani, 2019; Romeo and Conti, 2025). Additionally, distal evidence suggests that users may discount AI-caused failures relative to human-caused failures, suggesting a diminished inclination to scrutinize AI output critically (Chen, 2024). Further, overoptimism about product capabilities mirrors a complacent mindset, whereby users fail to account for potential errors or evolving conditions and thus neglect proper monitoring (Goodman and Imrak, 2013). Based on this observation, we propose the following set of hypotheses.

- H1a.* A low (vs. high) perception of AI monitoring accountability increases the likelihood of AI complacency.
- H1b.* A low (vs. high) perception of AI monitoring accountability will lead to a greater likelihood of a false confirmation decision. This effect is mediated by AI complacency.
- H1c.* AI complacency will decrease the willingness to oversee AI-generated output.
- H1d.* AI complacency will lead to a greater propensity for commission errors.

Regarding the moderating effect of task-based contingency, prior research on monitoring automated systems in aircraft suggests that the complexity of monitoring tasks (e.g. visual displays) can influence how human agents allocate their attentional resources (Molloy and Parasuraman, 1996). Tasks involving multiple steps and detailed activities can increase complexity due to the greater cognitive demands required to process information (Liu and Li, 2012). Complex tasks can strain cognitive resources, and in such a condition. In these conditions, users are more likely to rely on AI systems to offload the cognitive burden (Moriuchi and Murdy, 2024; Xu *et al.*, 2020). This tendency is especially pronounced when the system appears intelligent or human-like (So *et al.*, 2024). The use of AI agents is an appealing solution for cognitively demanding tasks because they can rapidly process prompts and generate highly personalized, seemingly intelligent outputs (Le and Cayrat, 2024). Thus, task complexity can encourage greater reliance on AI-generated output, particularly when individuals feel less personal accountability for their actions, thereby increasing the likelihood of complacency.

Additionally, when the subjective task load is high, it can further stress mental resources, making reliance on an algorithm a rational strategy to distribute the attentional burden

(Grinschgl *et al.*, 2023; Singh *et al.*, 1993). When monitoring accountability is low, complacency may surge under high task load because individuals divert resources from validating the generated output, unable to manage the monitoring task and other pressing duties (Spatola, 2024). Together, these contextual pressures amplify the influence of low AI monitoring accountability.

H2. A high (vs. low) level of task complexity combined with a high (vs. low) level of task load exacerbates the influence of low monitoring accountability on AI complacency.

Task interdependence is a core feature of work design that emphasizes the coordination of actions across members (Sharma and Yetton, 2003). It plays a significant role in structuring teamwork in human-robot collaboration settings (Le *et al.*, 2025), where interdependent tasks require ongoing synchronization (Le *et al.*, 2023). In such contexts, members perform work jointly, which increases the risk of diffusion of responsibility (Cooper, 2024). Under these conditions, individuals may assume that others will monitor overall quality, prompting them to reduce their own investment in collective processes (Saluja *et al.*, 2024). When an AI agent is embedded in interdependent task processing, this reduced sense of personal ownership over the team-level outcome can further dampen the motivation to scrutinize AI-generated output, especially for individuals already prone to AI complacency. Employees may perceive their contribution as only one part of a larger whole, making reliance on AI-generated output less conspicuous and less likely to be challenged. By contrast, in more independent task settings, accountability is concentrated on identifiable individuals. Thus, the employee who utilizes the AI agent is fully responsible for the outcome, regardless of how they use it. Even when AI complacency is high, task independence should constrain its adverse effect on the willingness to oversee AI-generated output, as oversight responsibility cannot be easily shifted onto others.

H3. Task interdependence (vs. task independence) exacerbates the influence of AI complacency on willingness to oversee AI-generated outputs.

In terms of the moderating effect of organization-related contingency, we propose that outcome-focused (vs. process-focused) performance accountability will mitigate the impact of AI complacency on commission error propensity. Outcome-focused performance accountability heightens the need to defend past decisions (Lerner and Tetlock, 1999). Prior research suggests that salient performance stakes heighten scrutiny of AI recommendations and improve decision accuracy (Haduong and Smith, 2024). In contrast, process-focused performance accountability emphasizes thoroughness in the work process regardless of the outcome (Simonson and Staw, 1992; Siegel-Jacobs and Yates, 1996).

When performance accountability is outcome-focused, users are motivated to ensure that their decisions involving AI-generated output can withstand scrutiny based on final results. This focus prompts them to engage more critically with those outputs to defend their choices in the future (Langer *et al.*, 2024). This heightened vigilance discourages passive acceptance of the generated output, thereby reducing the propensity to commit commission errors. By contrast, under process-focused accountability, the emphasis shifts toward adherence to prescribed guidelines rather than direct evaluation of outcomes. Thus, employees may equate procedural conformity with sound decision-making, focusing more on whether established steps have been followed than on whether the outcome is substantively correct. This orientation can create an illusion of due diligence, where compliance with procedures is mistaken for critical evaluation. Distal evidence suggests that interfaces and explanation features often raise users' surface-level confidence in an AI agent's reliability without materially improving the accuracy of its outputs (Afroogh *et al.*, 2024). When accountability emphasizes process adherence in this way, employees are more likely to accept AI-generated output at face value. Hence, process-focused accountability may inadvertently amplify AI complacency's influence on the likelihood of commission errors.

H4. An outcome-focused (vs process-focused) performance accountability weakens the influence of AI complacency on commission error propensity.

Empirical program overview

The current research comprises six experimental studies examining the influence of low monitoring accountability on the emergence of AI complacency and its downstream consequences. Table 2 summarizes empirical studies and the results of objective or hypothesis testing. Figure 2 presents the correlation matrices across all studies (see Web Appendix G and H for overall measurement validity and concept operationalization). Additionally, experiment

Table 2. An overview of empirical studies

Study	Design	Service task context	Hypothesis/Objective ^a
Study 1A	Single factor (AI monitoring accountability: High vs Low)	Managing a social media campaign with AI-generated posts	H1a: A low (vs. high) perception of AI monitoring accountability increases the likelihood of complacency H1b: A low (vs. high) perception of AI monitoring accountability will lead to a greater likelihood of making a false confirmation decision. This effect is mediated by AI complacency
Study 1B	Single factor (AI monitoring accountability: High vs Low)	Compiling a customer feedback analysis using an AI-generated report	Replicating Study 1A results. Further control for sample characteristics, learning effort, and readiness propensities in AI adoption
Study 2	Single factor (AI monitoring accountability: High vs Low)	Crafting promotional materials using AI-generated emails	Replicating Study 1A's result and ruling out an alternative mechanism – AI overreliance
Study 3	2 × 2 × 2 (AI monitoring accountability: High vs Low x Task complexity: High vs Low x Task load: High vs Low)	Processing a customer refund request with an AI-generated report	H1c: AI complacency will decrease the willingness to oversee AI-generated output H2a: A high (vs low) level of task complexity, combined with a high (vs low) level of task load, worsens the effect of low monitoring accountability on AI complacency
Study 4	2x2 (AI monitoring accountability: High vs Low x Task interdependence: Independent vs Interdependent)	Scheduling appointments with a GenAI assistant	H3: Task interdependence (vs. task independence) exacerbates the negative influence of AI complacency on willingness to oversee AI-generated output
Study 5	2x2 (AI monitoring accountability: High vs Low x Locus of Performance Accountability: Outcome vs Process)	Resolving a warranty claim using GenAI	H1d: AI complacency will lead to a greater commission error propensity H4a: An outcome-focused (vs process-focused) performance accountability weakens the influence of AI complacency on commission error propensity

Note(s): ^aConfirmed hypotheses are highlighted in italic

Source(s): Authors' own work

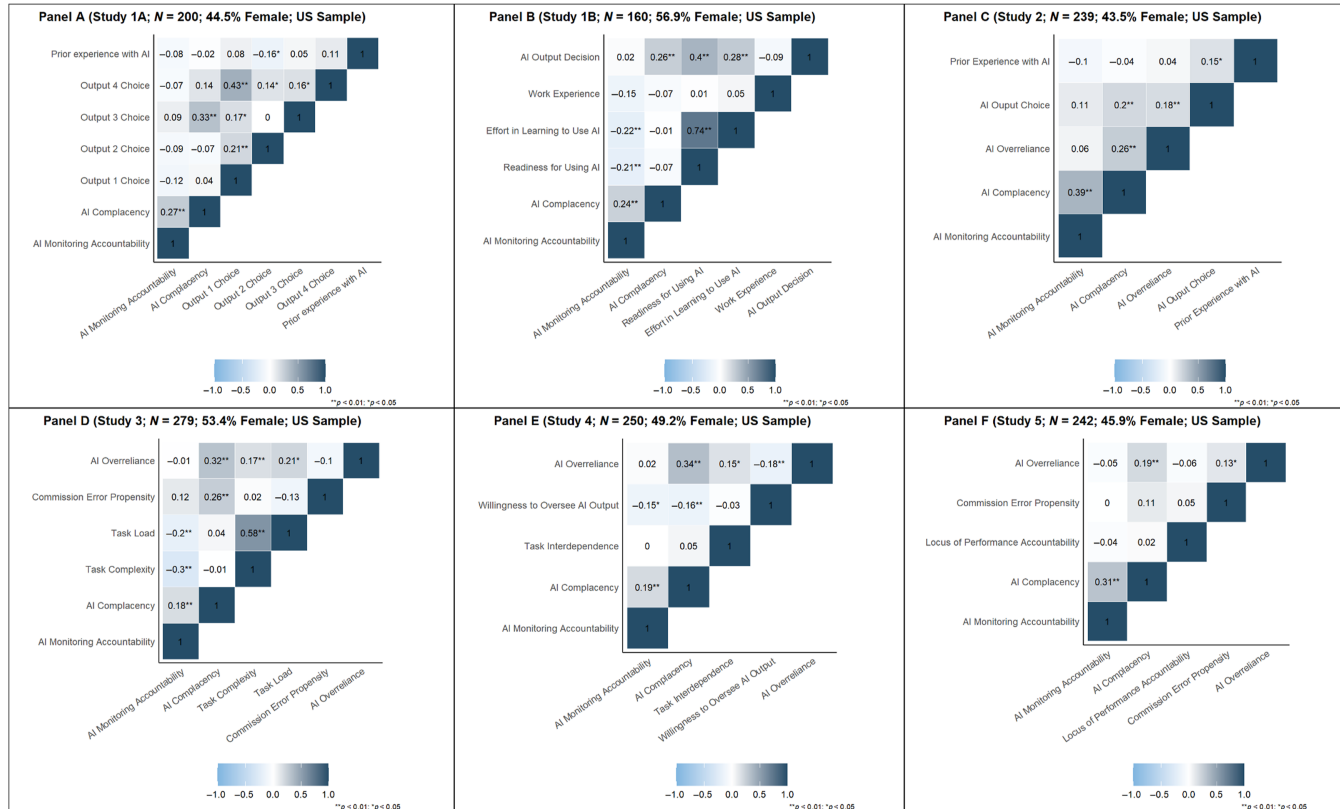


Figure 2. Correlation matrices across studies. AI monitoring accountability was reverse-coded prior to analyses. Source: Authors' own work

scenarios, measurement validity across all constructs across studies, and an overview of key concept definitions were documented in the [Web Appendices](#). Studies 1A, 1B and 2 focus on establishing evidence for the tendency toward AI complacency and distinguishing it from overreliance ([Banker and Khetani, 2019](#)). These studies also aim to demonstrate the role of low monitoring accountability in the development of AI complacency and its effects on subsequent outcomes.

Study 3 examines the moderation effect of task complexity and task load. Study 4 extends this inquiry to team-level contingencies, focusing on the moderating effect of task interdependence. Finally, Study 5 investigates the moderation effect of an organization-related contingency – locus of performance accountability. Across all experimental studies, various service-related contexts are utilized, including roles such as sales representatives or customer service agents. Despite these contextual differences, the findings consistently demonstrate the robust effect of low monitoring accountability on the development of AI complacency.

Study 1A

Design and procedure

The objective of Study 1A is to test [Hypotheses 1A](#) and [1B](#). A single-factor (AI monitoring accountability: high vs. low) between-subjects design was employed. The context of this experiment involved creating social media posts for an external client (see [Web Appendix A](#)). Participants were recruited through the Cloud Research platform.

Participants assumed the role of employees at a marketing agency tasked with creating social media posts for a fictitious client, “Funky Style,” a local clothing retailer. The client requested four posts for an upcoming monthly campaign, highlighting three key details: operating hours, a click-and-collect service, and free shipping for orders over \$50. This information was held constant in the base scenario to establish a reference point for the monitoring phase.

Next, participants were introduced to the framing of monitoring accountability. In the high monitoring accountability condition, participants were informed that they were responsible for thoroughly verifying all AI-generated content and submitting it to their manager for review. They also had to justify their decisions before the content could be sent to the client. In the low monitoring accountability condition, participants were told they were only responsible for basic verification and that the content could be sent to the client without managerial review.

Then, in the monitoring phase, participants reviewed four AI-generated social media posts and decided whether to accept or reject each post for the fictitious client. Two AI-generated outputs were factually accurate based on the provided information, while others contained deliberate inaccuracies that participants were unaware of beforehand. The study was designed to assess how effectively participants could verify AI-generated output against the baseline information. If participants accepted outputs containing inaccuracies without thorough verification, it signaled a deliberate omission of this activity – a marker of AI complacency ([Parasuraman and Manzey, 2010](#)).

The first two posts were factually aligned with the given information in the base scenario. However, the third post inaccurately stated that the store operates 24/7, whereas the provided details specified operating hours as Monday to Friday, 9:00 AM to 5:00 PM; Saturday, 10:00 AM to 5:00 PM; and Sunday, 10:00 AM to 3:00 PM. The fourth post falsely claimed free shipping for all orders, implying the inclusion of click-and-collect services, which contradicted the baseline details. In sum, outputs three and four were factually incorrect. To assess the believability and realism of the scenario, participants rated these aspects on a 7-point scale (1 = strongly disagree, 7 = strongly agree). The scenario was deemed acceptable, with participants giving high ratings for realism ($M = 5.86$, $SD = 1.06$) and believability ($M = 5.88$, $SD = 1.09$).

Manipulation check

An independent-samples *t*-test was conducted to verify the manipulation. In the high monitoring accountability condition, participants perceived a greater level of accountability ($M = 6.07$, $SD = 0.92$) than in the low accountability condition ($M = 5.63$, $SD = 1.32$), $t(178.22) = 2.75$, $p < 0.01$. Thus, the results confirmed the effectiveness of the manipulation. Perceived AI monitoring accountability was reverse-coded prior to analysis for ease of interpretation.

Results

Hypotheses were tested using the PROCESS macro in the *R* environment (Hayes, 2017). The regression results for Study 1A are presented in Table 3. The findings confirmed the presence of AI complacency resulting from a low level of monitoring accountability (β (SE) = 0.30 (0.08), $p < 0.01$), supporting Hypothesis 1A. Logistic regression results showed that for the factually correct generated outputs, participants were not susceptible to complacency (output 1: β (SE) = 0.18 (0.17), $p > 0.05$; output 2: β (SE) = -0.11 (0.14), $p > 0.05$). However, for the factually incorrect generated outputs (outputs 3 and 4), AI complacency significantly increased the likelihood of false confirmation decisions (output 3: β (SE) = 0.56 (0.13), $p < 0.01$; output 4: β (SE) = 0.47 (0.21), $p < 0.01$). Mediation analyses also revealed that AI complacency significantly mediated the effects of low monitoring accountability on generated outputs 3 and 4, but not on generated outputs 1 and 2, thus supporting Hypothesis 1B.

Table 3. Study 1A regression results

PROCESS model 4 (unstandardized scores)	AI complacency (2)	Separate analyses			
		AI output one decision [†] (3)	AI output two decision [†] (4)	AI output three decision [†] (5)	AI output four decision [†] (6)
Constant	2.05 (0.35)***	1.24 (0.82) ^{ns}	3.67 (0.87)***	-2.65 (0.71)***	0.32 (0.88)***
AI Monitoring Accountability [◆] (1)	0.30 (0.08)*** (H1a)	-0.29 (0.16) ^{ns}	-0.19 (0.16) ^{ns}	0.01 (0.14) ^{ns}	-0.28 (0.18) ^{ns}
AI Complacency (2)		0.18 (0.17) ^{ns}	-0.11 (0.14) ^{ns}	0.56 (0.13)***	0.47 (0.21)**
Control Variables					
Prior Experience with AI	-0.003 (0.09) ^{ns}	0.19 (0.19) ^{ns}	-0.44 (0.18)**	0.13 (0.16) ^{ns}	0.32 (0.21) ^{ns}
Model Summary					
R ² /Cox & Snell R ²	0.07	0.02	0.04	0.10	0.04
MSE/Nagelkerke R ²	1.53	0.04	0.06	0.14	0.08
F/ χ^2	7.64***	4.68 ^{ns}	7.94***	22.10***	9.05**
df	2,197	3,196	3,196	3,196	3,196
Mediation Analysis					
	Effect (Boot SE)	LLCI _{95%}	ULCI _{95%}		
(1)→(2)→(3)	0.05 (0.05)	-0.03	0.18		
(1)→(2)→(4)	-0.03 (0.05)	-0.15	0.05		
(1)→(2)→(5)	0.16 (0.07) (H1b)	0.06	0.35		
(1)→(2)→(6)	0.14 (0.08) (H1b)	0.03	0.34		

Note(s): *** $p < 0.01$; ** $p < 0.05$. Outputs one and two are factually correct outputs, whereas outputs three and four are factually incorrect compared to the given information (See Web Appendix A for full scenario setup). [◆] Reverse-coded prior to analysis. [†]Output decision: 0 = Reject, 1 = Accept. Results are in log-odds metrics. Prior Experience with AI was controlled since previous research suggests that this experience could have an impact on users' general attitude toward AI performance (Le et al., 2025)

Source(s): Authors' own work

Discussion

Study 1A provides preliminary evidence that low accountability in AI monitoring is associated with greater complacency, which can lead to false confirmation decisions in service processes. This extends prior research on resistance to and overreliance on AI agents (Castelo *et al.*, 2019; Dietvorst *et al.*, 2015; Longoni *et al.*, 2019) by identifying accountability as a structural factor that shapes frontline employees' engagement with AI agents. Recent frameworks have outlined directions for cultivating service climates that foster effective Employee-AI collaboration and for embedding responsible AI practices into service excellence (Alkire *et al.*, 2024; van Riel *et al.*, 2025). The present study offers initial empirical support for these propositions by showing that accountability (or its absence) functions as a critical structural lever. Although Employee-AI collaboration holds significant promises for enhancing value co-creation (Le *et al.*, 2025), these benefits are contingent on safeguards. Without appropriate accountability mechanisms, service organizations risk fostering complacency that ultimately undermines the quality of hybrid team service delivery.

Study 1B*Design and procedure*

The objective of Study 1B is twofold. First, it aims to replicate the findings of Study 1A. Second, it seeks to enhance ecological validity by controlling for sample characteristics, particularly given the specialized nature of the service task described in the previous experimental scenario. To this end, a single-factor between-subjects design was employed, manipulating monitoring accountability (high vs. low). In addition, as we posit that intentional negligence in monitoring AI outputs stems from low monitoring accountability, we also accounted for participants' AI readiness and their efforts to learn to use AI agents in service workplaces, as individual differences in readiness and learning motivation may shape how users engage with AI agents (Le *et al.*, 2023).

Study 1B was designed as a lab experiment with participants currently working in the retail industry and with professional experience in customer response in that sector. The experimental scenario involved compiling a customer feedback report for a fictional retail brand, Ecowear, which specializes in apparel. Participants were instructed to serve as a customer service analyst tasked with compiling a customer feedback report. Information about the customer review details was provided in the base scenario to establish a reference point. Next, in alignment with the previous study, participants read the framing of monitoring accountability. The manipulated information pertained to the AI-generated summary. The provided details did not include any comments on color options, whereas the AI-generated report incorrectly summarized that color options were appreciated in the reviews (see [Web Appendix B](#)).

The control process for sample characteristics consisted of two steps. First, we conducted a dedicated screening on the Cloud Research platform to recruit professional participants currently employed in the retail sector. Second, within the experimental protocol, we collected additional data regarding participants' accumulated work experience in the sector. This deliberate sampling strategy ensures that participants' day-to-day familiarity with retail work aligns with the experimental context, increasing the relevance of the AI-based task. [Table 4](#) below provides an overview of the sample characteristics. Overall, we recruited 160 participants who are currently working in the retail industry. The scenario was deemed acceptable, with participants giving high ratings for realism ($M = 5.76$, $SD = 0.96$) and believability ($M = 5.83$, $SD = 0.92$).

Manipulation check

An independent-samples *t*-test was conducted to verify the manipulation. In the high monitoring accountability condition, participants perceived a greater level of accountability

Table 4. An overview of sample descriptive statistics in study 1B

Retail sector	Sample distribution
Apparel and fashion	33 (20.6%)
Electronics and technology	12 (7.5%)
Grocery and food	48 (30%)
Home improvement/hardware	7 (4.4%)
Department store	15 (9.4%)
Luxury	2 (1.3%)
Health and beauty	5 (3.1%)
Furniture and home goods	2 (1.3%)
Convenience store	10 (6.3%)
Others	26 (16.3%)
Total	160
<hr/>	
Accumulated work experience in the retail industry	Sample distribution
Less than a year	4 (2.5%)
1–2 years	24 (15%)
2–3 years	19 (11.9%)
3–4 years	15 (9.4%)
4–5 years	9 (5.6%)
5–6 years	10 (6.3%)
6–7 years	9 (5.6%)
7–8 years	15 (9.4%)
8–9 years	7 (4.4%)
9–10 years	6 (3.8%)
More than 10 years	42 (26.3%)
Total	160
<hr/>	
Source(s): Authors' own work	

($M = 6.03$, $SD = 0.89$) than in the low accountability condition ($M = 4.87$, $SD = 1.37$), $t(133.61) = 6.32$, $p < 0.01$. Thus, the results confirmed the effectiveness of the manipulation.

Results

The regression results for Study 1B are presented in Table 5. The findings reconfirmed the presence of AI complacency resulting from a low level of monitoring accountability (β (SE) = 0.20 (0.07), $p < 0.01$). Logistic regression results also showed that AI complacency significantly increased the likelihood of the false confirmation decision (β (SE) = 0.53 (0.17), $p < 0.01$). Mediation analyses revealed that AI complacency significantly mediated the effects of low monitoring accountability (Effect (BootSE) = 0.11 (0.06), $CI_{95\%} = [0.03 \text{ to } 0.25]$). The presence of work experience (β (SE) = -0.01 (0.02), $p > 0.05$); AI readiness (β (SE) = -0.07 (0.07), $p > 0.05$); Effort in learning to use AI (β (SE) = 0.07 (0.07), $p > 0.05$) were all insignificant.

Discussion

By targeting currently employed retail professionals and aligning the task with industry-specific experience, the study enhances ecological validity. Individual differences, such as AI readiness, learning effort, and work experience, did not significantly affect the development of AI complacency. This suggests that structural design choices, particularly around accountability, play a more prominent role in employees' tendency to be complacent with AI-generated output.

Table 5. Study 1B regression results

PROCESS model 4 (unstandardized scores)	Model 1		Model 2	
	AI complacency (2)	AI output decision [†] (3)	AI complacency (2)	AI output decision [†] (3)
Constant	2.05 (0.18)***	-0.99 (0.49)**	2.13 (0.38)***	-4.42 (1.08)***
AI Monitoring Accountability ♦ (1)	0.20 (0.07)***	-0.08 (0.13) ^{ns}	0.20 (0.07)***	0.07 (0.15) ^{ns}
AI Complacency (2)		0.53 (0.17)***		0.69 (0.19)***
Control Variables				
Perceived Readiness for Using AI			-0.07 (0.07) ^{ns}	0.72 (0.18)***
Perceived Effort in Learning to Use AI			0.07 (0.07) ^{ns}	-0.08 (0.16) ^{ns}
Work Experience			-0.01 (0.02) ^{ns}	-0.06 (0.05) ^{ns}
Model Summary				
R ² /Cox & Snell R ²	0.06	0.07	0.06	0.24
MSE/Nagelkerke R ²	1.12	0.09	1.13	0.32
F/χ ²	9.60***	11.11***	2.68**	44.29***
df	1,158	2,157	4,155	5,154
Mediation Analysis	Effect (Boot SE)	LLCI _{95%}	ULCI _{95%}	
(1)→(2)→(3) (Model 1)	0.11 (0.06)	0.03	0.25	

Note(s): *** $p < 0.01$; ** $p < 0.05$. ♦ Reverse-coded prior to analysis. †Output decision: 0 = Reject, 1 = Accept. Results are in log-odds metrics

Source(s): Authors' own work

Study 2

Design and procedure

As AI complacency and overreliance are closely interconnected (Parasuraman and Manzey, 2010), this study aims to rule out AI overreliance as a competing mechanism. Thus, a single-factor (monitoring accountability: high vs. low), between-subjects experimental design was employed. The context involved crafting promotional materials for a local car dealership (see Web Appendix C). Participants were recruited through the Cloud Research platform.

This study used the same experimental setup as in the previous study. Participants played the role of a marketing coordinator tasked with creating promotional materials for the year-end sales event of a fictitious company, “Prime Auto Group,” a local car dealership. Information about the event’s details was provided in the base scenario to establish a reference point. Next, in alignment with the previous study, participants read the framing of AI monitoring accountability. Regardless of the condition, the AI-generated email was always factually incorrect, and this information was unknown to participants.

The manipulated information pertained to the SUV deal. The provided details stated that the offer was valid only for online purchases, whereas the AI-generated email incorrectly indicated that it was valid both in-store and online. All other content in the AI-generated email remained consistent with the original information. Then, in the monitoring phase, participants reviewed the AI-generated email and decided whether to accept or reject it. The scenario was deemed acceptable, with participants giving high ratings for realism ($M = 5.55$, $SD = 1.23$) and believability ($M = 5.59$, $SD = 1.31$). The measurement of AI overreliance is documented in the Web Appendix.

Manipulation check

An independent-samples t -test was conducted to check the manipulation. In the high monitoring accountability condition, participants reported greater accountability ($M = 6.32$, $SD = 0.96$) than in the low monitoring accountability condition ($M = 3.56$, $SD = 1.89$), $t(172.78) = 14.17$, $p < 0.01$. Thus, the results confirmed the manipulation.

Results

Study 2 regression results are documented in Table 6. The results replicated the effect of low monitoring accountability on AI complacency (β (SE) = 0.29 (0.04), $p < 0.01$). The chain mediating effect on AI output choice behavior was also significant (Effect (Boot SE) = 0.10 (0.05), $CI_{95\%} = [0.01 \text{ to } 0.23]$). This result reconfirmed Hypotheses 1A and 1B. To rule out the alternative mechanism, we modeled AI overreliance as a parallel mediator. The results indicate that low perception of monitoring accountability did not influence overreliance (β (SE) = 0.05 (0.05), $p > 0.05$). Further, the mediation from low monitoring accountability to AI output choice through overreliance was insignificant (Effect (Boot SE) = 0.01 (0.02), $CI_{95\%} = [-0.01 \text{ to } 0.05]$).

Discussion

This study contributes to research on AI abuse (Table 1) by empirically distinguishing between AI complacency and overreliance. While prior work has highlighted the risk of overreliance on emerging technologies (Banker and Khetani, 2019), it has not clearly distinguished this bias from complacency, nor has it examined the role of monitoring accountability in shaping these phenomena. Our findings suggest that low monitoring accountability increases the likelihood of false confirmation errors primarily through complacency, thereby shedding light on the mechanisms underlying observed decision failures. By differentiating these two phenomena, the study provides a more precise basis for exploring the boundary conditions for designing accountability structures that support more effective Employee-AI collaboration in service provision contexts.

Study 3

Design and procedure

The objective of Study 3 is to test Hypotheses 1C and 2. The hypotheses propose that AI complacency reduces the willingness to oversee AI-generated output and that task complexity and task load moderate the effect of monitoring accountability on AI complacency. A 2 (AI

Table 6. Study 2 regression results

PROCESS model 4 (unstandardized Scores)	AI complacency (2)	AI overreliance (3)	AI output decision [†] (4)
Constant	2.13 (0.35)***	2.79 (0.37)***	-1.90 (0.81)**
AI Monitoring Accountability \blacklozenge (1)	0.29 (0.04)*** (H1a)	0.05 (0.05) ^{ns}	0.08 (0.10) ^{ns}
AI Complacency (2)			0.34 (0.15)**
AI Overreliance (3)			0.26 (0.14) ^{ns}
Control Variables			
Prior Experience with AI	-0.01 (0.09) ^{ns}	0.07 (0.09) ^{ns}	0.46 (0.18)**
Model Summary			
R ² /Cox & Snell R ²	0.15	0.01	0.09
MSE/Nagelkerke R ²	1.91	2.10	0.15
F/ χ^2	21.58***	0.77 ^{ns}	22.27***
df	2,236	2,236	4,234
Mediation Analysis	Effect (Boot SE)	LLCI _{95%}	ULCI _{95%}
[A] (1)→(2)→(4)	0.10 (0.05) (H1b)	0.01	0.23
[B] (1)→(3)→(4)	0.01 (0.02)	-0.01	0.05
Pairwise comparison: [A] - [B]	0.09 (0.06)	-0.003	0.21

Note(s): *** $p < 0.01$; ** $p < 0.05$. \blacklozenge Reverse-coded prior to analysis. \dagger Output decision: 0 = Reject, 1 = Accept, Results are in log-odds metrics

Source(s): Authors' own work

monitoring accountability (high vs. low) x 2 (task complexity: high vs. low) x 2 (task load: high vs. low) full factorial between-subjects design was employed. The experiment context focused on processing a customer refund request (see [Web Appendix D](#)). Participants were recruited through Cloud Research.

Participants assumed the role of customer service representative at a local consumer electronics retailer, “TechMart Electronics.” Similar to previous studies, baseline information about the refund request was provided to establish a reference point. In this study, the key information was the missing charger – the customer indicated a missing charger in the baseline information, but the AI-generated refund processing request omitted this detail.

Next, participants read the framing for monitoring accountability, task complexity, and task load. The manipulation for monitoring accountability was similar to that in previous studies. Task complexity was manipulated by varying the amount of information regarding the refund process. In the high task complexity condition, participants read a 5-step procedure with extensive details about each step, whereas in the low task complexity condition, the same 5-step procedure was presented with simplified information. This operationalization of task complexity was based on the amount of available information in taskwork processing ([Liu and Li, 2012](#)).

Task load was manipulated based on the subjective perception of task demands ([Meijman and Mulder, 2013](#)). Specifically, in the high task load condition, participants read that they were handling multiple urgent tasks in addition to processing the refund request. In the low task-load condition, participants were told they had only a few requests to process and no other pressing deadlines. Then, in the monitoring phase, participants reviewed the AI-generated report and decided whether to accept or reject it before sending it to the finance department for processing the refund request. The scenario was deemed acceptable, with high ratings of realism ($M = 5.92$, $SD = 0.95$) and believability ($M = 5.93$, $SD = 0.99$).

Manipulation check

Independent-samples *t*-tests were conducted to verify the manipulations. In the high monitoring accountability condition, participants perceived a greater level of accountability ($M = 5.63$, $SD = 1.22$) than in the low monitoring accountability condition ($M = 4.75$, $SD = 1.42$), $t(270.22) = 5.56$, $p < 0.01$. Similarly, participants in the high task complexity condition reported higher task complexity ($M = 4.14$, $SD = 1.57$) than those in the low task complexity condition ($M = 3.61$, $SD = 1.53$), $t(277) = 2.84$, $p < 0.01$. Task load was also higher in the high task load condition ($M = 3.67$, $SD = 0.84$) compared to the low task load condition ($M = 2.56$, $SD = 0.80$), $t(277) = 11.34$, $p < 0.01$. Thus, the results confirmed the effectiveness of the manipulations.

Results

Study 3 regression results are documented in [Table 7](#). The baseline model reconfirmed the influence of low monitoring accountability on the development of AI complacency (β (SE) = 0.16 (0.05), $p < 0.01$). Further, we found that AI complacency negatively affects willingness to oversee AI output (β (SE) = -0.14 (0.05), $p < 0.01$), thus confirming [Hypothesis 1C](#). The full model revealed a three-way interaction effect of task complexity and task load on the relationship between low AI monitoring accountability and complacency (β (SE) = 0.12 (0.06), $p < 0.05$). Conditional effect analysis indicated that the impact was strongest under conditions of high task complexity and high task load (Effect (Boot SE) = 0.37 (0.11), $CI_{95\%} = [0.15 \text{ to } 0.59]$). This result confirms [Hypothesis 2](#). [Figure 3](#) provides a visualization of this interaction.

Discussion

Within the service literature, prior scholarship has explored how AI reshapes customer experiences and employee roles, often emphasizing both opportunities and risks ([Wirtz et al., 2023](#);

Table 7. Study 3 regression results

PROCESS models (unstandardized Scores)	Model 1		Model 2		Model 3	
	AI complacency	Willingness to oversee AI output	AI complacency	Willingness to oversee AI output	AI complacency	Willingness to oversee AI output
Constant	1.68 (0.32)***	4.26 (0.29)***	2.14 (0.28)***	3.90 (0.26)***	2.15 (0.29)***	3.90 (0.26)***
AI Monitoring Accountability ♦ (1)	0.16 (0.05)***	-0.13 (0.04)***	0.17 (0.05)***	-0.13 (0.04)***	0.13 (0.07) ^{ns}	-0.13 (0.04)***
Task Complexity (2)			-0.01 (0.05) ^{ns}		-0.02 (0.07) ^{ns}	
Task Load† (3)					0.003 (0.15) ^{ns}	
(1)*(2)			0.03 (0.03) ^{ns}		-0.01 (0.04) ^{ns}	
(1)*(3)					0.05 (0.11) ^{ns}	
(2)*(3)					0.06 (0.10) ^{ns}	
(1)*(2)*(3)					0.12 (0.06)** (H2)	
AI Complacency Control Variables		-0.14 (0.05)*** (H1c)		-0.14 (0.05)***		-0.14 (0.05)***
Prior Experience with AI	-0.10 (0.07) ^{ns}	0.14 (0.06)**	-0.10 (0.07) ^{ns}	0.14 (0.06)**	-0.10 (0.07) ^{ns}	0.14 (0.06)**
AI Overreliance	0.29 (0.05)***	-0.04 (0.05) ^{ns}	0.29 (0.05)***	-0.04 (0.05) ^{ns}	0.28 (0.05)***	-0.04 (0.05) ^{ns}
Model Summary						
R ²	0.14	0.10	0.14	0.10	0.16	0.10
MSE	1.37	1.03	1.38	1.03	1.37	1.03
F	14.91***	7.69***	9.25***	7.69***	5.65***	7.69***
df	3,275	4,274	5,273	4,274	9,269	4,274
Conditional Effects on The Influence of AI Monitoring Accountability on Complacency				Effect (SE)	LLCI _{95%}	ULCI _{95%}
Low Task Complexity at Low Task Load				0.14 (0.08)	-0.01	0.30
Low Task Complexity at High Task Load				0.001 (0.11)	-0.22	0.23
High Task Complexity at Low Task Load				0.12 (0.11)	-0.10	0.33
High Task Complexity at High Task Load				0.37 (0.11)	0.15	0.59

Note(s): *** $p < 0.01$; ** $p < 0.05$. ♦ Reverse-coded prior to analysis. †0 = Low Task Load, 1 = High Task Load. Previously, we established that AI overreliance is a closely related concept of interest. Thus, in this study, we control for prior experience and also factor in the potential influence of this concept in the model

Source(s): Authors' own work

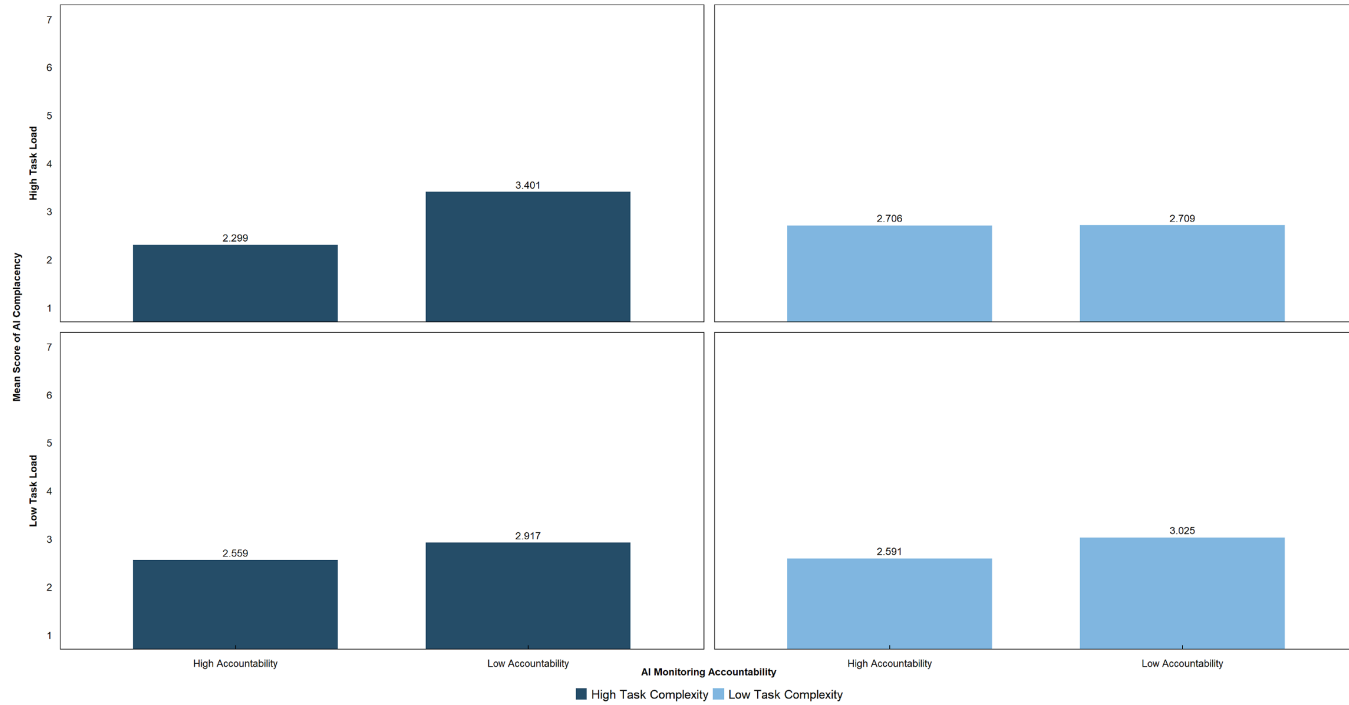


Figure 3. A Visualization of the interaction effect in study 3. Source: Authors' own work

Alkire *et al.*, 2024). Further, recent studies have primarily examined the attributes of decision-makers (Banker and Khetani, 2019; Bonezzi and Osinelli, 2021; Castelo, 2024). While this line of work has been valuable, it has paid comparatively little attention to the situational demands of taskwork that shape how employees interact with AI agents. The current study advances this stream by investigating how task-oriented conditions moderate the relationship between monitoring accountability and AI complacency. The findings suggest that when accountability pressures are low, cognitive stressors such as task complexity and workload can inadvertently heighten complacency. By examining how AI monitoring accountability interacts with task-level stressors, this study advances a more nuanced understanding of when to leverage monitoring accountability. Hence, this study underscores that AI oversight cannot be separated from the broader service task context in which it operates (Le *et al.*, 2023).

Study 4

Design and procedure

Study 4 aims to test [Hypothesis 3](#), which proposes that task interdependence can exacerbate the influence of AI complacency on the willingness to oversee AI output. A 2 (monitoring accountability: high vs. low) \times 2 (task interdependence: interdependent vs. independent) full factorial between-subjects design was employed. The experiment's context was scheduling appointments. Participants were recruited through the Cloud Research platform. Participants served as service coordinators at a local clinic called "Healthy Life," coordinating appointments. In the experiment, the AI assistant failed to account for the doctor's request for a mini break, resulting in a scheduling conflict. The provided information specified the break, but the AI-generated schedule omitted it. Participants then decided whether to send this schedule to the doctor and patient (see [Web Appendix E](#)). The manipulations for AI monitoring accountability were consistent with those in previous studies. Additionally, in the task interdependence condition, participants were informed that their work depended on team input, was closely integrated into the clinic's workflow, and that errors could disrupt the team operations. In the task independence condition, participants were told they would work independently, relying solely on system-provided information, with errors having minimal impact on the operation. The scenario was deemed acceptable for realism ($M = 5.68$, $SD = 0.97$) and believability ($M = 5.70$, $SD = 0.93$).

Manipulation check

Independent-samples *t*-tests were conducted to verify the manipulations. In the high monitoring accountability condition, participants perceived a greater level of accountability ($M = 6.04$, $SD = 1.13$) than in the low accountability condition ($M = 3.66$, $SD = 1.73$), $t(215.88) = 12.87$, $p < 0.01$. In the task interdependence condition, participants perceived a greater level of interdependency ($M = 5.56$, $SD = 0.97$) than in the task independence condition ($M = 3.62$, $SD = 1.79$), $t(186.15) = 10.60$, $p < 0.01$. Thus, the results confirmed the manipulations.

Results

Study 4 regression results are documented in [Table 8](#). The analysis revealed a significant main effect of low monitoring accountability on AI complacency in the base model (β (SE) = 0.14 (0.04), $p < 0.01$), replicating previous findings. In addition, a negative influence of complacency on willingness to oversee AI-generated output was observed (β (SE) = -0.11 (0.05), $p < 0.05$), thereby reconfirming [Hypothesis 1C](#). When task interdependence was added to the model, a significant negative moderation effect emerged. This result confirmed [Hypothesis 3](#), which proposed that the effect of AI complacency on willingness to oversee AI-generated output is intensified by task interdependence (β (SE) = -0.26 (0.09), $p < 0.01$). This interaction effect is visualized in [Figure 4](#), Panel A. Further probing revealed that the effect

Table 8. Study 4 regression results

PROCESS models (unstandardized Scores)	Model 1		Model 2		Model 3	
	AI complacency (3)	Willingness to oversee AI output (4)	AI complacency (3)	Willingness to oversee AI output (4)	AI complacency (3)	Willingness to oversee AI output (4)
Constant	2.49 (0.16)***	4.29 (0.17)***	1.48 (0.23)***	4.42 (0.20)***	-1.43 (0.23)***	4.23 (0.20)***
AI Monitoring Accountability † (1)	0.14 (0.04)***	-0.07 (0.03) ^{ns}	0.13 (0.04)***	-0.07 (0.03)**	0.13 (0.04)***	-0.06 (0.03) ^{ns}
Task Interdependence † (2)						0.06 (0.12) ^{ns}
AI Complacency (3) (2)*(3)		-0.11 (0.05)**		-0.07 (0.05) ^{ns}		0.05 (0.07) ^{ns}
Control Variables						-0.26 (0.09)*** (H3)
AI Overreliance			0.32 (0.06)***	-0.10 (0.05)**	0.32 (0.06)***	-0.10 (0.05)**
Model Summary						
R ²	0.04	0.04	0.15	0.06	0.15	0.09
MSE	1.71	0.98	1.52	0.96	1.52	0.94
F	9.50***	5.41***	21.81***	5.23***	21.81***	4.79***
df	1,248	2,247	2,247	3,246	2, 247	5, 244
Mediation Analysis (Model 1) (1)→(3)→(4)	Effect (Boot SE)	LLCI _{95%}	ULCI _{95%}			
	-0.01 (0.01)	-0.04	0.0001			

Note(s): *** $p < 0.01$; ** $p < 0.05$. † Reverse coded prior to analysis. †0 = Task Independence; 1 = Task Interdependence. As we demonstrated in previous studies, the effect of prior experience with AI was consistently negligible. Thus, in this experiment and onward, we controlled AI overreliance due to its relationship with complacency

Source(s): Authors' own work

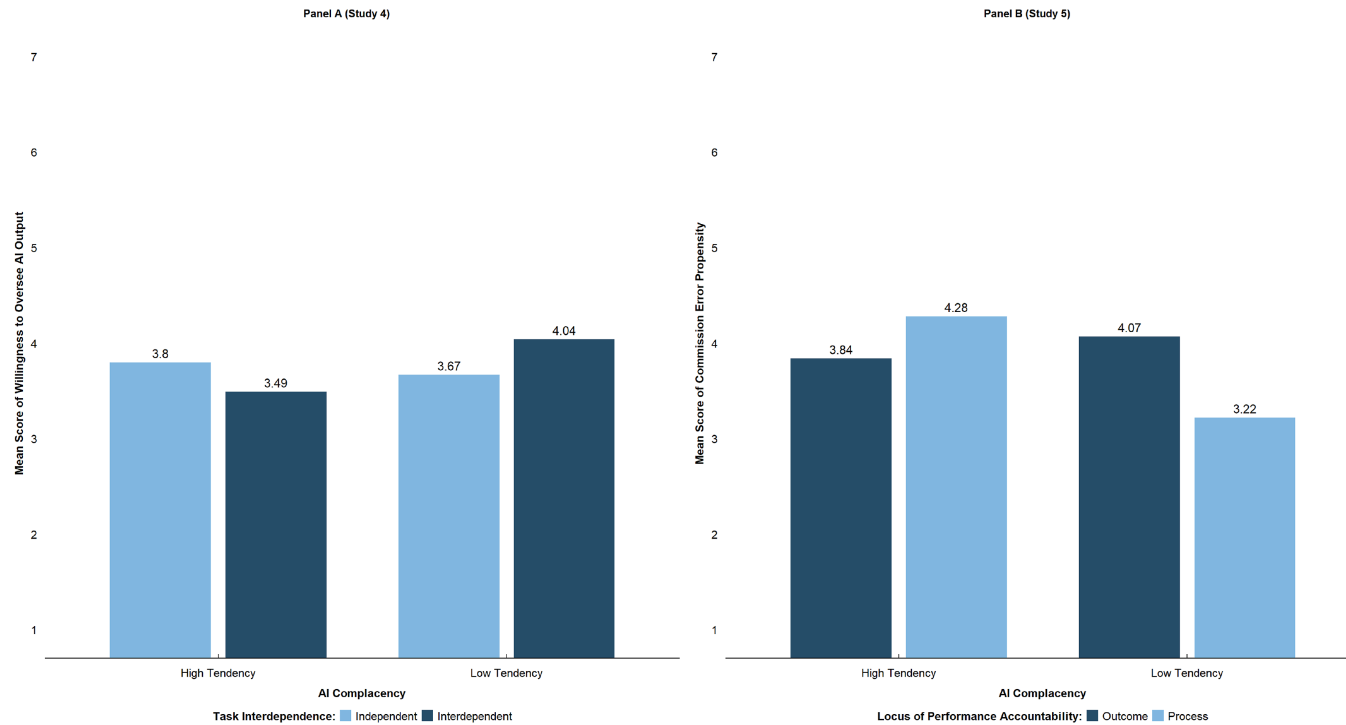


Figure 4. Visualizations of interaction effects in study 4 and 5. Source: Authors' own work

was not significant under task independence (Effect (SE) = 0.05 (0.07), $CI_{95\%} = [-0.08 \text{ to } 0.18]$) but became pronounced under task interdependence (Effect (SE) = -0.21 (0.07), $CI_{95\%} = [-0.35 \text{ to } -0.07]$).

Discussion

This study extends research on Employee-AI collaboration in service provision (Blaurock *et al.*, 2025) by identifying task interdependence as a novel boundary condition in how employees interact with AI agents. While prior work has examined technological substitution and augmentation of frontline employees (Wirtz *et al.*, 2023), the moderating role of team-level contingencies has received little attention (Le *et al.*, 2023). Our findings demonstrate that task interdependence can intensify the negative effects of AI complacency, reducing employees' willingness to monitor AI-generated output. This offers new insights into an inadvertent effect of task interdependence in technology-infused service systems (Noble *et al.*, 2022). Our study underscores that responsible AI adoption is contingent not only on employee attitudes and capabilities (van Riel *et al.*, 2025) but also on the degree of interdependence embedded in service process design.

Study 5

Design and procedure

The objective of Study 5 is to test Hypothesis 4, which proposes that emphasizing outcome-oriented performance accountability will weaken the influence of AI complacency on commission error propensity. A 2 (AI monitoring accountability: high vs. low) x 2 (locus of performance accountability: process vs. outcome) between-subjects design was employed. The study focused on resolving warranty claims, and participants were recruited through the Cloud Research platform. Participants assumed the role of a customer service representative at "ElectroLux Electronics," responsible for handling warranty claims. The key piece of information pertained to the deviation in the reporting of power irregularities. The diagnostic report (i.e. the given information) indicated that the power irregularity might stem from a defect in the TV's internal power management unit. However, the AI-generated decision letter attributed the issue solely to the customer's home power supply and overlooked the warranty policy covering manufacturing defects (see Web Appendix F). Participants then reviewed the report and decided whether to use the AI-generated letter. The monitoring accountability manipulations were consistent with those used in prior studies. Further, in the process-oriented accountability condition, participants were informed that their performance would be evaluated based on adherence to the company's warranty claim review procedures, regardless of the outcome. Conversely, in the outcome-oriented condition, performance was linked to the decision's results, including customer satisfaction and alignment with company interests. The scenario was rated acceptable for realism ($M = 5.97$, $SD = 0.86$) and believability ($M = 6.02$, $SD = 0.84$).

Manipulation check

Independent samples *t*-tests were conducted to verify the manipulations. In the high monitoring accountability condition, participants perceived a significantly greater level of accountability ($M = 5.96$, $SD = 0.99$) than in the low accountability condition ($M = 5.14$, $SD = 1.32$), $t(221) = 5.46$, $p < 0.01$. Locus of performance accountability was measured separately for outcome-oriented and process-oriented aspects. In the outcome-oriented accountability condition, participants rated their performance accountability for the outcome ($M = 5.93$, $SD = 1.02$) higher than for the process ($M = 4.13$, $SD = 1.59$), $t(200.59) = 10.40$, $p < 0.01$. Conversely, in the process-oriented performance accountability condition, participants rated their performance accountability for the process ($M = 6.08$, $SD = 1.01$)

higher than for the outcome ($M = 5.28$, $SD = 1.46$), $t(214.99) = 4.99$, $p < 0.01$. These results confirm the effectiveness of the manipulations.

Results

Study 5 regression results are documented in [Table 9](#). These results first replicated previous findings, demonstrating that low monitoring accountability leads to AI complacency (β (SE) = 0.32 (0.06), $p < 0.01$). Adding the locus of performance accountability in the model revealed a significant moderation effect on the relationship between AI complacency and commission error propensity (β (SE) = -0.53 (0.21), $p < 0.05$). Panel B in [Figure 4](#) illustrates this interaction effect. Conditional effect analysis indicated that the relationship was strongest under process-oriented accountability (Effect (SE) = 0.43 (0.16), $CI_{95\%} = [0.12 \text{ to } 0.74]$), but it weakened and became statistically insignificant under outcome-oriented accountability (Effect (SE) = -0.09 (0.15), $CI_{95\%} = [-0.39 \text{ to } 0.21]$), thereby confirming [Hypothesis 4](#).

Discussion

This study adds to ongoing efforts to support the responsible integration of AI agents in the service workplace ([Cillo and Rubera, 2024](#)) by illustrating how organizational design can help reduce the risk of AI complacency. Placing greater emphasis on outcome-based performance accountability may encourage employees to remain attentive when working with AI agents. In doing so, the findings build on prior research that underscores the importance of Employee-AI collaborative systems, in which employee responsibility for service outcomes jointly produced with AI agents serves as a key governance mechanism ([Blaurock et al., 2025](#)).

General discussion

As AI technology becomes an integral part of productive service work and continues to transform the workspace, a new challenge emerges. Employees may lack vigilance toward potentially fabricated information generated by AI agents. This research underscores the importance of AI monitoring accountability in co-producing services with AI systems ([Alkire et al., 2024](#)). While improving the accuracy and efficiency of LLMs is crucial for addressing confabulated information from a technological standpoint, our study emphasizes the role of employees' accountability in monitoring AI-generated output as a managerial leverage to address AI complacency. Across studies, this research examines how low monitoring accountability influences AI complacency and its consequences, as well as the contingent factors that shape these effects. Studies 1A and 1B established the link between low monitoring accountability and AI complacency. Study 2 further differentiated AI complacency from overreliance. Study 3 demonstrated that task complexity and task load jointly exacerbate the effect of low monitoring accountability on AI complacency. Study 4 revealed that task interdependence worsens the effect of AI complacency on the willingness to oversee AI output. Study 5 showed that outcome-oriented accountability reduced the effect of AI complacency on commission error propensity.

Theoretical implications

Prior research has sought to boost AI adoption by explaining algorithm aversion, mainly from a consumer perspective ([Castelo et al., 2019](#); [Longoni et al., 2019](#); [Longoni and Cian, 2022](#)). It has also highlighted the role of contextual factors ([Garvey et al., 2023](#)) and decision-maker characteristics ([Srinivasan and Sarial-Abi, 2021](#); [Yalcin et al., 2022](#)). In contrast, we engage the emerging "AI abuse" research stream (e.g. [Banker and Khetani, 2019](#)). This perspective focuses on the operational challenges of working with AI agents in service processes. Such a focus marks a necessary shift for effective Employee-AI collaboration ([Le et al., 2025](#)). Responding to calls for responsible, governance-first AI deployment ([Alkire et al., 2024](#);

Table 9. Study 5 regression results

PROCESS models (unstandardized scores)	Model 1		Model 2		Model 3	
	AI complacency (3)	Commission error propensity (4)	AI complacency (3)	Commission error propensity (4)	AI complacency (3)	Commission error propensity (4)
Constant	1.99 (0.17)***	3.46 (0.36)***	1.40 (0.23)***	3.05 (0.44)***	-1.35 (0.24)***	3.43 (0.47)***
AI Monitoring Accountability ♦ (1)	0.31 (0.06)***	-0.07 (0.11) ^{ns}	0.32 (0.06)***	-0.05 (0.11) ^{ns}	0.32 (0.06)***	-0.08 (0.11) ^{ns}
AI Complacency (3)		0.20 (0.11) ^{ns}		0.16 (0.11) ^{ns}		0.43 (0.16)***
Locus of Performance Accountability† (2) (2)*(3)						0.20 (0.26) ^{ns} -0.53 (0.21)** (H4)
Control Variables						
AI Overreliance			0.18 (0.05)***	0.15 (0.09) ^{ns}	0.18 (0.05)***	0.16 (0.09) ^{ns}
Model Summary						
R ²	0.10	0.01	0.14	0.02	0.14	0.05
MSE	1.36	4.00	1.30	3.97	1.30	3.89
F	26.06***	1.71 ^{ns}	19.69***	2.01 ^{ns}	19.69***	2.62**
df	1,239	2,238	2,238	3,237	2, 238	5, 235
Mediation Analysis (Model 1) (1)→(3)→(4)	Effect (Boot SE)	LLCI _{95%}	ULCI _{95%}			
	0.06 (0.04)	-0.01	0.15			

Note(s): ****p* < 0.01; ***p* < 0.05. ♦ Reverse coded prior to analysis. †0 = Process; 1 = Outcome

Source(s): Authors' own work

Sigala *et al.*, 2024; Wirtz and Stock-Homburg, 2025), the current research draws on contingency theory's principle of preemptive self-criticism (Lerner and Tetlock, 1999). It highlights the malleability of accountability in the context of Employee-AI collaboration in co-producing services (Cillo and Rubera, 2024). Through this lens, the current research demonstrates that weak monitoring accountability fosters AI complacency. Thus, we identify and provide firsthand evidence of a prominent phenomenon in AI abuse in service operations and outline avenues for further inquiry in this stream. We also differentiate AI complacency from overreliance (Banker and Khetani, 2019), in which weak monitoring and accountability are more closely associated with complacency than with overreliance. This distinction adds nuance to accounts of human bias toward emergent technologies in marketing (Dowling *et al.*, 2020).

Additionally, this research demonstrates how environmental factors collectively moderate the relationship between AI monitoring accountability, complacency, and its downstream consequences. By drawing on task-, team-, and organization-level contingencies, we offer a coherent account of when organizational contexts enable AI complacency and what can be done to curb it. Although calls for responsible AI agent deployment are growing (Alkire *et al.*, 2024), recent work still concentrates on transactive interactions in Employee-AI collaboration (e.g. Blaurock *et al.*, 2025; Le *et al.*, 2025). This focus underplays the formative role of the service work environment (vvan Riel *et al.*, 2025). We extend research on environmental influences in Employee-AI interaction (Kaartemo and Helkkula, 2025) by demonstrating how structural features across task, team, and organizational strata interact to shape this dynamic (Le *et al.*, 2023), particularly with the emergence of AI complacency. This perspective aligns with the growing emphasis in service research on designing resilient, human-centered service systems amid increasing automation (Alkire *et al.*, 2024). Accordingly, we provide initial theoretical guidance on configuring accountability structures and the service work environment to preserve employees' oversight of AI-generated content.

Managerial implications

As AI rapidly reshapes service processes (Cillo and Rubera, 2024), its practical integration into workflows has become not only inevitable but strategically critical (Grewal *et al.*, 2024a, b). While new AI technologies offer clear gains in productivity, it also introduces novel operational challenges. One such challenge is AI complacency, which we investigate throughout this research. If left unchecked, AI complacency can undermine productivity and, more broadly, compromise the quality of service provision. This study provides empirically grounded, actionable guidance for service leaders seeking to harness the advantages of AI agents while minimizing the associated complacency challenge. Specifically, we identify and address a novel lever – employee monitoring accountability. We offer a set of practical recommendations, summarized in Figure 5, that managers can use to establish robust oversight structures, define human-in-the-loop roles, and implement safeguards to mitigate complacency-driven failures. The goal is to guide the thoughtful integration of AI agents into service processes, balancing innovation with accountability (Alkire *et al.*, 2024).

Encouraging systematic accountability mechanisms when using AI agents in service processes, with detailed guiding frameworks, will lay a foundation to enhance human employees' vigilance when using AI-generated outputs. In a recent survey conducted by Harvard Business Review Analytic Services with 257 corporate respondents, 89% agreed that an organizational guiding principle is crucial for navigating the risks associated with AI agents (HBR Analytic Service, 2024). Our findings consistently demonstrate that a low perception of accountability is the primary driver of AI complacency. Therefore, implementing systematic accountability guidelines to foster a culture of responsibility among employees could mitigate AI complacency. In AI governance, existing protocols for trustworthy AI design should focus on system design and emphasize human involvement and accountability (i.e. human-in-the-loop) in these systems (Deloitte AI Institute, 2023b). When individuals are aware that their use

	Managerial Recommendations	Insights in Practice
Key Empirical Findings		
A perception of low AI monitoring accountability increases likelihood of making false confirmation decision due to an increase in AI complacency.	Encouraging systematic accountability mechanisms when using AI agents in service processes with detailed guiding frameworks.	Deloitte's trustworthy AI framework fosters accountability by holding employees and leadership responsible for AI outcomes and enabling effective scrutiny (Deloitte 2023).
A compound impact of high task complexity and task load can intensify the effect of low AI monitoring accountability on the development of AI complacency.	Streamlining service workflows by eliminating unnecessary steps and standardizing procedures through process automation to reduce burden on employees' attentional resources.	IBM utilizes Business Automation Workflow to simplify processes and reduce task load by automating responses to FAQs, thereby freeing human cognitive resources (ibm.com/workflow).
Task interdependence can exacerbate the negative influence of the AI complacency on willingness to oversee AI-generated output.	Clearly define individual accountability and foster a shared accountability culture in teamwork, especially when employees utilize AI agents for their respective contributions.	Microsoft Team Copilot fosters cross-team collaborations and members are individually accountable for verifying its accuracy when using it for their part in team meetings (Carter 2024).
Outcome (vs process) accountability weakens the influence of AI complacency on commission error propensity.	Setting clear performance expectations and providing frameworks to promote employees' accountability for the outcomes of their decisions, even when assisted by AI.	Accenture's AI compliance program documents enterprise-wide governance structures with clear roles, policies, and expectations to ensure responsible AI use (Accenture 2024).

Figure 5. An overview of managerial recommendations. Source: Authors' own work

of AI is monitored, they are more likely to critically validate the information produced by these agents. This heightened vigilance empowers employees to proactively address potential issues with AI-generated output before incorporating it into other aspects of marketing processes, thereby limiting the development of AI complacency. For example, Deloitte's trustworthy AI framework highlights employee accountability when using AI agents at work and guides the ethical use of AI agents in the workplace (Deloitte AI Institute, 2023a).

Streamlining service workflows by eliminating unnecessary steps and standardizing procedures through process automation to reduce the burden on employees' attentional resources. By simplifying processes, firms can free up employees' cognitive resources from unnecessary tasks, enabling them to engage more effectively with AI agents in meaningful work. This increased engagement could help mitigate the impact of low monitoring accountability on AI complacency, as employees would have more cognitive resources available to monitor the generated output. For example, IBM uses business automation workflows to simplify processes and reduce task loads, freeing employees' cognitive resources for higher-value tasks (ibm.com/workflow).

Clearly define individual accountability and foster a shared accountability culture within the team, especially when employees use AI agents to make their contributions. Task interdependence is a defining characteristic of cross-functional collaboration (Wageman, 1995), and our findings indicate that it can amplify the negative influence of AI complacency on employees' willingness to oversee AI-generated output. While teamwork is essential in any organizational setting, marketing managers should be mindful of the challenges that can emerge as AI agents become increasingly integrated into various service workflows. Establishing clear lines of individual accountability, alongside a culture of shared accountability for team efforts, encourages vigilance in using AI agents for group tasks. This approach helps mitigate the unintended negative impact of task interdependence.

A practical illustration of this dynamic is Microsoft Teams Copilot, which helps coordinate cross-team meetings by automating tasks such as scheduling, agenda creation, and follow-ups. In these settings, team members collectively share responsibility for ensuring the accuracy and relevance of Copilot's outputs before acting on them (Carter, 2024). For example, they double-check generated schedules and assigned tasks to verify their correctness before initiating any work. This shared vigilance enhances the effectiveness of AI-generated output by promoting a proactive oversight culture within team workflows.

Setting clear performance expectations and providing frameworks to promote employees' accountability for the outcomes of their decisions, even when assisted by AI, as we demonstrated that outcome accountability weakens the influence of complacent tendencies on commission error propensity. Outcome accountability emphasizes the need to justify past decisions, creating psychological pressure that motivates users to remain vigilant. Thus,

setting clear expectations can serve as a safeguard to reduce the impact of complacent tendencies on commission error propensity. An illustrative example is Accenture's AI compliance program (Accenture, 2024). Accenture has established enterprise-wide governance structures that outline clear policies and performance expectations to ensure the responsible use of AI. By defining comprehensive policies, Accenture creates a framework that holds employees accountable for the outcomes of their decisions.

Limitations and future research

The current research is not without limitations. Our studies employ scenario-based experiments with online participants rather than practicing employees in service organizations. Role assumption cannot fully reproduce the behavior of professionals in the industries represented by our scenarios. Nonetheless, the controlled design affords a novel, causally identified test of how insufficient monitoring accountability drives AI complacency. The results remain robust under selective recruitment that matched the scenario characteristics (Study 1B). Nonetheless, future work should examine more naturalistic settings. Promising paths include field or quasi-experiments with partner firms to observe employees' monitoring behavior when interacting with AI agents and using the generated outputs, as well as analyses of operational data to corroborate the findings.

Additionally, this research utilizes simplified task scenarios, which might limit the ecological validity of the findings; nonetheless, it is noteworthy that the scenario realism ratings across studies are acceptable. In high-stakes service contexts, the decision to trust or reject AI-generated output could also be shaped by other extraneous factors such as system design (Shin, 2021) or subjective norms (Mehrotra *et al.*, 2024). These factors were not adequately taken into account in the experiments, given the current research's focus on the organizational design level (e.g. Blaurock *et al.*, 2025), particularly the malleability of AI monitoring accountability. Future research could examine these aspects in light of the emergence of AI complacency.

Conceptually, the findings would benefit from a different perspective on how employees process AI-generated outputs and allocate attention. Future theorizing and tests could focus on user-oriented mechanisms such as confidence calibration to explain the emergence of AI complacency. Integrating these perspectives would sharpen the understanding of AI complacency in service provision settings. Further, the present studies do not benchmark the salience of AI complacency in a human-to-AI collaboration setting against when using AI agents in human-to-human collaboration, where validation practices are more socially robust. Future research should adopt comparative settings to examine the implications of AI monitoring, accountability, and complacency, as this could reveal novel boundary conditions and outcomes; this aligns with recent calls for a deeper understanding of Employee-AI collaboration in service contexts. (Kunz *et al.*, 2025).

Although AI is an innovative technology poised to transform work processes (Cillo and Rubera, 2024), the risks associated with its use must not be underestimated (Grewal *et al.*, 2024a, b). We hope this research contributes to the ongoing effort to integrate AI agents into service work in a responsible manner. To that end, we propose additional research avenues, documented in Figure 6, to encourage future scholars to explore the challenging aspects of working with AI agents. We highlight four investigation frontiers.

On the organizational frontier, recent research has underscored the importance of aligning control and accountability measures to mitigate AI-related risks (Grote *et al.*, 2024). Building on our empirical findings, future work could examine how such alignment might be integrated into marketing processes (e.g. Le *et al.*, 2025). In addition, there is a need to design and implement comprehensive frameworks for responsible AI adoption that offer clear guidance on user accountability. Extending this perspective, firms could explore initiatives such as corporate digital responsibility as an alternative means of managing AI complacency. Further, since complacency has been linked to limited attentional resources (Parasuraman and Manzey,

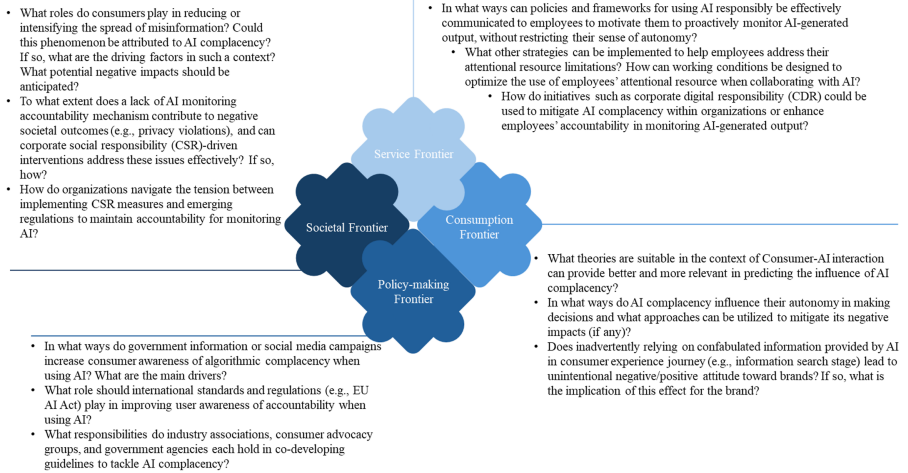


Figure 6. Future research avenues. Source: Authors' own work

2010), firms should explore strategies to optimize work allocation to utilize employees' attentional resources better.

On the consumption frontier, previous research has highlighted consumers' susceptibility to overdependence on AI (Banker and Khetani, 2019). Exploring AI complacency within consumer contexts could yield valuable insights into its impact on consumer well-being. Researchers might examine the implications of AI monitoring accountability and complacency throughout the customer experience journey. Another promising direction involves examining the influence of AI complacency on consumer-robot power relationships. Additionally, future studies should investigate the implications for consumer autonomy when complacency is present in the consumer journey. Moreover, as this study draws on contingency theory, future research could explore alternative theoretical perspectives to provide new insights into the consumer-focused dimensions of complacency. This approach might offer a broader theoretical understanding of the phenomenon. Addressing these issues would contribute to a more nuanced understanding of AI complacency and foster responsible interactions with AI technologies in consumer settings.

On the policy-making frontier, future research should explore how government or social media campaigns can be leveraged to educate people about AI complacency effectively. Further, a broader inquiry could examine the role of international standards and regulations, such as the EU AI Act, in enhancing user awareness of accountability when AI agents are used for personal purposes. Concurrently, additional studies could investigate how industry-level organizations contribute to the development of ethical guidelines for AI usage.

On the societal frontier, AI has been used to spread misinformation (Shin, 2024). One promising direction is to explore the impact of enforcing accountability measures to address this issue, especially given recent research showing that lower AI literacy can lead to greater receptivity to AI (Tully *et al.*, 2025). The role of consumers warrants investigation to understand the conditions under which they may either reduce or intensify the spread of misinformation. Further, corporate social responsibility could be examined for its potential to address complacency in light of emerging regulations on AI use.

Supplementary material

The supplementary material for this article can be found online

References

- Acar, A.O. (2023), "A practical guide for marketers who want to use GenAI", *Harvard Business Review*, available at: <https://hbr.org/2023/12/a-practical-guide-for-marketers-who-want-to-use-genai> (accessed 14 October 2024).
- Accenture (2024), *Accenture's Blueprint for Responsible AI*, Accenture, available at: <https://www.accenture.com/us-en/case-studies/data-ai/blueprint-responsible-ai> (accessed 23 December 2024).
- Afroogh, S., Akbari, A., Malone, E., Kargar, M. and Alambeigi, H. (2024), "Trust in AI: progress, challenges, and future directions", *Humanities and Social Sciences Communications*, Vol. 11 No. 1, pp. 1-30, doi: [10.1057/s41599-024-04044-8](https://doi.org/10.1057/s41599-024-04044-8).
- Alkire, L., Bilgihan, A., Bui, M., Buoye, A.J., Dogan, S. and Kim, S. (2024), "RAISE: leveraging responsible AI for service excellence", *Journal of Service Management*, Vol. 35 No. 4, pp. 490-511, doi: [10.1108/josm-11-2023-0448](https://doi.org/10.1108/josm-11-2023-0448).
- Bahner, J.E., Hüper, A.D. and Manzey, D. (2008), "Misuse of automated decision aids: complacency, automation bias and the impact of training experience", *International Journal of Human-Computer Studies*, Vol. 66 No. 9, pp. 688-699, doi: [10.1016/j.ijhcs.2008.06.001](https://doi.org/10.1016/j.ijhcs.2008.06.001).
- Banker, S. and Khetani, S. (2019), "Algorithm overdependence: how the use of algorithmic recommendation systems can increase risks to consumer well-being", *Journal of Public Policy and Marketing*, Vol. 38 No. 4, pp. 500-515, doi: [10.1177/0743915619858057](https://doi.org/10.1177/0743915619858057).
- Blaurock, M., Büttgen, M. and Schepers, J. (2025), "Designing collaborative intelligence systems for employee-AI service Co-production", *Journal of Service Research*, Vol. 28 No. 4, pp. 544-562, doi: [10.1177/10946705241238751](https://doi.org/10.1177/10946705241238751).
- Bohannon, M. (2023), "Judge fines two lawyers for using fake cases from ChatGPT", *Forbes*, available at: <https://www.forbes.com/sites/mollybohannon/2023/06/22/judge-fines-two-lawyers-for-using-fake-cases-from-chatgpt/> (accessed 4 December 2024).
- Bonezzi, A. and Ostinelli, M. (2021), "Can algorithms legitimize discrimination?", *Journal of Experimental Psychology: Applied*, Vol. 27 No. 2, pp. 447-459, doi: [10.1037/xap0000294](https://doi.org/10.1037/xap0000294).
- Carter, R. (2024), *Can AI Be Part of Cross-Team Collaboration?*, UC Today, available at: <https://www.uctoday.com/collaboration/can-ai-be-part-of-cross-team-collaboration/> (accessed 26 November 2024).
- Castelo, N. (2024), "Perceived corruption reduces algorithm aversion", *Journal of Consumer Psychology*, Vol. 34 No. 2, pp. 326-333, doi: [10.1002/jcpy.1373](https://doi.org/10.1002/jcpy.1373).
- Castelo, N., Bos, M.W. and Lehmann, D.R. (2019), "Task-dependent algorithm aversion", *Journal of Marketing Research*, Vol. 56 No. 5, pp. 809-825, doi: [10.1177/0022243719851788](https://doi.org/10.1177/0022243719851788).
- Chen, C. (2024), "How consumers respond to service failures caused by algorithmic mistakes: the role of algorithmic interpretability", *Journal of Business Research*, Vol. 176, pp. 1-14, doi: [10.1016/j.jbusres.2024.114610](https://doi.org/10.1016/j.jbusres.2024.114610).
- Christian, J. (2023), "Magazine publishes serious errors in first AI-generated health article", *Futurism*, available at: <https://futurism.com/neoscope/magazine-mens-journal-errors-ai-health-article> (accessed 19 August 2024).
- Chui, M., Hazan, E., Roberts, R., Singla, A., Smaje, K., Sukharrevsky, A., Yee, L. and Zimmel, R. (2023), *The Economic Potential of Generative AI: The Next Productivity Frontier*, McKinsey&Company, available at: <https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/the-economic-potential-of-generative-ai-the-next-productivity-frontier#/> (accessed 20 September 2024).
- Cillo, P. and Rubera, G. (2024), "Generative AI in innovation and marketing processes: a roadmap of research opportunities", *Journal of the Academy of Marketing Science*, Vol. 53 No. 3, pp. 1-18, doi: [10.1007/s11747-024-01044-7](https://doi.org/10.1007/s11747-024-01044-7).
- Cooper, D.A. (2024), "Diffusion of responsibility for actions with advice", *Journal of Behavioral Decision Making*, Vol. 37 No. 4, pp. 1-15, doi: [10.1002/bdm.2415](https://doi.org/10.1002/bdm.2415).

- Courtright, S.H., Thurgood, G.R., Stewart, G.L. and Pierotti, A.J. (2015), "Structural interdependence in teams: an integrative framework and meta-analysis", *Journal of Applied Psychology*, Vol. 100 No. 6, pp. 1825-1846, doi: [10.1037/apl0000027](https://doi.org/10.1037/apl0000027).
- Davenport, T. and High, P. (2024), "How gen AI and analytical AI differ – and when to use each", *Harvard Business Review*, available at: <https://hbr.org/2024/12/how-gen-ai-and-analytical-ai-differ-and-when-to-use-each> (accessed 4 January 2025).
- Deloitte AI Institute (2023a), *The National Institute of Standards and Technology Artificial Intelligence Risk Management Framework (NIST AI RMF) Emphasizes the Need for Trustworthy AI™*, Deloitte, available at: <https://www2.deloitte.com/content/dam/Deloitte/us/Documents/deloitte-analytics/us-ai-institute-NIST-ai-risk-management.pdf> (accessed 15 January 2025).
- Deloitte AI Institute (2023b), *Proactive Risk Management in Generative AI*, Deloitte, available at: <https://www2.deloitte.com/us/en/pages/consulting/articles/responsible-use-of-generative-ai.html> (accessed 15 January 2025).
- Dietvorst, B.J., Simmons, J.P. and Massey, C. (2015), "Algorithm aversion: people erroneously avoid algorithms after seeing them err", *Journal of Experimental Psychology: General*, Vol. 144 No. 1, pp. 114-126, doi: [10.1037/xge0000033](https://doi.org/10.1037/xge0000033).
- Donaldson, L. (2001), *The Contingency Theory of Organizations*, Sage, CA.
- Dowling, K., Guhl, D., Klapper, D., Spann, M., Stich, L. and Yegoryan, N. (2020), "Behavioral biases in marketing", *Journal of the Academy of Marketing Science*, Vol. 48 No. 3, pp. 449-477, doi: [10.1007/s11747-019-00699-x](https://doi.org/10.1007/s11747-019-00699-x).
- Ede-Osifo, U. (2023), *College Instructor Put on Blast for Accusing Students of Using ChatGPT on Final Assignments*, NBC News, available at: <https://www.nbcnews.com/tech/chatgpt-texas-college-instructor-backlash-rcna84888> (accessed 18 January 2025).
- Ekin, M., Krejtz, K., Duarte, C., Duchowski, A.T. and Krejtz, I. (2025), "Prediction of intrinsic and extraneous cognitive load with oculometric and biometric indicators", *Scientific Reports*, Vol. 15 No. 1, pp. 1-14, doi: [10.1038/s41598-025-89336-y](https://doi.org/10.1038/s41598-025-89336-y).
- Endsley, M.R. (1995), "Toward a theory of situation awareness in dynamic systems", *Human Factors*, Vol. 37 No. 1, pp. 32-64, doi: [10.1518/001872095779049543](https://doi.org/10.1518/001872095779049543).
- Endsley, M.R. (2023), "Ironies of artificial intelligence", *Ergonomics*, Vol. 66 No. 11, pp. 1656-1668, doi: [10.1080/00140139.2023.2243404](https://doi.org/10.1080/00140139.2023.2243404).
- Funke, J. (2010), "Complex problem solving: a case for complex cognition?", *Cognitive Processing*, Vol. 11 No. 2, pp. 133-142, doi: [10.1007/s10339-009-0345-0](https://doi.org/10.1007/s10339-009-0345-0).
- Gaczek, P., Pozharliev, R., Leszczyński, G. and Zieliński, M. (2023), "Overcoming consumer resistance to AI in general health care", *Journal of Interactive Marketing*, Vol. 58 Nos 2-3, pp. 321-338, doi: [10.1177/10949968221151061](https://doi.org/10.1177/10949968221151061).
- Garvey, A.M., Kim, T. and Duhachek, A. (2023), "Bad news? Send an AI. Good news? Send a human", *Journal of Marketing*, Vol. 87 No. 1, pp. 10-25, doi: [10.1177/00222429211066972](https://doi.org/10.1177/00222429211066972).
- Goodman, J.K. and Irmak, C. (2013), "Having versus consuming: failure to estimate usage frequency makes consumers prefer multifeature products", *Journal of Marketing Research*, Vol. 50 No. 1, pp. 44-54, doi: [10.1509/jmr.10.0396](https://doi.org/10.1509/jmr.10.0396).
- Grewal, D., Guha, A. and Becker, M. (2024a), "AI is changing the world: achieving the promise, minimizing the peril", *Journal of Macromarketing*, Vol. 44 No. 4, pp. 936-947, doi: [10.1177/02761467241289573](https://doi.org/10.1177/02761467241289573).
- Grewal, D., Satornino, C.B., Davenport, T. and Guha, A. (2024b), "How generative AI is shaping the future of marketing", *Journal of the Academy of Marketing Science*, Vol. 53 No. 3, pp. 1-21, doi: [10.1007/s11747-024-01064-3](https://doi.org/10.1007/s11747-024-01064-3).
- Grinschgl, S., Papenmeier, F. and Meyerhoff, H.S. (2023), "Mutual interplay between cognitive offloading and secondary task performance", *Psychonomic Bulletin and Review*, Vol. 30 No. 6, pp. 2250-2261, doi: [10.3758/s13423-023-02312-3](https://doi.org/10.3758/s13423-023-02312-3).

- Grote, G., Parker, S.K. and Crowston, K. (2024), "Taming artificial intelligence: a theory of control-accountability alignment among AI developers and users", *Academy of Management Review*. doi: [10.5465/amr.2023.0117](https://doi.org/10.5465/amr.2023.0117).
- Haduong, N. and Smith, N.A. (2024), *Raising the Stakes: Performance Pressure Improves AI-Assisted Decision Making*, arXiv e-prints, pp. 1-20.
- Hagafors, R. and Brehmer, B. (1983), "Does having to justify one's judgments change the nature of the judgment process?", *Organizational Behavior and Human Performance*, Vol. 31 No. 2, pp. 223-232, doi: [10.1016/0030-5073\(83\)90122-8](https://doi.org/10.1016/0030-5073(83)90122-8).
- Hart, S.G. and Staveland, L.E. (1988), "Development of NASA-TLX (Task Load Index): results of empirical and theoretical research", *Advances in Psychology*, Elsevier, pp. 139-183.
- Harvard Business Review Analytic Services (2024), "Realizing the generative AI opportunity: embracing change to create business value", *Harvard Business Review*, available at: <https://hbr.org/sponsored/2025/01/realizing-the-generative-ai-opportunity> (accessed 14 February 2025).
- Hayes, A.F. (2017), *Introduction to Mediation, Moderation, and Conditional Process Analysis: A Regression-Based Approach*, Guilford Publications, New York.
- Helal, M.Y., Elgendy, I.A., Albashrawi, M.A., Dwivedi, Y.K., Al-Ahmadi, M.S. and Jeon, I. (2025), "The impact of generative AI on critical thinking skills: a systematic review, conceptual framework and future research directions", *Information Discovery and Delivery*, Vol. ahead-of-print No. ahead-of-print, doi: [10.1108/idd-05-2025-0125](https://doi.org/10.1108/idd-05-2025-0125).
- Hermann, E. and Puntoni, S. (2025), "Generative AI in marketing and principles for ethical design and deployment", *Journal of Public Policy and Marketing*, Vol. 44 No. 3, pp. 332-349, doi: [10.1177/07439156241309874](https://doi.org/10.1177/07439156241309874).
- Huang, M.H. and Rust, R.T. (2024), "The caring machine: feeling AI for customer care", *Journal of Marketing*, Vol. 88 No. 5, pp. 1-23, doi: [10.1177/00222429231224748](https://doi.org/10.1177/00222429231224748).
- Hult, G.T.M. (2011), "Toward a theory of the boundary-spanning marketing organization and insights from 31 organization theories", *Journal of the Academy of Marketing Science*, Vol. 39 No. 4, pp. 509-536, doi: [10.1007/s11747-011-0253-6](https://doi.org/10.1007/s11747-011-0253-6).
- Kaartemo, V. and Helkkula, A. (2025), "Human-AI resource relations in value co-creation in service ecosystems", *Journal of Service Management*, Vol. 36 No. 2, pp. 291-306, doi: [10.1108/josm-03-2023-0104](https://doi.org/10.1108/josm-03-2023-0104).
- Karakas, C., Brock, D. and Lakhotia, A. (2023), "Leveraging ChatGPT in the pediatric neurology clinic: practical considerations for use to improve efficiency and outcomes", *Pediatric Neurology*, Vol. 148, pp. 157-163, doi: [10.1016/j.pediatrneurol.2023.08.035](https://doi.org/10.1016/j.pediatrneurol.2023.08.035).
- Karakikes, M. and Nathanael, D. (2023), "The effect of cognitive workload on decision authority assignment in human-robot collaboration", *Cognition, Technology and Work*, Vol. 25 No. 1, pp. 31-43, doi: [10.1007/s10111-022-00719-x](https://doi.org/10.1007/s10111-022-00719-x).
- Keller, R.T. (1994), "Technology-information processing fit and the performance of R&D project groups: a test of contingency theory", *Academy of Management Journal*, Vol. 37 No. 1, pp. 167-179.
- Kunz, W.H. and Wirtz, J. (2024), "Corporate digital responsibility (CDR) in the age of AI: implications for interactive marketing", *Journal of Research in Interactive Marketing*, Vol. 18 No. 1, pp. 31-37, doi: [10.1108/jrim-06-2023-0176](https://doi.org/10.1108/jrim-06-2023-0176).
- Kunz, W.H., Sajtos, L. and Flavián, C. (2025), "Beyond replacement: human-machine collaboration in the age of AI", *Journal of Service Management*, Vol. 36 No. 4, pp. 1-19, doi: [10.1108/josm-04-2025-0194](https://doi.org/10.1108/josm-04-2025-0194).
- Langer, M., Baum, K. and Schlicker, N. (2024), "Effective human oversight of AI-based systems: a signal detection perspective on the detection of inaccurate and unfair outputs", *Minds and Machines*, Vol. 35 No. 1, pp. 1-30, doi: [10.1007/s11023-024-09701-0](https://doi.org/10.1007/s11023-024-09701-0).
- Le, K.B.Q. and Cayrat, C. (2024), "Howdy, Robo-Partner: exploring artificial companionship and its stress-alleviating potential for service employees", *Journal of Service Management*, Vol. 36 No. 4, pp. 631-658, doi: [10.1108/josm-09-2023-0408](https://doi.org/10.1108/josm-09-2023-0408).

- Le, K.B.Q., Sajtos, L. and Fernandez, K.V. (2023), "Employee-(ro)bot collaboration in service: an interdependence perspective", *Journal of Service Management*, Vol. 34 No. 2, pp. 176-207, doi: [10.1108/josm-06-2021-0232](https://doi.org/10.1108/josm-06-2021-0232).
- Le, K.B.Q., Sajtos, L., Kunz, W.H. and Fernandez, K.V. (2025), "The future of work: understanding the effectiveness of collaboration between human and digital employees in service", *Journal of Service Research*, Vol. 28 No. 1, pp. 186-205, doi: [10.1177/10946705241229419](https://doi.org/10.1177/10946705241229419).
- Lerner, J.S. and Tetlock, P.E. (1999), "Accounting for the effects of accountability", *Psychological Bulletin*, Vol. 125 No. 2, pp. 255-275, doi: [10.1037/0033-2909.125.2.255](https://doi.org/10.1037/0033-2909.125.2.255).
- Lin, B. (2023), "How did companies use generative AI in 2023? Here's a look at five early adopters", *The Wall Street Journal*, available at: <https://www.wsj.com/articles/how-did-companies-use-generative-ai-in-2023-heres-a-look-at-five-early-adopters-6e09c6b3>
- Liu, P. and Li, Z. (2012), "Task complexity: a review and conceptualization framework", *International Journal of Industrial Ergonomics*, Vol. 42 No. 6, pp. 553-568, doi: [10.1016/j.ergon.2012.09.001](https://doi.org/10.1016/j.ergon.2012.09.001).
- Longoni, C. and Cian, L. (2022), "Artificial intelligence in utilitarian vs. hedonic contexts: the 'word-of-machine' effect", *Journal of Marketing*, Vol. 86 No. 1, pp. 91-108, doi: [10.1177/0022242920957347](https://doi.org/10.1177/0022242920957347).
- Longoni, C., Bonezzi, A. and Morewedge, C.K. (2019), "Resistance to medical artificial intelligence", *Journal of Consumer Research*, Vol. 46 No. 4, pp. 629-650, doi: [10.1093/jcr/ucz013](https://doi.org/10.1093/jcr/ucz013).
- Longoni, C., Cian, L. and Kyung, E.J. (2023), "Algorithmic transference: people overgeneralize failures of AI in the government", *Journal of Marketing Research*, Vol. 60 No. 1, pp. 170-188, doi: [10.1177/00222437221110139](https://doi.org/10.1177/00222437221110139).
- Lopez, A. and Garza, R. (2023), "Consumer bias against evaluations received by artificial intelligence: the mediation effect of lack of transparency anxiety", *Journal of Research in Interactive Marketing*, Vol. 17 No. 6, pp. 831-847, doi: [10.1108/jrim-07-2021-0192](https://doi.org/10.1108/jrim-07-2021-0192).
- Mehrotra, S., Degachi, C., Vereschak, O., Jonker, C.M. and Tielman, M.L. (2024), "A systematic review on fostering appropriate trust in Human-AI interaction: trends, opportunities and challenges", *ACM Journal on Responsible Computing*, Vol. 1 No. 4, pp. 1-45, doi: [10.1145/3696449](https://doi.org/10.1145/3696449).
- Meijman, T.F. and Mulder, G. (2013), "Psychological aspects of workload", in Drenth, P.J.D., Thierry, H. and de Wolff, C.J. (Eds), *A Handbook of Work and Organizational Psychology*, Psychology Press, Sussex, pp. 5-33.
- Menguc, B., Auh, S. and Uslu, A. (2013), "Customer knowledge creation capability and performance in sales teams", *The Journal of the Academy of Marketing Science*, Vol. 41 No. 1, pp. 19-39.
- Metzger, U., Duley, J.A., Abbas, R. and Parasuraman, R. (2000), "Effects of variable-priority training on automation-related complacency: performance and eye movements", *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, Los Angeles, SAGE Publications, pp. 346-349.
- Molloy, R. and Parasuraman, R. (1996), "Monitoring an automated system for a single failure: vigilance and task complexity effects", *Human Factors*, Vol. 38 No. 2, pp. 311-322, doi: [10.1518/001872096779048093](https://doi.org/10.1518/001872096779048093).
- Moriuchi, E. and Murdy, S. (2024), "The role of robots in the service industry: factors affecting human-robot interactions", *International Journal of Hospitality Management*, Vol. 118, pp. 1-11, doi: [10.1016/j.ijhm.2023.103682](https://doi.org/10.1016/j.ijhm.2023.103682).
- Noble, S.M., Mende, M., Grewal, D. and Parasuraman, A. (2022), "The Fifth Industrial Revolution: how harmonious human-machine collaboration is triggering a retail and service [r]evolution", *Journal of Retailing*, Vol. 98 No. 2, pp. 199-208, doi: [10.1016/j.jretai.2022.04.003](https://doi.org/10.1016/j.jretai.2022.04.003).
- Parasuraman, R. and Manzey, D.H. (2010), "Complacency and bias in human use of automation: an attentional integration", *Human Factors*, Vol. 52 No. 3, pp. 381-410, doi: [10.1177/0018720810376055](https://doi.org/10.1177/0018720810376055).
- Romeo, G. and Conti, D. (2025), "Exploring automation bias in human-AI collaboration: a review and implications for explainable AI", *AI & Society*, Vol. 41, pp. 1-20, doi: [10.1007/s00146-025-02422-7](https://doi.org/10.1007/s00146-025-02422-7).

- Saluja, S., Sinha, S. and Goel, S. (2024), "Loafing in the era of generative AI", *Organizational Dynamics*, Vol. 54 No. 3, pp. 1-6, doi: [10.1016/j.orgdyn.2024.101101](https://doi.org/10.1016/j.orgdyn.2024.101101).
- Sarter, N.B. and Woods, D.D. (1995), "How in the world did we ever get into that mode? Mode error and awareness in supervisory control", *Human Factors*, Vol. 37 No. 1, pp. 5-19, doi: [10.1518/001872095779049516](https://doi.org/10.1518/001872095779049516).
- Sharma, R. and Yetton, P. (2003), "The contingent effects of management support and task interdependence on successful information systems implementation", *MIS Quarterly*, Vol. 27 No. 4, pp. 533-556, doi: [10.2307/30036548](https://doi.org/10.2307/30036548).
- Shin, D. (2021), "The effects of explainability and causability on perception, trust, and acceptance: implications for explainable AI", *International Journal of Human-Computer Studies*, Vol. 146, pp. 1-10, doi: [10.1016/j.ijhcs.2020.102551](https://doi.org/10.1016/j.ijhcs.2020.102551).
- Shin, D. (2024), "Misinformation and generative AI: how users construe their sense of diagnostic misinformation", in *Artificial Misinformation: Exploring Human-Algorithm Interaction Online*, Springer Nature Switzerland, pp. 227-258.
- Siegel-Jacobs, K. and Yates, J.F. (1996), "Effects of procedural and outcome accountability on judgment quality", *Organizational Behavior and Human Decision Processes*, Vol. 65, pp. 1-17, doi: [10.1006/obhd.1996.0001](https://doi.org/10.1006/obhd.1996.0001).
- Sigala, M., Ooi, K.B., Tan, G.W.H., Aw, E.C.X., Cham, T.H., Dwivedi, Y.K., Kunz, W.H., Letheren, K., Mishra, A., Russell-Bennett, R. and Wirtz, J. (2024), "ChatGPT and service: opportunities, challenges, and research directions", *Journal of Service Theory and Practice*, Vol. 34 No. 5, pp. 726-737, doi: [10.1108/jstp-11-2023-0292](https://doi.org/10.1108/jstp-11-2023-0292).
- Simonson, I. and Staw, B.M. (1992), "Deescalation strategies: a comparison of techniques for reducing commitment to losing courses of action", *Journal of Applied Psychology*, Vol. 77 No. 4, pp. 419-426, doi: [10.1037/0021-9010.77.4.419](https://doi.org/10.1037/0021-9010.77.4.419).
- Singh, I.L., Molloy, R. and Parasuraman, R. (1993), "Automation-induced 'complacency': development of the complacency-potential rating scale", *The International Journal of Aviation Psychology*, Vol. 3 No. 2, pp. 111-122, doi: [10.1207/s15327108ijap0302_2](https://doi.org/10.1207/s15327108ijap0302_2).
- Skitka, L.J., Mosier, K. and Burdick, M.D. (2000), "Accountability and automation bias", *International Journal of Human-Computer Studies*, Vol. 52 No. 4, pp. 701-717, doi: [10.1006/ijhc.1999.0349](https://doi.org/10.1006/ijhc.1999.0349).
- So, K.K.F., Kim, H., Liu, S.Q., Fang, X. and Wirtz, J. (2024), "Service robots: the dynamic effects of anthropomorphism and functional perceptions on consumers' responses", *European Journal of Marketing*, Vol. 58 No. 1, pp. 1-32, doi: [10.1108/ejm-03-2022-0176](https://doi.org/10.1108/ejm-03-2022-0176).
- Spatola, N. (2024), "The efficiency-accountability tradeoff in AI integration: effects on human performance and over-reliance", *Computers in Human Behavior: Artificial Humans*, Vol. 2 No. 2, pp. 1-10, doi: [10.1016/j.chbah.2024.100099](https://doi.org/10.1016/j.chbah.2024.100099).
- Srinivasan, R. and Sarial-Abi, G. (2021), "When algorithms fail: consumers' responses to brand harm crises caused by algorithm errors", *Journal of Marketing*, Vol. 85 No. 5, pp. 74-91, doi: [10.1177/0022242921997082](https://doi.org/10.1177/0022242921997082).
- Sweller, J. (1988), "Cognitive load during problem solving: effects on learning", *Cognitive Science*, Vol. 12 No. 2, pp. 257-285, doi: [10.1207/s15516709cog1202_4](https://doi.org/10.1207/s15516709cog1202_4).
- Tetlock, P.E. (1992), "The impact of accountability on judgment and choice: toward a social contingency model", *Advances in Experimental Social Psychology*, Vol. 25, pp. 331-376.
- Tully, S., Longoni, C. and Appel, G. (2025), "Lower artificial intelligence literacy predicts greater AI receptivity", *Journal of Marketing*, Vol. 89 No. 5, pp. 1-20, doi: [10.1177/00222429251314491](https://doi.org/10.1177/00222429251314491).
- van Riel, A.C., Tabatabaei, F., Yang, X., Maslowska, E., Palanichamy, V., Clark, D. and Luongo, M. (2025), "A new competitive edge: crafting a service climate that facilitates optimal human-AI collaboration", *Journal of Service Management*, Vol. 36 No. 1, pp. 27-49, doi: [10.1108/josm-05-2024-0220](https://doi.org/10.1108/josm-05-2024-0220).
- Wageman, R. (1995), "Interdependence and group effectiveness", *Administrative Science Quarterly*, Vol. 40 No. 1, pp. 145-180, doi: [10.2307/2393703](https://doi.org/10.2307/2393703).
- Wirtz, J. and Stock-Homburg, R. (2025), "Generative AI meets service robots", *Journal of Service Research*, Vol. 28 No. 4, pp. 527-543, doi: [10.1177/10946705251340487](https://doi.org/10.1177/10946705251340487).

- Wirtz, J., Kunz, W.H., Hartley, N. and Tarbit, J. (2023), "Corporate digital responsibility in service firms and their ecosystems", *Journal of Service Research*, Vol. 26 No. 2, pp. 173-190, doi: [10.1177/10946705221130467](https://doi.org/10.1177/10946705221130467).
- Wood, R.E. (1986), "Task complexity: definition of the construct", *Organizational Behavior and Human Decision Processes*, Vol. 37 No. 1, pp. 60-82, doi: [10.1016/0749-5978\(86\)90044-0](https://doi.org/10.1016/0749-5978(86)90044-0).
- Xu, Y., Shieh, C.H., van Esch, P. and Ling, I.L. (2020), "AI customer service: task complexity, problem-solving ability, and usage intention", *Australasian Marketing Journal*, Vol. 28 No. 4, pp. 189-199, doi: [10.1016/j.ausmj.2020.03.005](https://doi.org/10.1016/j.ausmj.2020.03.005).
- Yalcin, G., Lim, S., Puntoni, S. and van Osselaer, S.M. (2022), "Thumbs up or down: consumer reactions to decisions by algorithms versus humans", *Journal of Marketing Research*, Vol. 59 No. 4, pp. 696-717, doi: [10.1177/00222437211070016](https://doi.org/10.1177/00222437211070016).
- Yang, S., Guo, Y., Chen, Y., Li, Y., Zhou, S. and Jiang, H. (2025), "Understanding service performance of frontline employees from employee-AI interaction perspective", *Journal of Service Management*, Vol. 36 No. 3, pp. 448-475, doi: [10.1108/josm-10-2023-0425](https://doi.org/10.1108/josm-10-2023-0425).

Corresponding author

Khanh Bao Quang Le can be contacted at: khanh.le@aut.ac.nz