informs.
https://pubsonline.informs.org/journal/orsc

ORGANIZATION SCIENCE
Articles in Advance, pp. 1–20
ISSN 1047-7039 (print), ISSN 1526-5455 (online)

More to Lose: The Adverse Effect of High Performance Ranking on Employees' Preimplementation Attitudes Toward the Integration of Powerful AI Aids

Ilanit SimanTov-Nachlieli[a]

[a] Coller School of Management, Tel Aviv University, Tel Aviv 6997801, Israel
Contact: ilanitsn@tauex.tau.ac.il, https://orcid.org/0000-0002-2713-7686 (IS-N)

Received: March 23, 2023
Revised: September 26, 2023; October 30, 2023; April 23, 2024; July 14, 2024
Accepted: September 18, 2024
Published Online in Articles in Advance: October 18, 2024

https://doi.org/10.1287/orsc.2023.17515

Abstract. Despite the growing availability of algorithm-augmented work, algorithm aversion is prevalent among employees, hindering successful implementations of powerful artificial intelligence (AI) aids. Applying a social comparison perspective, this article examines the adverse effect of employees' high performance ranking on their preimplementation attitudes toward the integration of powerful AI aids within their area of advantage. Five studies, using a weight estimation simulation (Studies 1–3), recall of actual job tasks (Study 4), and a workplace scenario (Study 5), provided consistent causal evidence for this effect by manipulating performance ranking (performance advantage compared with peers versus no advantage). Studies 3–4 revealed that this effect was driven in part by employees' perceived potential loss of standing compared with peers, a novel social-based mechanism complementing the extant explanation operating via one's confidence in own (versus AI) ability. Stronger causal evidence for this mechanism was provided in Study 5 using a "moderation-of-process" design. It showed that the adverse effect of high performance ranking on preimplementation AI attitudes was reversed when bolstering the stability of future performance rankings (presumably counteracting one's concern with potential loss of standing). Finally, pointing to the power of symbolic threats, this adverse effect was evident both in the absence of financial incentives for high performance (Study 1) and in various incentive-based settings (Studies 2–3). Implications for understanding and managing high performers' aversion toward the integration of powerful algorithmic aids are discussed.

Funding: This work was supported by the Coller Foundation.
Supplemental Material: The supplemental material is available at https://doi.org/10.1287/orsc.2023.17515.

Keywords: high performers • algorithm aversion • AI aids • social comparison • stability of ranking

Introduction

It is increasingly common for employees working in diverse domains to use algorithmic aids to augment their work. Recent developments in artificial intelligence (AI)[1] technology have increased the scope of work tasks and decisions that can be algorithmically augmented (Von Krogh 2018, Shrestha et al. 2021). These AI aids augment human work in two distinct ways. First is "human in the loop" (Lebovitz et al. 2022, p. 127), whereby humans and AI aids collaborate to accomplish a task or a decision (e.g., with AI providing data to support decision making). For example, AI aids can be used to analyse medical imaging data, offer diagnostics, and recommend treatments, as well as estimate financial and security risks. Second is automating some of the work tasks—simple, standardized, and repetitive tasks (Huang and Rust 2018)—but not replacing human labour entirely (when able to take over all of a job's tasks). For example, AI aids can match candidates' qualifications and job specifications to speed up hiring or generate code snippets to speed up development processes.

Despite the ongoing debate about whether AI may replace humans or merely augment their work (Lichtenthaler 2018), here, the focus is on AI aids aimed at augmenting human work by supporting employees in concrete decisions and tasks (Jarrahi 2018, Shrestha et al. 2021). Often, people reach better work outcomes if they use such aids. Indeed, meta-analyses have found that, mechanical, algorithm-based predictions generally outperform those of human experts (Grove et al. 2000, Kuncel et al. 2013). Given their utility in augmenting employees' work, and ultimately organizational performance (Jarrahi 2018, Metcalf et al. 2019, Shrestha et al. 2021), we are witnessing growing management and organizational initiatives to integrate such aids into various work domains.

Yet, recent research demonstrates persistent algorithm aversion, defined as "the reluctance of human decision makers to use superior but imperfect algorithms" (Burton et al. 2020, p. 220; see also Dietvorst et al. 2015 and Mahmud et al. 2022 for a review). This raises questions about whether, in practice, joint human-algorithm collaboration is at all feasible (Burton et al. 2020). Indeed, even within the clinical diagnostic domain in which AI advancements are rapid and the benefits are clear (Hosny et al. 2018), employees often hold negative preimplementation attitudes (e.g., Hanemaayer 2021, Jussupow et al. 2022), attitudes formed even before they gain direct experience with such algorithmic aids. These early attitudes have the power to shape experiences and behaviors further along the implementation process, thus posing barriers to successful implementation and even leading to failures (Lichtenthaler 2020, Prakash and Das 2021; see Herold et al. 1995 for the role of preimplementation attitudes in technology implementation success). Thus, we need to better understand employees' preimplementation AI attitudes, and particularly explore factors that adversely affect them. A better understanding of the roots of such early attitudes is vital as it may enable us to actively address and influence these attitudes in practice.

The literature typically groups factors influencing algorithm aversion into four broad themes: individual factors such as expertise, algorithm factors such as accuracy, task factors such as complexity, and higher-level factors such as organizational culture (see Mahmud et al. 2022 for a review). However, there is currently a theoretical gap in our understanding of how employees' social context may adversely affect their attitudes toward such aids. Specifically, what if powerful AI aids are aligned within one's unique advantage compared with peers (i.e., able to augment employees' work in domains or tasks where one currently enjoys a performance advantage over peers)? Would they support the integration of such algorithmic aids that have the power to level the playing field?

Notably, there is a significant difference between AI aids and older technologies such as various information systems or even predictive algorithms using traditional statistical methods. Although the latter may also support employees in decision making, they often require employees' own skills to interpret and apply the information themselves to reach sound decisions. In contrast, AI aids can recommend the decision itself or complete a work task on behalf of employees based on complex algorithmic analyses and interpretation. Thus, as elaborated below, such powerful AI aids, accessible to all employees, have or are perceived to have the power to level the playing field and take existing skill differentials out of the equation, a difference central to the current theorizing on the novel social-based mechanism underlying algorithm aversion. Indeed, with respect to generative AI, particularly large language models like GPT—recognized for their unexpectedly high capabilities in analytical, creative, and writing tasks and for achieving top scores in professional examinations (Dell'Acqua et al. 2023)—prior studies have found that using such AIs not only positively impacts knowledge workers' performance but also disproportionately benefits workers who struggle with tasks (e.g., low performers) compared with those who excel (e.g., high performers), leading to equalizing effects (e.g., Brynjolfsson et al. 2023, Dell'Acqua et al. 2023, Noy and Zhang 2023).

Applying a social comparison perspective, the current research examined whether employees' high performance ranking—performance *advantage compared with peers (versus no advantage)*—may have an adverse impact on their preimplementation attitudes toward the integration of powerful AI aids within their area of advantage (hereafter, *employee preimplementation AI attitudes*). Given that social comparisons are "embedded deeply into the fabric of organizational life" (Greenberg et al. 2007, p. 23), taking such a social-based perspective may explain further variance in employees' AI attitudes left unexplained by prior research. Moreover, complementing extant theorizing (Burton et al. 2020) suggesting that this effect may operate via one's confidence in own (versus AI) ability, the current research examined an additional, novel social-based mechanism: high performers' concern with potential loss of standing compared with peers.

This study has several key contributions. First, extending research on factors contributing to algorithm aversion (Mahmud et al. 2022) and recent theorizing on the role of perceived risks (or losses) in human-AI aid relationships (Solberg et al. 2022), it addresses the gap in our current understanding of how employee social context may adversely affect AI attitudes. Specifically, complementing the current focus on one's confidence in own (versus AI) ability, this study offers a novel mechanism to explain why employees' high performance ranking may adversely impact their preimplementation AI attitudes. As such, it also extends our understanding of why experts (compared with novices) show greater algorithm aversion (Arkes et al. 1986, Logg et al. 2019, Allen and Choudhury 2022). Second, although trust in AI's ability is typically related to positive AI attitudes (Glikson and Woolley 2020), here I theorize, and find, that even when clearly able to improve their outcomes, employees' attitudes toward powerful AI aids can be negatively affected by yet unexplored social factors. Practically, because high performers' preimplementation AI attitudes set the stage for the success or failure of AI implementation, the current findings provide managerial insights into how to better incorporate powerful algorithmic aids in organizations.

Theory and Hypotheses

Consistent with the notion of algorithmic technologies as representing the "new contested terrain of control"

(Kellogg et al. 2020, p. 366), managers continuously innovate to maximize the value of labour whereas workers resist. In the following, I review the crucial role of high-performing employees in the successful implementation of new management initiatives. Next, I take a new social comparison perspective to explain the adverse hypothesized impact of high performance ranking on employees' preimplementation AI attitudes.

The Role of High-Performing Employees in the Diffusion of New Technological Innovations. High performers are employees who outperform their peers (Call et al. 2015, Campbell et al. 2017, Hendricks et al. 2023). Research on attracting and retaining these employees, who bring disproportionally substantial value to their organizations, dominates discussions among management scholars and business leaders (Sutton 2007, Campbell et al. 2017). High performers and experts (individuals with high absolute performance within a specific domain (Ericsson and Towne 2010)) can lead by example and affect the diffusion of new technologies. Specifically, end users are influenced in their adoption attitudes by advice from such respected peers (Rogers 1982, Leonard-Barton 1985), who function as opinion leaders to "influence other individuals' attitudes or overt behavior informally" (Rogers 1982, p. 27). Leonard-Barton (1985, p. 941) similarly noted that in the diffusion of controversial technological innovations, these employees "influence the rate and extent of acceptance by serving as negative or positive opinion leaders."

Thus, from a management and strategic perspective, any implementation process must involve these employees. Without their volitional "buy-in," implementation will fail before it has begun. As noted by Michlitsch (2000, p. 28), "Strategy implementation is best accomplished through high-performing people." Buy-in from such employees is also critical in that high performers positively impact peer performance through useful knowledge spillovers and role modelling (Hendricks et al. 2023). Thus, when soliciting input during organizational change initiatives, implementers often target end users, those with subject matter expertise, or "high performers" (Lewis and Russ 2012, p. 277)—employees the management wants to have fully on board—seek their feedback, and look for early signs of their resentment (Mayfield 2014).

Employees' AI Attitudes and Algorithm Aversion. Research investigating factors affecting employees' AI attitudes has focused on human trust in AI (Glikson and Woolley 2020). Factors affecting this trust include the human actor's characteristics (e.g., age, gender, personality traits, and expertise/ability); the AI's characteristics (e.g., accuracy/ability (Kaplan et al. 2021) and perceived ease of use and usefulness (Sánchez-Prieto et al. 2020)); and situational (e.g., task nature, task difficulty, and employee workload) and higher-level or organizational (e.g., culture, communication, and training) factors (see Mahmud et al. 2022 for a review of sources of algorithm aversion, grouped into similar four broad themes).

In their review of algorithm aversion in augmented decision making (the current focus), Burton et al. (2020, p. 223) emphasize the role of employees' own expertise and expectations "as to what an algorithm can do." They predict that employees with high domain expertise would feel confident *without* the algorithm and deem the effort needed to consult it unnecessary. Goddard et al. (2012) similarly propose that reliance on algorithms is essentially a trade-off between self-confidence in own performance and trust in algorithm performance. Supporting this proposition, Logg et al. (2019) found that experts use algorithmic advice less than do lay participants, and reasoned that experts tend to rely more on their own (versus algorithmic) judgment or ability. This proposition is consistent with findings indicating that algorithmic aversion is prevalent among experts and findings linking experts' confidence in their abilities to underutilization of seemingly unnecessary algorithmic aids (Arkes et al. 1986, Allen and Choudhury 2022).

Although in all these studies, the focus is on one's domain expertise, as naturally observed (measured), rather than on one's (manipulated) performance advantage compared with peers,[2] the above theorizing and findings provide initial support to propose a causal adverse effect of an employee's high performance ranking on preimplementation AI attitudes through one's greater confidence in own (versus AI) ability. Complementing this prior theorized confidence/ability-based mechanism, I propose next that this adverse effect is also driven by a novel *social-based mechanism*—one's greater concern with a potential loss of standing compared with peers.

A New Social Comparison Perspective on Preimplementation AI Attitudes. A recent conceptual model highlights the important role of perceived losses (or risks of undesirable threatening outcomes) that could result in negative AI attitudes and underutilization of the AI even if the person experiences high trust in its ability (Solberg et al. 2022). Given the current focus on powerful AI aids aimed at augmenting employees' work, employees' fear of being replaced by such aids is beyond the current scope. However, as further explained below, even when the employees' job is secure and employees can trust the AI to perform well (and augment their own performance), it does not necessarily mean they would support its integration. The fact that aversion is still prevalent suggests that employees' AI attitudes depend not only on the technology's objective benefits, but also on how it is subjectively perceived by employees.

Indeed, AI can pose symbolic, psychological risks to human identity (e.g., to the employees' professional prestige and work autonomy (Hanemaayer 2021, Jussupow et al. 2022)). Drawing on social comparison literature (Smith 2000, Buunk and Gibbons 2007), I focus on one's concern with potential loss of standing relative to peers to explain why high performance ranking may adversely impact preimplementation AI attitudes.

Social comparison processes and outcomes have been studied in a wide variety of areas, including organizational settings (e.g., Brown et al. 2007, Goodman and Haisley 2007). As mentioned, complementing the extant focus on one's confidence in own (versus AI) ability, I reason next that high performance ranking may also heighten an employee's perceived potential loss of standing—one's belief that integrating powerful AI aids within their area of advantage could worsen their comparative standing within a group or position in the organizational hierarchy. Put differently, although powerful AI aids may have (or be perceived as having) the potential to augment human work and improve the utility for the organization as a whole, their perceived personal utility for high performers is questionable, as those employees are already performing well, whereas their perceived potential risk to the employees' high relative standing may become salient. Prior research indeed suggests that high performers may prioritize their personal over organizational goals—for example, they may control key resources to reduce others' learning opportunities (see Asgari et al. 2021 for a review).

Employee performance is a critical dimension for social comparisons at work, in particular among peers well informed about each other's performance (Molleman et al. 2007), who attach importance to such information (Barr and Conlon 1994). Indeed, we frequently compare ourselves to colleagues whose current or even future good performance may directly threaten us (Johnson 2012). Prior research has highlighted that various threats (and consequent responses) resulting from upward or downward comparisons depend on whether individuals contrast or assimilate themselves in comparison with others (Smith 2000, Mussweiler et al. 2004; see also Matta and Van Dyne 2020). Pertinent to the current focus, when an employee outperforms others, downward assimilative comparisons in which one expects a potential high future performance similarity in comparison with others may be self-threatening. This is because they heighten one's likelihood of losing their high standing and the unpleasant emotions associated with it.[3] Indeed, given that one's basic motivation is to maintain (high) standing on any self-relevant dimension, motivated to prevent future threats, "people tend *not* to recommend individuals who surpass them on the relevant dimension on which they have high standing" (Garcia et al. 2010, p. 97). Similarly, an anticipated status threat by faster-rising coworkers induced employees to undermine them (Reh et al. 2018).

Accordingly, I posit that a downward assimilation-driven social comparison (Smith 2000) is also involved in the adverse impact of high performance ranking on preimplementation AI attitudes. When peers can benefit from powerful AI aids, employees with high performance rankings may believe that their performance advantage over peers (or their high standing) is becoming unstable, inducing concern with potential loss of standing.

An individual's social rank in the organizational hierarchy (or the influence and attention one holds within a group (Mitchell et al. 2020)) may be rooted in characteristics reflecting task competence, which may include prior high task performance (Magee and Galinsky 2008, Kehoe et al. 2018). Moreover, individuals' fundamental desire for high social rank (and motivation to protect it) is partly driven by the intrapersonal, symbolic benefits it affords, such as self-esteem and autonomy or interpersonal benefits such as peer recognition and influence (Mitchell et al. 2020). As such, perceived potential loss of standing presents a symbolic risk that may harm the formation of positive preimplementation AI attitudes among high performers. Thus, I hypothesize:

Hypothesis 1. *Employees' high performance ranking has an adverse impact on their preimplementation attitudes toward the integration of powerful AI aids within their area of advantage.*

Hypothesis 2. *Complementing the extant theorized mechanism operating via one's confidence in own (versus AI) ability, the adverse effect of high performance ranking on preimplementation AI attitudes is also mediated by one's perceived potential loss of standing compared with peers.*

Perceived potential loss of standing, besides posing a symbolic risk, can also incur realistic downstream consequences involving financial costs. I propose next that certain organizational incentive structures—those characterized by *relative* pay determination criteria (Lazear and Oyer 2013)—may exacerbate the link between perceived potential loss of standing and AI attitudes.

The Second-Stage Moderating Role of Relative Pay Determination Criteria. Because employees tend to perceive organizational resources as scarce (Huijsmans et al. 2019), they may be motivated to maintain their high relative standing (Mittone and Savadori 2009, Chernyak-Hai and Davidai 2022), not only for its intrapersonal (e.g., self-esteem and autonomy) benefits, but also to improve their access to concrete organizational resources (Mitchell et al. 2020).

Indeed, pay is largely contingent upon employee performance and pay-for-performance (PFP) is pervasive (Gerhart et al. 2009). Thus, the way performance is translated into actual pay (i.e., pay determination criteria) is important to consider. Whereas some organizations base their PFP on performance *relative* to peers, others (concerned with competition) ground their PFP on *absolute* standards. Indeed, relative (versus absolute) pay criteria induce competition (Gerhart et al. 2009, Belogolovsky and Bamberger 2014). Moreover, competition increases as one expects to be more closely ranked in comparison with others (Garcia et al. 2013). Finally, Reh et al. (2018) found that employees reacted to faster-rising coworkers with socially undermining behavior when the organizational climate was more competitive. Thus, I hypothesize:

Hypothesis 3. *The negative link between perceived potential loss of standing and preimplementation AI attitudes is amplified when pay determination criteria are relative (versus absolute), resulting in a second-stage moderated mediation model.*

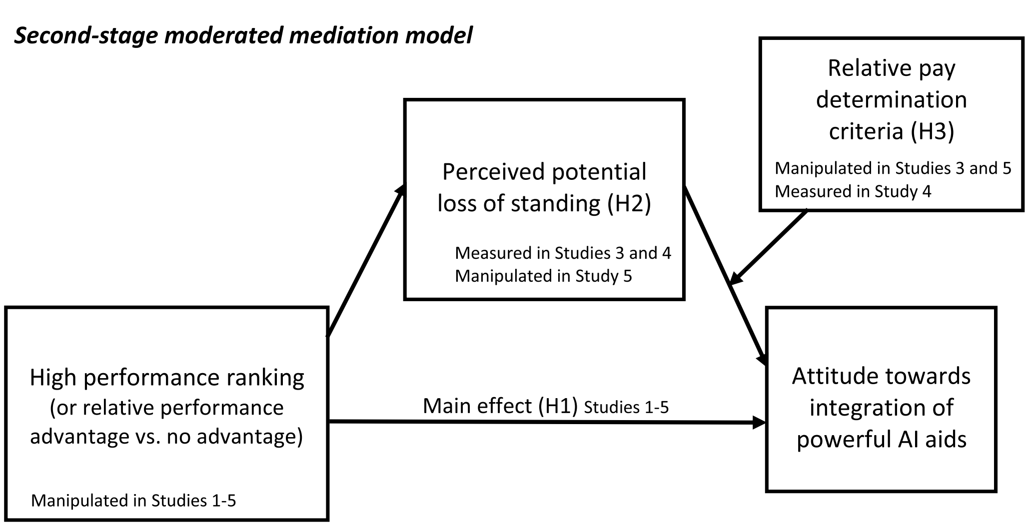
Overview of Studies and Methods

The above hypotheses were tested in five studies using a variety of experimental settings and designs (see Figure 1). Study 1, conducted among undergraduate students with a weight estimation task simulation manipulating one's relative performance, tested the main effect (Hypothesis 1). It also provided an initial qualitative indication for the mediating role of perceived potential loss of standing. Study 2, conducted among employees using the same simulation but with an incentive-compatible setting and manipulating both absolute and relative performance, provided further indication that social comparisons were involved by showing that the effect of relative performance was evident across different levels of absolute performance. Studies 3–5 tested the novel, complementary social-based mechanism using different approaches. Specifically, Studies 3 and 4, conducted among employees, used the weight estimation task (Study 3) and employees' recall of actual job tasks (Study 4) to test this mechanism (controlling for extant theorized ability-based mechanism) using measurement-of-mediation designs (Spencer et al. 2005). They also tested the full second-stage moderated mediation model (i.e., Hypotheses 1–3). Study 5 tested this mechanism in a causal manner using a moderation-of-process design (Spencer et al. 2005) by manipulating the stability of future rankings in a workplace scenario. All studies were approved by an ethics committee.

Given the focus on *powerful* AI aids, their key characteristic—AI accuracy (or its perceived ability-based trustworthiness, a key antecedent of trust in AI (Kaplan et al. 2021))—was held constant and high in all studies and across all conditions, describing an easy-to-use AI aid with exceptional proven capabilities in the focal task domain. The AI's utility to everybody was further made explicit in Studies 2 and 5. As such, participants' supportive preimplementation attitudes were generally high (significantly above the 4 midpoint; $p < 0.001$) across all conditions and studies.

The employees in Studies 2–5 (U.S. employees) were recruited via the Prolific platform (See the supplemental material (SM) for more details on Prolific and its advantages compared to other platforms). Invitations were sent to employees with a minimum of secondary education, working at least half-time, and having peers. I also restricted the sample to "high-reputation" workers (≥95% approval rate). Studies 2, 4, and 5 were preregistered (see https://aspredicted.org/QQJ_Z1J, https://aspredicted.org/GYV_PGJ, and https://aspredicted.org/THZ_9JK). Study 3 adhered to the same hypotheses,

Figure 1. Theorized Model with Focal Social-Based Mediator



analyses, and measures preregistered in Study 4. All studies' materials, data, and syntax are available at https://osf.io/rjsu6/?view_only=49a4a7e3699e4e2fb98365cca96913d4. Samples' descriptions, sizes, and exclusions (e.g., based on participants failing instructional manipulation checks (Oppenheimer et al. 2009)) were determined a priori and fully reported in the SM. The number of participants reported in the text is after such exclusions. Finally, as preregistered, data were analysed using IBM SPSS Statistics and the Hayes (2017) PROCESS macro (version 4.0) for mediation and moderated mediation analyses (bootstrap sample size = 5,000).

Study 1

Conducted among students participating for credit points, Study 1 aimed to establish the adverse causal effect of high performance ranking on preimplementation AI attitude (Hypothesis 1).

Method

Participants were 209 business administration undergraduates (96 women; M_{age} = 22.54, SD = 2.35). They were informed that the study included a simulation consisting of two sessions, in which they would be asked to assess individuals' weight based on photographs (see Moore and Klein 2008, Logg et al. 2019).[4] They were further told that they would receive feedback about their relative performance. Given the absence of financial incentives,[5] prior to the simulation, participants indicated their a priori motivation to succeed in the task using three items adapted from Grant and Dweck (2003), for example, "It is important for me to do well on the task" (α = 0.95).

In the first session, the participants assessed the weight of 10 individuals. Once submitting their estimations, they were told to stand by "while we evaluate participants' weight estimations." After waiting for 15 seconds (constant across conditions), participants received randomly assigned (bogus) performance feedback. Those assigned to the high (versus average, no advantage over peers) relative performance condition were told that "In this first session, your performance level relative to the performance of other participants is: *high* (versus *average*)." A manipulation check was then conducted—participants were asked: "According to the performance feedback you've just received, between 0 (worst performers) and 10 (best performers), where do you stand?" Participants also indicated their positive and negative state affect (α = 0.71 and α = 0.80, respectively) following this feedback, using 10 items on a 5-point scale from the short-form Positive and Negative Affect Schedule (PANAS (Thompson 2007)), because incidental positive affect was associated with greater preference for the general status quo option (e.g., Yen and Chuang 2008).

Then, before supposedly moving on to the next session (assessing the weights of other individuals), participants were told that "An Artificial Intelligence system *with exceptional proven capabilities in image processing* may be integrated into the second session. If integrated, *everybody will be able to use its weights estimations easily* (if they choose to do so) to make their weight assessments in the second session." After verifying that participants read this information, they indicated their attitude toward the integration of this AI using two items (*Spearman-Brown* = 0.82) adapted from Kim and Kankanhalli (2009), randomized in order, on a seven-point scale (1 = *strongly disagree* to 7 = *strongly agree*): "I support the integration of this AI system in the next session," and "I oppose the integration of this AI system in the next session" (reverse coded).

As an additional and more behavioral construct, participants were also asked, "If it were up to you, would you like to integrate this AI system in the next session?," with only three response options available: "No," "Yes," and "Indifferent." Next, for exploratory purposes only, participants were given the option to indicate, using an open textbox, the reasons leading them to respond the way they had (97% responded). Finally, students provided their demographics as well as AI knowledge using three items adopted from Chiu et al. (2021), for example, "I do *not* feel very knowledgeable about artificial intelligence" (α = 0.64), to control for in a robust analysis.

Results and Discussion

Results indicated that students' a priori motivation to succeed in the estimation task was similarly high (significantly above the 4 midpoint, ps < 0.001) across conditions, M = 5.78, SD = 1.31, and M = 5.68, SD = 1.17, in the average and high relative performance conditions, respectively. Moreover, as fully elaborated in the SM, the manipulation of performance ranking was successful.

Main Analysis. A univariate analysis revealed a significant effect of the performance ranking manipulation on preimplementation AI attitudes, $F(1, 207) = 4.86$, $p = 0.03$. Specifically, supporting Hypothesis 1, students assigned to the high relative performance condition indicated *lower* supportive preimplementation AI attitudes (M = 4.92, SD = 1.72) than did those assigned to the average performance condition (M = 5.39, SD = 1.33), $B = -0.47$, $SE = 0.21$, $p = 0.03$, $\eta_p^2 = 0.02$. Adding to the robustness of this finding, the effect persisted while controlling for gender, age, a priori motivation to succeed in the task, positive and negative state affect, and knowledge of AI, $B = -0.48$, $SE = 0.23$, $p = 0.04$, $\eta_p^2 = 0.02$.

A similar effect was observed in students' responses when asked to choose whether they would like to integrate this AI in the next session ("No," "Yes," "Indifferent"). A multinomial logistic regression was statistically significant, $\chi^2 = 9.01$, $df = 2$, $p = 0.01$, Nagelkerke $R^2 = 0.05$. Although participants were more likely to choose "Yes" over "No" in both conditions (45.8% versus 24.3% in the high relative performance condition, 63.7% versus 10.8% in the average condition), participants in the high relative (versus average) performance condition were about three times more likely to choose "No" over "Yes," $OR = 3.13$, $p < 0.01$. There was no statistically significant difference in the percentage of those who chose "Indifferent" across the two conditions (25.5% and 29.9% in the average and high conditions, respectively), $p = 0.13$.

Coding Reasons for Responses. Two research assistants (RAs), blind to experimental conditions, independently coded the participants' open responses (RAs' codings are provided in the OSF link). Coding categories were predetermined: 1A for *general trust in AI's ability to help everybody* (e.g., "technology is a good thing to help us make decisions") or 1B for *general aversion* (e.g., "A human eye is always better"); 2A (or 2B) for high (low) *confidence in AI's ability to help the participant* (e.g., "Because I think it will help me guess the weight better"); 3A (or 3B) for high (or low) *confidence in own ability* (e.g., "Because I'm really good without the system"); 4A to -C for reasons related to *social comparison*—specifically, 4A for a potential improvement in standing compared with peers (e.g., "Because if I'm average I would like something that might help me raise my grade"), 4B for a potential loss of standing (e.g., "My performance is good now, and I don't need others to compare their results to me"; "I'm already among the best and I don't want the bad ones to be good"), and 4C for stability in standing (e.g., "Doesn't matter because even if they introduce the system, everyone will improve on the same scale"); and 5A and -B for other reasons—specifically, 5A for study-related concerns (e.g., "I assume that the purpose of the research is to guess on your own and not to be helped by anything," "It is more real to guess alone"), and 5B for individual tendencies such as curiosity and challenge (e.g., "Just out of curiosity," "I like the challenge of the question").

Here, I present key findings related to this coding. Agreement between RAs was 95%. They were unable to code a similar percentage of responses across the two conditions (14% and 15% in the average and high conditions, respectively). Most interestingly, focusing on "Yes" and "No" decisions and including only key coded categories—those with at least two fitting responses per condition (and decision)—I found that categories related to "Yes" were similar across the two conditions: specifically, 66% and 65% were coded 1A (general trust in AI's ability to help everybody) in the average and high conditions, respectively, and 20% and 22% were coded 2A (confidence in AI's ability to help the participant) in these two conditions, respectively. Coded categories related to "No" were quite different: whereas in the average condition, 36% were coded as 1B (general AI aversion) and 18% were coded as 4C (stability in standing), in the high relative condition, 23% were coded 3A (high confidence in own ability) and 27% were coded 4B (the social-based mechanism operating via concern with loss of standing) (see the SM, Table S1, for more detailed results). As such, this coding provided an initial qualitative support for the complementary social-based mechanism.

In sum, Study 1 revealed the hypothesized adverse impact of high performance ranking on preimplementation AI attitudes by combining the strengths of both quantitative and qualitative data to answer the what and why questions (respectively), without priming participants with questions related to any theorized mediators. Pointing to the power of symbolic threats, it also established the causality of this adverse effect, net of any financial incentive for high performance.

Study 2

Study 2 aimed to (a) replicate the adverse effect of high performance ranking on preimplementation AI attitudes among actual employees (rather than students), (b) provide an additional (quantitative) indication that social comparisons are involved in this effect, and (c) make it even more explicit that everybody (even those with high absolute performance) can benefit from consulting the AI aid. The study used the same weight estimation simulation as in Study 1, but, as preregistered, participants were randomly assigned to receive performance feedback combining both absolute and relative ratings. Specifically, the study had a 2 (*Absolute* performance: high level, held constant at 80% success rate; moderate level, held constant at 60% success rate) × 2 (*Relative* performance: high compared with other participants, average compared with other participants) between-subjects factorial design. Moreover, AI's (higher) performance level was held constant (at 90.5% success rate) across all four experimental conditions.

This study aimed to show that both absolute performance (consistent with previous research on experts versus novices and the extant theorized ability-based mechanism) and relative performance (consistent with current social comparison approach) matter. Hence, as preregistered, I expected to reveal a main effect of relative performance in addition to the prior theorized main effect of absolute performance. If concerns with high standing are indeed at play, then relative performance should matter regardless of absolute performance (high or moderate).

Method

Participants were 810 U.S. employees (417 women, $M_{age} = 39.14$, $SD = 12.00$). They were compensated with $1 for participation in a 5-minute simulation (see the SM for sample description, predetermined size, and exclusions). Study 1's procedure was generally followed with a few changes: after submitting their estimations (and before receiving their feedback), participants were asked to indicate how difficult it was to estimate people's weights from pictures, to verify that such an estimation task was not simple as well as to control for perceived difficulty in robust analysis.

Next, as mentioned, participants were randomly assigned to receive performance feedback combining both absolute and relative performance ratings. Specifically, those randomly assigned to the high (moderate) *absolute* performance were told that "your absolute (that is, objective) performance level is: *high* (*moderate*). Specifically, out of 10 guesses you made, *8* (*6*) were correct. Thus, your performance level stands on: *80%*, (*60%*)." As in Study 1, those randomly assigned to the high (average) relative performance were told that "Moreover, your performance level relative to the performance of other participants connected online is: *high* (*average*)." As such, participants were randomly assigned into four conditions—high absolute performance with high ranking, high absolute performance with average ranking, moderate absolute performance with high ranking, and moderate absolute performance with average ranking.

The vividness of the setting and its relevance to future work were enhanced by telling the participants that "As you've just done, employees often make complex estimations/decisions on their jobs such as estimating security, financial, and medical risks. It is increasingly common for knowledge workers in organizations to use algorithmic tools to augment their work-related tasks and decisions." This study also used an incentive-compatible setting—specifically, participants were encouraged to respond seriously and told that "further bonus may be allocated based on the total output (accuracy in weight estimations) of all the participants *in the second session*."[6] Similar to Study 1, they were informed about the potential integration of an AI aid specialized in image processing into the next session. Here, to make it explicit that this AI may benefit *everybody*, participants further read: "The performance level of this AI typically stands on 90.5%."

Next, to make it explicit that preimplementation attitudes carry consequences, participants were further told that "You and the other participants will be asked for your recommendations concerning the possible integration of this AI in the next session. Important! If recommendations do not exceed a certain threshold among the participants, this AI will NOT be integrated in the next session, and all the participants will continue to make their weight estimations on their own (as done in the first session)." Participants' comprehension was verified using a short quiz.

Employees then indicated to the research directors their recommendation about integrating that AI aid into the second session, as well as their preimplementation AI attitudes (see below). Finally, the absolute and relative manipulations were checked using two items on a seven-point scale: "My *absolute* (objective) performance level was high" and "My *relative* performance *compared to that of other participants* was high" (1 = *strongly disagree* to 7 = *strongly agree*). To further verify that the performance combinations were perceived as feasible, participants indicated, "In your view, how likely it was for someone (not necessarily you) to achieve this performance?" (1 = *impossible* to 7 = *extremely likely*). Employees indicated their demographics, working hours, organizational tenure, occupational level, and AI knowledge (see Study 1; a = 0.78), as well as AI usage in their current workplace (1 = *never* to 5 = *always*).

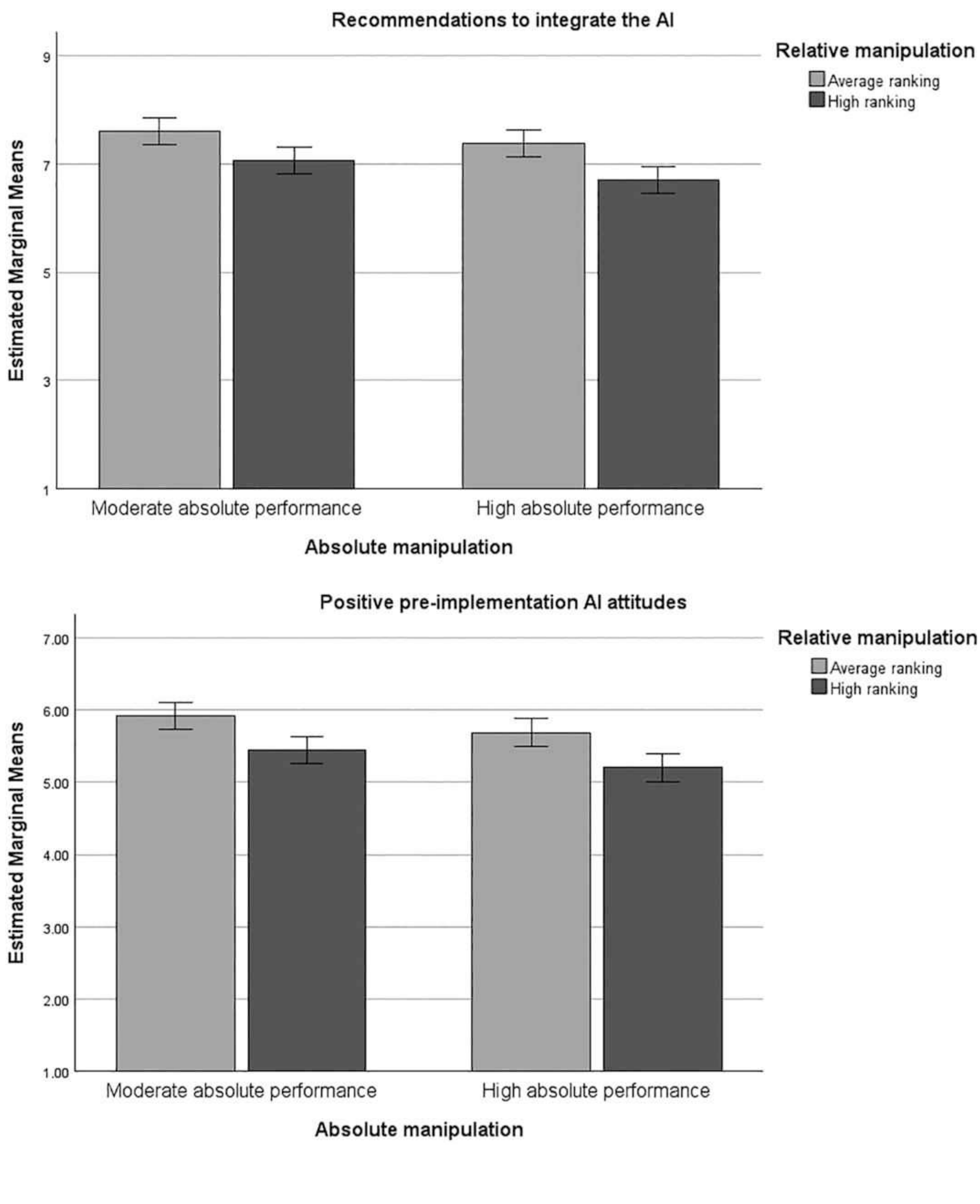
Dependent Variables (DVs). As preregistered, one item on a nine-point scale (1 = *Do not recommend at all* to 9 = *Very highly recommend*), adapted from prior research on organizational action (Heilman et al. 1997), was used: "Would you recommend the research directors/supervisors to incorporate this AI (for the use of all participants) in the next session?" Another three items on a seven-point scale (1 = *strongly disagree* to 7 = *strongly agree*) were adapted from prior studies to measure preimplementation AI attitudes. As in Study 1, one item (reverse worded) was adapted from Kim and Kankanhalli (2009): "I support the integration of this AI in the next session." Two additional items were adapted from Taylor and Todd (1995; see also Dwivedi et al. 2017): "I like the idea of integrating this AI in the next session" and "Integrating this AI in the next session is a good idea" (a = 0.95 for all three).

Results and Discussion

As fully elaborated in the SM, both the absolute and relative performance manipulations were successful. Moreover, perceived task difficulty was similarly high (significantly above the 4 midpoint, $ps < 0.001$) across conditions (M = 4.93, SD = 1.48; M = 5.02, SD = 1.47; M = 5.22, SD = 1.33; and M = 5.00, SD = 1.46, in the high absolute with high ranking, high absolute with average ranking, moderate absolute with high ranking, and moderate absolute with average ranking conditions, respectively). Perceived feasibility of the performance feedback was also high (significantly above the 4 midpoint, $ps < 0.001$) across these conditions (M = 4.89, SD = 1.30; M = 5.12, SD = 1.22; M = 4.81, SD = 1.07; and M = 4.97, SD = 1.01). No interaction effect emerged, $p = 0.62$ (thus, the performance rating combinations were perceived as similarly likely). All the effects persisted when controlling for demographics, work- and AI-related variables, perceived difficulty, and feasibility in robust analyses.[7]

Main Analysis. A multivariate two-way analysis was conducted to test the effects of absolute performance, relative performance, and their interaction on participants' recommendations and AI attitudes. As seen in Figure 2, this analysis revealed the expected two main effects of *absolute* and *relative* performance manipulations on both DVs. Specifically, consistent with the extant theorized ability-based mechanism, on average, across the two relative conditions, participants in the *high absolute* condition indicated lower recommendations (M = 7.04, SD = 1.90) as well as lower positive AI attitudes (M = 5.45, SD = 1.46) than did those in the *moderate absolute* condition (M = 7.34, SD = 1.59, and M = 5.68, SD = 1.13), $F(1, 806) = 6.09$, $p = 0.014$, $\eta_p^2 = 0.01$, and $F(1, 806) = 5.89$, $p = 0.015$, $\eta_p^2 = 0.01$, respectively. Consistent with the social comparison approach, on average, across the two absolute conditions, participants in the *high relative* condition indicated lower recommendations (M = 6.88, SD = 1.96) as well as lower positive AI attitudes (M = 5.33, SD = 1.53) than did those in the average condition (M = 7.49, SD = 1.47, and M = 5.81, SD = 1.19), $F(1, 806) = 24.94$, $p < 0.001$, $\eta_p^2 = 0.03$, and $F(1, 806) = 24.91$, $p < 0.001$, $\eta_p^2 = 0.03$, respectively. Results persisted when combining the two DVs. Finally, no interaction effects emerged, ps > 0.56 (i.e.,

Figure 2. Main Effects of Relative and Absolute Performance Manipulations on Recommendations to Integrate and Positive AI Attitudes (Study 2)



the adverse effect of high relative performance was evident both for those with high absolute performance and for those with moderate absolute performance; see the SM for detailed results).

In sum, although it was clear that everybody could benefit from AI because of its higher success rate, using an incentive-compatible setting with consequences for their recommendations showed that relative performance ratings were influential. Specifically, high (versus average) ranking adversely affected recommendations and AI attitudes across the two absolute conditions. Thus, the addition of information to the participants' advantage compared with peers affected their recommendations and attitudes, providing further quantitative indication that social comparisons are *also* involved. Notably, these effects were replicated in a similar study conducted among 400[8] UK employees, prescreened for using AI at work at least once a week (see the SM, "UK Replication of Study 2").

Study 3

Study 3 had two goals. First, it aimed to directly measure the novel complementary social-based mechanism (perceived potential loss of standing compared with peers) and test its mediating role when considering the extant mechanism operating via confidence in own (versus AI) ability (Hypothesis 2). Second, it aimed to replicate the adverse effect of high performance ranking on preimplementation AI attitudes using performance-based incentives (Gerhart et al. 2009)—manipulating relative (versus absolute) pay determination criteria for one's performance in the second session. As such, it also tested Hypothesis 3 (the second-stage moderated mediation effect) using a 2 (performance ranking: high, average) × 2 (pay determination criteria: relative, absolute) between-subjects factorial design.

Method

Participants were 395 U.S. employees (190 women, M_{age} = 37.36, SD = 10.69) compensated with $1 for participation in a 5-minute simulation (see the SM for sample description). This study followed Study 1's procedure with one exception. After receiving their (randomly assigned) relative performance feedback in the first session and assessing the manipulation check (see Study 1), participants were told that they may receive a bonus in the next session based on their relative performance compared with peers or on an absolute (predetermined) basis according to their own level of performance only (in the relative versus absolute conditions, respectively; see Belogolovsky and Bamberger 2014). Three items adapted from Belogolovsky and Bamberger (2014) and rated on a seven-point scale were used as a manipulation check (e.g., "Bonus will be influenced by my performance relative to the performance of others in the second session"; a = 0.82). Then, participants indicated their motivation to succeed in the next session using the items used in Study 1 (a = 0.92), but measured *after* the manipulations, to rule it out as an alternative explanation.

Next, as in previous studies, participants were informed about the potential integration of a powerful AI aid into the second session. Once verifying that participants read this information, various measures rated on a seven-point scale were presented in a randomized order to test the mediating roles of perceived potential loss of standing (novel social-based mediator) and confidence in own (versus AI) ability (prior theorized mediator). Finally, employees indicated their preimplementation AI attitudes using three items (a = 0.95; see Study 2). All measures were adapted from prior research. A confirmatory factor analysis (CFA) (conducted using Mplus, version 8.4) confirmed that all items loaded on their respective factors (ps < 0.001). Further analyses (chi-square difference tests comparing alternative models) confirmed the discriminant validity of these measures.[9]

Social-Based Mediator—Perceived Potential Loss of Standing. Two items (*Spearman-Brown* = 0.65) on a seven-point scale (1 = *strongly disagree* to 7 = *strongly agree*), based on Chernyak-Hai and Davidai (2022), assessed employees' perceptions of how their standing compared with peers would (negatively) change if the AI aid were integrated into the second session: "*worsen* my standing relative to others" and "*improve* my standing relative to others" (reverse coded).[10]

An additional item, "*not change* my standing relative to others," was measured so as not to prime participants in any specific direction as well as to show that in contrast to the negative expected link between perceived potential loss of standing and (positive) AI attitude, this item (perceived stability of future rankings) may not show this association. This item is *not* part of the above social-based mediator scale (it can be similarly low for participants perceiving a high potential loss as well as for those perceiving a high potential gain). The three items were randomized.

Prior Theorized Mediator—Confidence in Own (vs. AI's) Ability. Participants indicated how much confidence they had in own and AI's (future) weight estimates (1 = *none* to 7 = *a lot*; adapted from Dietvorst et al. 2015). These items were randomized. The difference between the two (confidence in own minus AI's ability) was then calculated.

Results and Discussion

Means and standard deviations for variables measured per experimental condition are provided in Table 1. As elaborated in the SM, both the performance ranking and pay determination criterion manipulations were successful. Notably, all the effects reported below persisted

Table 1. Means and Standard Deviations of Key Measured Variables, per Condition (Study 3)

	Absolute pay criteria				Relative pay criteria			
	Average performance (N = 101)		High performance (N = 101)		Average performance (N = 95)		High performance (N = 98)	
Variable	M	SD	M	SD	M	SD	M	SD
MC of performance ranking	5.28_a	0.72	7.97_b	1.12	5.38_a	0.86	8.01_b	1.58
MC of relative pay criteria	1.75_a	1.17	1.59_a	0.99	5.46_b	1.21	5.48_b	1.13
Motivation (after manipulations)	6.69_a	0.57	6.64_a	0.62	6.59_a	0.67	6.71_a	0.60
Confidence in self	4.40_a	1.12	5.32_b	1.07	4.44_a	1.41	5.38_b	1.08
Confidence in AI	5.37_a	1.20	5.21_a	1.34	5.28_a	1.10	5.25_a	1.26
Confidence in self *minus* in AI	-0.97_a	1.68	0.11_b	1.40	-0.84_a	1.60	0.12_b	1.65
AI usefulness to self	5.48_a	1.15	5.09_b	1.26	5.56_a	1.22	5.05_b	1.35
AI usefulness to others	5.12_a	1.22	5.26_a	1.20	5.42_a	1.21	5.48_a	1.30
AI usefulness to others *minus* to self	-0.36_a	0.93	0.17_b	0.89	-0.14_a	0.69	0.43_c	0.80
Perceived loss of standing	2.99_a	1.28	3.50_b	1.42	2.81_a	1.26	3.63_b	1.48
Perceived stability	3.96_a	1.84	3.94_a	1.58	3.66_a	1.70	3.74_a	1.75
(Positive) AI attitude	5.94_a	1.38	5.39_b	1.71	5.94_a	1.37	4.96_c	1.97

Notes. N = 395. MC refers to manipulation check. In each row, means with different subscripts differ significantly from each other ($p < 0.05$). The difference between "b" and "c" subscripts in the last row is marginal ($p = 0.06$).

when controlling for all demographics and work- and AI-related variables in further robust analyses.

Replication: The Adverse Effect of High Performance Ranking on Preimplementation AI Attitudes. Supporting Hypothesis 1, a univariate two-way analysis revealed the expected main effect of performance ranking on preimplementation AI attitudes. Specifically, on average, across the two pay determination criterion conditions, participants in the high performance ranking condition indicated a lower positive AI attitude (M = 5.18, SD = 1.85) than did those in the average performance condition (M = 5.94, SD = 1.37), $F(1, 386) = 21.56$, $p < 0.001$, $\eta_p^2 = 0.05$. The main effect of pay determination criteria and the independent variable (IV) × moderator interaction were not significant, ps = 0.19.

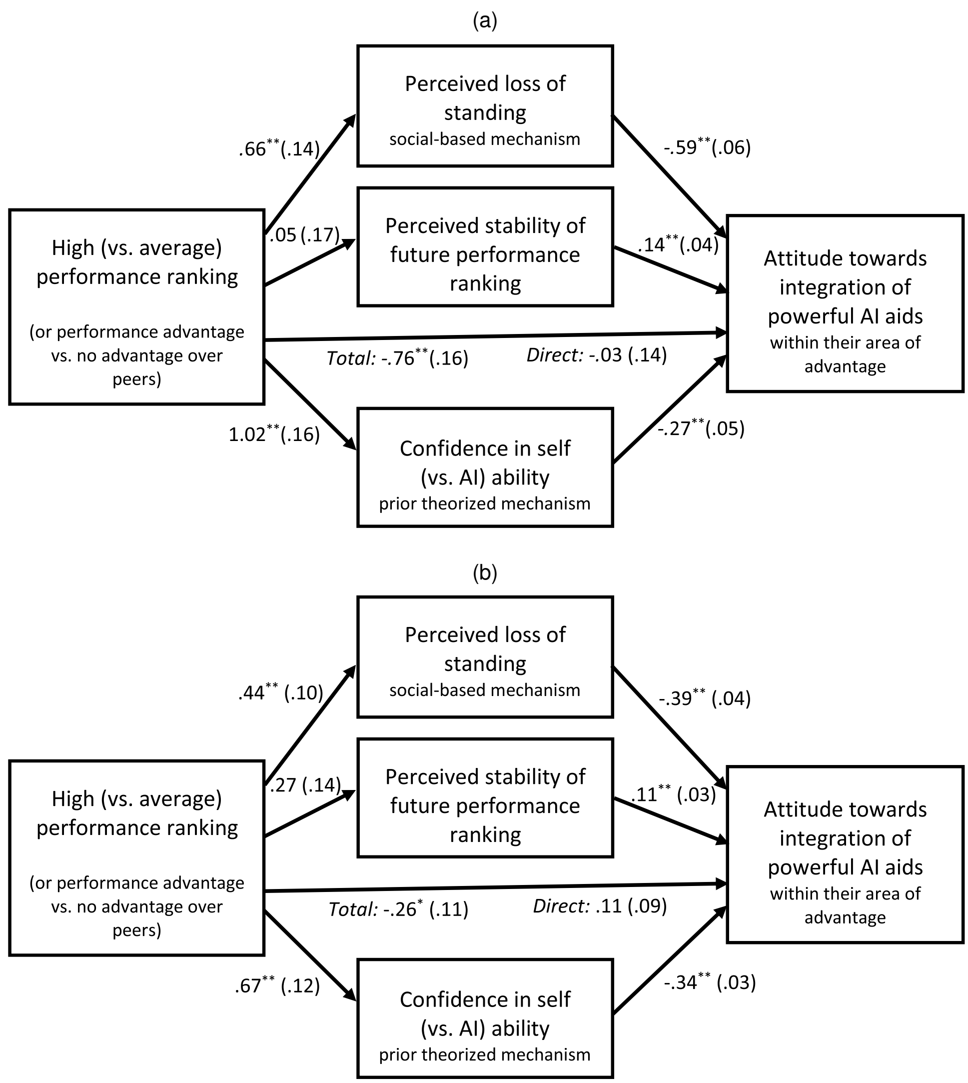
Testing the Mediating Role of Perceived Potential Loss of Standing. I used Hayes' (2017) PROCESS macro, model 4, with parallel mediators to test whether the effect of performance ranking (coded one and zero, for the high and average performance ranking conditions, respectively) on preimplementation AI attitudes (DV) was mediated via both confidence in own (versus AI) ability and perceived loss of standing (Hypothesis 2). Analysis controlled for pay criteria (coded one and zero for the relative and absolute conditions, respectively). Results persisted without this control.

As seen in Figure 3(a), consistent with theorizing, employees assigned to the high (versus average) performance ranking condition indicated greater confidence in own (versus AI) ability, $B = 1.02$, $SE = 0.16$, $p < 0.001$, $\eta_p^2 = 0.13$, as well as greater perceived potential loss of standing, $B = 0.66$, $SE = 0.14$, $p < 0.001$, $\eta_p^2 = 0.05$. The performance ranking manipulation had no effect on perceived stability of future rankings, $B = 0.05$, $SE = 0.17$, $p = 0.75$, with mean rating *lower* than the 4 midpoint (suggesting that future rankings were *not* perceived to be stable).

Moreover, as predicted and further illustrated in Figure 3(a), both greater confidence in their own (versus AI) ability (prior theorized mediator) and perceived potential loss of standing (novel social-based mediator) were uniquely and *negatively* related to positive AI attitudes, $B = -0.27$, $SE = 0.04$, $p < 0.001$, and $B = -0.59$, $SE = 0.06$, $p < 0.001$, respectively, whereas the *direct* effect of performance ranking on preimplementation AI attitude was not significant, $B = -0.03$, $SE = 0.14$, $p = 0.834$. Thus, supporting a full mediation model, the significant total effect of performance ranking on preimplementation AI attitudes was eliminated once including the two mediators as predicting preimplementation AI attitudes. Perceived stability of future rankings—although (as mentioned) not affected by the performance ranking manipulation—was *positively* related to AI attitudes, $B = 0.14$, $SE = 0.04$, $p < 0.001$.

Supporting Hypothesis 2, this analysis revealed the hypothesized negative indirect effect—via greater perceived loss of standing, *Estimate* = −0.39, $SE = 0.09$, 95% confidence interval (CI) (−0.567, −0.223), net of the indirect effect via greater confidence in their own (versus AI) ability, *Estimate* = −0.28, $SE = 0.06$, 95% CI (−0.399, −0.123). These two indirect effects did not differ, *Estimate* = 0.11, $SE = 0.10$, 95% CI (−0.079, 0.324). Ranking stability did not mediate this effect, *Estimate* = 0.01, $SE = 0.03$, 95% CI (−0.047, 0.060).

Figure 3. Mediation Results with Unstandardized Coefficients



Notes. (a) Study 3. (b) Study 4. SEs are in parentheses. $^*p < 0.05$; $^{**}p < 0.001$.

Testing the Second-Stage Moderated Mediation Model. Finally, Hayes' (2017) PROCESS macro, model 14,[11] was used to test the second-stage moderated mediation model (Hypothesis 3). In this model, pay determination criteria (coded one and zero for the relative and absolute conditions, respectively) represented the second-stage moderator.

As theorized, the perceived potential loss of standing × pay determination criterion interaction on AI attitudes was significant, $B = -0.24$, $SE = 0.11$, $p = 0.035$. Specifically, perceived potential loss of standing was negatively associated with positive AI attitudes, $B = -0.47$, $SE = 0.08$, $p < 0.001$, within the absolute condition. Partly supporting Hypothesis 3, this adverse effect was indeed amplified in the relative condition, $B = -0.71$, $SE = 0.08$, $p < 0.001$. However, although the indirect effect via the social-based mechanism was higher in the relative, *Estimate* $= -0.47$, $SE = 0.11$, 95% CI (-0.701, -0.258), than in the absolute, *Estimate* $= -0.31$, $SE = 0.09$, 95% CI (-0.508, -0.147), condition, there was no statistically significant difference between them, *Estimate* $= -0.16$, $SE = 0.11$, 95% CI (-0.392, 0.031). Note that the confidence in own (versus AI) ability × pay determination criterion interaction on AI attitudes was not significant, $p = 0.28$.

In sum, supporting Hypothesis 1 and replicating Studies 1–2, but with distinct PFP settings, high performance ranking had an adverse impact on AI attitudes regardless of PFP nature (relative versus absolute). Supporting Hypothesis 2, the ability- and the social-based mechanisms were found to be two independent (and equally influential) mechanisms driving the adverse impact of high performance ranking on AI attitudes. Partly supporting Hypothesis 3, the negative link between perceived loss of standing and positive AI attitudes was amplified in the relative (versus absolute) condition, but the indirect effects did not statistically differ between absolute and relative PFP criteria.

Study 4

Although Studies 1–3 supported Hypothesis 1 and Hypothesis 2 and partly supported Hypothesis 3 using an experimental design high on internal validity—manipulating performance ranking via a weight-estimation task—this task may be too specific to be generalizable across jobs. Study 4 aimed to replicate these findings in a more natural setting by manipulating performance ranking using employees' recall of actual tasks they (and peers) do on their jobs, thus providing greater external and ecological validity.

Method

Participants were 619 U.S. employees (299 women, $M_{age} = 36.02$, $SD_{age} = 11.05$) compensated with $1 for their participation (see the SM for sample description). They were told that "A prominent trend that characterizes current (and future) workplaces is the use of artificial intelligence (AI) tools that can assist employees on their work tasks. Such AI tools can be used to program code, screen applicants for job positions, diagnose patients, estimate financial risks, etc."

Next, as preregistered, they were allowed to participate only if they were able to recall a work task in their current job (1) that was important for them to do well, (2) in which they were considered to be one of the *high* (*average*) performers among peers (randomly assigned between-subjects factor), and (3) for which they believed there may be currently (or perhaps in the near future) a powerful AI tool that could help them and their peers. Consistent with preregistration, employees who indicated that they were unable to recall such a task were unable to continue and received a partial pay of $0.1 (see the SM for more details on addressing selection concerns).

As mentioned, employees were randomly assigned to one of two performance (high versus average) ranking conditions using a recall task. Employees were asked to describe their recalled task (e.g., "manual fraud detection for online orders," "scanning medical records," and "recoding SQL into Python"). Further categorization of recalled tasks (in general and per condition) are reported in the SM (see Table S2 and Figure S1). One item was then used as a manipulation check, asking participants to indicate—between 0 (worst performers) and 10 (best performers)—where they stood in terms of performance on the particular recalled task.

Next, the extent to which their organization allocates rewards based on relative criteria (second-stage moderator) was measured using two items (*Spearman-Brown* = 0.79) adapted from Belogolovsky and Bamberger (2014): "In my company, my rewards (salary increases, bonuses, and promotions) are typically influenced by" the following: "my performance relative to the performance of others (i.e., allocated on a relative basis)" and "how well others perform compared to me."[12]

The participants were then asked to imagine that their organization was considering "the integration of an easy-to-use AI tool with *exceptional proven capabilities in the recalled task domain*. If integrated, all employees will be able to use it (if they wish) to improve their task outcomes." Consistent with Study 3 and using similar measures on a seven-point scale (see preregistration at https://aspredicted.org/GYV_PGJ), two items (*worsen*/*improve* my standing relative to others; *Spearman-Brown* = 0.43)[13] measured the social-based mechanism and an additional item measured perceived stability (*not change* my standing). Confidence in own (versus AI) ability (the extant ability-based mechanism) was also measured as the difference between two items measuring ability to perform the specific task recalled. All these measures were randomized. Next, employees indicated their preimplementation AI attitude (again using the three items in Studies 2 and 3, $a = 0.96$).

Finally, participants provided their demographics, work-related variables, as well as their AI knowledge (three items, $a = 0.84$; see Study 1) and actual AI usage (see Study 2). Addressing selection concerns (see the SM), taking a conservative approach, all results reported below controlled for all demographics and work- and AI-related factors. All persisted even without these controls.

Results and Discussion

Means (and SDs) of variables per experimental condition are provided in Table 2. As fully elaborated in the SM, the performance ranking manipulation was successful.

Replication: The Adverse Effect of High Performance Ranking on Preimplementation AI Attitudes. A univariate analysis again revealed an adverse effect of performance ranking on preimplementation AI attitudes, $F(1, 608) = 5.55$, $p = 0.02$. Supporting Hypothesis 1, despite the inherent variation in personal recalled tasks, employees assigned to the high performance ranking condition indicated lower positive preimplementation AI attitudes ($M = 5.15$, $SD = 1.52$) than did those assigned to the average ranking condition ($M = 5.39$, $SD = 1.33$), $B = -0.25$, $SE = 0.11$, $p = 0.02$, $\eta_p^2 = 0.01$. This effect persisted when including the proposed second-stage moderator (relative pay criteria; centred) and its interaction with the IV on DV. Specifically, this analysis revealed a main effect of performance ranking on AI attitude, $B = -0.26$, $SE = 0.11$, $p = 0.02$, $\eta_p^2 = 0.01$. As in Study 3, the main effect of relative pay criteria

Table 2. Means and Standard Deviations of Key Measured Variables, per Condition (Study 4)

	Average performance (N = 311)		High performance (N = 308)	
Variable	M	SD	M	SD
MC of performance ranking	6.59_a	1.53	8.30_b	1.29
Confidence in self	5.59_a	1.00	6.13_b	1.00
Confidence in AI	5.54_a	1.11	5.36_b	1.27
Confidence in self *minus* in AI	0.04_a	1.50	0.76_b	1.55
Perceived loss of standing	2.91_a	1.11	3.31_b	1.40
Perceived stability	4.12_a	1.81	4.34_a	1.83
(Positive) AI attitude	5.39_a	1.33	5.15_b	1.52

Notes. N = 619. MC refers to manipulation check. In each row, means with different subscripts differ significantly from each other ($p < 0.05$). The difference between "a" and "b" subscripts in the second row is marginal ($p = 0.06$).

and the IV × moderator interaction were not significant, $p = 0.28$ and $p = 0.15$, respectively.

Testing the Mediating Role of Perceived Potential Loss of Standing. As in Study 3, Hayes' (2017) PROCESS macro, model 4, with parallel mediators was used to test Hypothesis 2, controlling for relative pay criteria, all demographics, and work- and AI-related factors. As in Study 3 and as illustrated in Figure 3(b), employees assigned to the high (versus average) ranking condition indicated greater confidence in self (versus AI), $B = 0.67$, $SE = 0.12$, $p < 0.001$, as well as greater perceived loss of standing, $B = 0.44$, $SE = 0.10$, $p < 0.001$. Moreover, as predicted, the two (ability- and social-based) mediators were uniquely and negatively related to (positive) preimplementation AI attitudes, $B = -0.34$, $SE = 0.03$, $p < 0.001$, and $B = -0.39$, $SE = 0.04$, $p < 0.001$, respectively, whereas the *direct* effect of performance ranking on preimplementation AI attitude was not significant, $B = 0.11$, $SE = 0.09$, $p = 0.239$ (again, pointing to full mediation by both mediators). Supporting Hypothesis 2, both negative indirect effects (via the respective mechanisms) were significant, *Estimate* = −0.23, $SE = 0.05$, 95% CI (−0.331, −0.138) and *Estimate* = −0.17, $SE = 0.04$, 95% CI (−0.256, −0.086), respectively. Again, these indirect effects did not differ significantly, *Estimate* = −0.06, $SE = 0.06$, 95% CI (−0.180, 0.062).

Finally, as in Study 3, although perceived ranking stability was positively related to (positive) AI attitudes, $B = 0.11$, $SE = 0.03$, $p < 0.001$, it did not mediate the effect of performance ranking on AI attitudes, *Estimate* = 0.03, $SE = 0.02$, 95% CI (−0.004, 0.066).

Testing the Second-Stage Moderated Mediation Model. Next, as in Study 3, Hayes' (2017) PROCESS macro, model 14, was used to test the second-stage moderated mediation (see Figure 1)—here, with the *measured* moderator of relative pay criteria (centred). Greater confidence in own (versus AI) ability and perceived potential loss of standing were again both uniquely and negatively related to (positive) preimplementation AI attitudes, $B = -0.33$, $SE = 0.03$, $p < 0.001$, and $B = -0.38$, $SE = 0.04$, $p < 0.001$, respectively. Here, inconsistent with Hypothesis 3, the perceived potential loss of standing × relative pay criterion interaction on AI attitude was not significant, $p = 0.87$ (thus, the negative indirect effects of performance ranking on AI attitudes via perceived potential loss of standing at different levels of the moderator did not statistically differ).

In sum, supporting Hypothesis 1 and replicating Studies 1–3, high performance ranking had an adverse impact on AI attitudes toward powerful AI aids in the context of employees' actual work tasks. Supporting Hypothesis 2, findings revealed again the mediating role of perceived loss of standing in addition to the mediation via confidence in own (versus AI) ability. Although here, unlike in Study 3, an amplification mediator-DV effect was not observed, as in Study 3, the negative indirect effects were evident regardless of the extent to which relative pay criteria were used in one's organization.

Study 5

Study 5 sought to provide stronger causal evidence for the social-based mechanism, using a moderation-of-process design (Spencer et al. 2005) by manipulating the stability of future rankings in a realistic workplace scenario. Specifically, it aimed to examine whether bolstering the stability of future rankings (presumably counteracting one's concern with potential loss of standing) can eliminate the adverse effect of high performance ranking on preimplementation AI attitudes.

Only employees who could imagine AI aids assisting them (perhaps in the near future) on their work tasks were included. Eligible participants were randomly assigned to a 2 (performance ranking: high, average) × 2 (ranking stability: stable, unstable) × 2 (pay determination criteria: relative, absolute) between-subjects factorial design. As preregistered, the study primarily aimed to test the two-way performance ranking × ranking stability interaction on AI attitudes.

Method

Participants were 892 U.S. employees (398 women, M_{age} = 39.74, SD = 11.78) compensated with $1 for their participation (see the SM for sample description). They were asked to imagine they "work for a company that is considered a good place to work. You work as one of the programmers in one of its largest divisions … It is very important for you to do well on your job." Next, performance ranking was manipulated: "Imagine that you are currently one of the *high* (*average*) performers (among peers) in your division." The success of this manipulation was verified, using again the 0 = *worst performers* to 10 = *best performers* scale.

Participants were then asked to "Imagine that your company is considering to integrate a powerful and easy-to-use AI tool, *with exceptional proven programming capabilities that can assist you (and your peers) to write basic code faster so that you can focus on more complex programming*. If integrated, all employees will be able to use it to improve their work outcomes." The manipulation of ranking (stability/instability) was then conducted. Specifically, participants further read that in other companies that integrated this AI aid, "(all employees benefited equally, perhaps in different ways/some employees benefited more or less than others) from its integration. Thus, if integrated in your division, one's performance ranking relative to peers is (NOT expected/expected) to change." A manipulation check was conducted: "According to the above, if this powerful AI is integrated, one's performance ranking relative to peers is expected to change" (1 = *strongly disagree* to 7 = *strongly agree*).

Next, pay determination criteria were manipulated. Those in the relative (versus absolute) condition were told this company rewarded employees based on their relative performance (versus based on their own level of performance only). This manipulation was verified using three items (see Study 3; a = 0.91). Finally, we reminded participants about the (manipulated) scenario key factors—performance ranking, ranking stability, and organizational pay criteria—and asked them to note their preimplementation AI attitude using three items (a = 0.94; see Studies 2–4). As in previous studies, participants finally indicated their demographics and work- and AI-related variables.

Results and Discussion

As fully elaborated in the SM, the three (performance ranking, ranking stability, and pay criteria) manipulations were successful. Notably, all the effects reported below persisted when controlling for all demographics and work- and AI-related variables in further robust analyses.

As preregistered, a univariate three-way analysis was conducted including the three key factors, all their two-way interactions, and the three-way interaction, focusing primarily on the performance ranking × stability interaction, which (as expected) was significant, $F(1, 884) = 18.23$, $p < 0.001$, $\eta_p^2 = 0.02$. Given that the three-way interaction was not significant, $F(1, 884) = 0.28$, $p = 0.60$ (implying that the pay criterion manipulation did not change the pattern of the key performance ranking × stability interaction), to simplify the presentation of the results, I utilized a univariate two-way analysis (the three-way analysis is fully interpreted in the SM).

The two-way analysis included performance ranking, stability, and performance ranking × stability interaction (the focus of Study 5) on AI attitudes and *controlling* for pay criteria. Reassuringly, this analysis revealed the same performance ranking × stability interaction effect, $F(1, 887) = 18.23$, $p < 0.001$, $\eta_p^2 = 0.02$. As seen in Figure 4, this analysis revealed that the adverse effect of high (versus average) performance ranking on AI attitude, $M = 5.48$, $SD = 1.26$, versus $M = 5.82$, $SD = 1.00$, $B = -0.34$, $SE = 0.10$, $p = 0.001$, $\eta_p^2 = 0.01$, was evident only in the *unstable* condition. This effect was reversed in the *stable* condition. Specifically, when stability of future rankings was assured, the effect of high (versus average) performance ranking on AI attitude was positive, $M = 5.91$, $SD = 1.08$, versus $M = 5.62$, $SD = 1.08$, $B = 0.29$, $SE = 0.10$, $p = 0.006$, $\eta_p^2 = 0.01$.

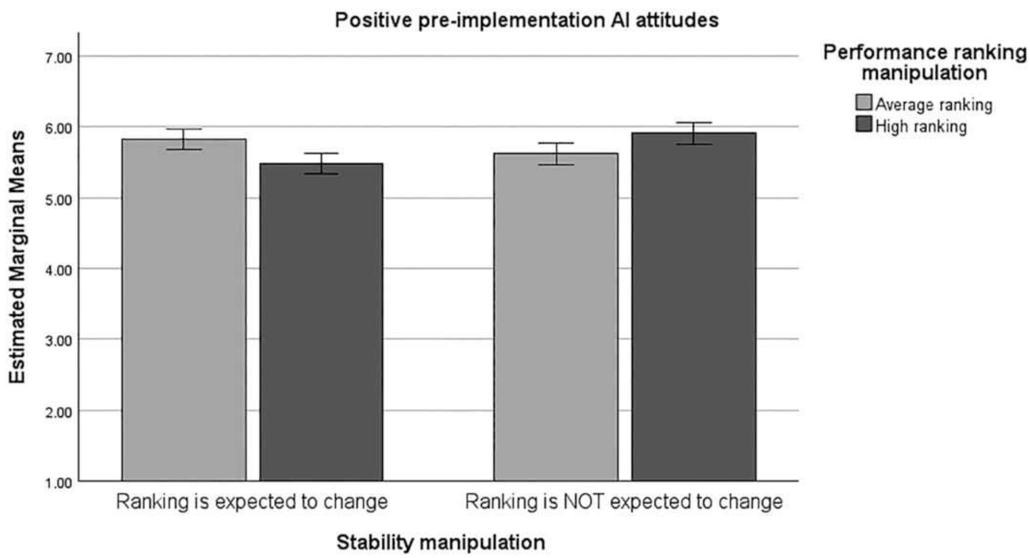
In sum, the adverse effect of high performance ranking on AI attitudes was evident in the unstable condition. This effect was not only eliminated but even reversed once it was made explicit that future rankings were *not* expected to change. It may be that counteracting high performers' concern with future standing sensitized them to explore this new opportunity. In contrast, perhaps it reduced average performers' motivation to use an AI aid not expected to improve their standing.

General Discussion

Powerful AI aids are increasingly entering the workplace. Yet, algorithm aversion is still prevalent among employees. Given the key role of employee preimplementation AI attitudes in successful implementation, scholars and practitioners have devoted substantial efforts to identifying factors affecting AI attitudes. The current work uncovers the adverse effect of employees' high performance ranking on their preimplementation attitudes toward powerful AI aids.

Five studies conducted among undergraduates (Study 1) and U.S. employees (Studies 2–5) and using a variety of experimental settings and designs—a weight estimation simulation (Studies 1–3), recall of actual work tasks (Study 4), and a workplace scenario (Study 5)—provided consistent support for this causal effect and the mediating role of perceived potential loss of standing. Pointing to the power of symbolic threats and consistent with employees' desire to protect their standing (Mitchell et al. 2020), the adverse effect of

Figure 4. Performance Ranking × Stability Interaction on AI Attitudes, Controlling for Pay Criteria (Study 5)



high performance ranking on AI attitudes was evident in the absence of financial incentives for high performance (Study 1), in various incentive-based settings (Studies 2–3), as well as across different PFPs used in one's organization (Study 4).

This research offers several theoretical implications. First, extending research on factors contributing to algorithm aversion (Mahmud et al. 2022) and the role of perceived risks in human-AI relationships (Solberg et al. 2022), it addresses the gap in our understanding of how the social context may influence AI implementation. Specifically, if one is concerned with future standing, limiting the advancement of lower-ranking peers (by limiting their access to powerful AI aids) maintains one's high position in the organizational hierarchy. As such, the current research also extends prior findings on employee behavior aimed at protecting one's standing (Garcia et al. 2010)—being less supportive of powerful AI aids as an indirect means to keep peers in their lower place.

Second, the current research reveals that performance ranking may adversely impact employee pre-implementation AI attitudes not only as a function of one's confidence in own (versus AI) ability as theorized by Burton et al. (2020), but also because of a novel social-based mechanism—high performers' concern with loss of standing. This extends our understanding of why experts show greater aversion toward algorithms (Arkes et al. 1986, Logg et al. 2019, Allen and Choudhury 2022).

Third, although trust in AI's ability is typically related to positive AI attitudes (Glikson and Woolley 2020), the current findings indicate that even when AI is clearly able to improve their work, employees' preimplementation AI attitudes can be negatively affected by social factors involving downward assimilative comparisons. Prior research posits that threatening social comparisons may lead to behaviors that can benefit the individual in the short term, but have long-term consequences for relationships, groups, and organizations (Johnson 2012). Extending these findings to the realm of AI aids, the current research suggests that social comparisons can trigger self-protective motives that ultimately jeopardize the successful implementation of powerful AI aids in organizations.

From a practical perspective, employees may unwittingly fail to promote organizational performance by protecting their personal standing. The current findings highlight the complex nature of future work, which managers should consider when integrating AI. Enhancing management or peer support for AI (Venkatesh and Bala 2008) may do little to mitigate the adverse impact of perceived loss of standing. Thus, managers should consider not only technological aspects in the *employee-AI* interface, but also how the integration of AI aids may affect *employee-employee* relations, social comparison processes, and the general competitive climate within workgroups.

Limitations and Future Directions

Several limitations and corresponding avenues for future research should be noted. First, the results related to the amplification role of relative pay determination criteria in the link between perceived potential loss of standing and AI attitudes (Hypothesis 3) were inconsistent: in Study 3, involving a PFP criterion *manipulated within the focal task*, this amplification was evident, whereas in Study 4, a more *general measured* PFP at one's workplace did not moderate this link. It may be that in Study 3, involving a more controlled short-term setting, participants cared more about their financial gains or less about their psychological standing. However, in a real job setting (Study 4), psychological standing may play such an important role that relative pay criteria do not add significantly more. Moreover, in such natural settings, pay is not the only factor contributing to one's standing in the organizational hierarchy. Relatedly, in both studies, the moderated mediation hypothesis was not supported because the negative indirect effects for different levels of the moderator did not statistically differ. Future studies may delve into the moderating role of pay determination criteria by using contexts characterized by greater employee access to concrete organizational resources.

Second, although the adverse causal effect of performance ranking on AI attitudes was replicated using diverse experimental designs and settings, its effect size was relatively small. Several reasons may explain this. First, given the AI's ability was held constant and high across all studies and conditions, participants' supportive AI attitudes were relatively high, making it more difficult to observe the hypothesized effect. Further noise in Study 4 (inherent variation in personal recalls) may also have prevented the observation of larger effect size. Finally, although the use of experiments is common in the study of employee attitudes (given their high realism and strong evidence for causality), they involve relatively low stakes. Thus, the effects observed may have been underestimated compared with real-life settings. Relatedly, performance ranking is only one antecedent driving employee concern with a potential loss of standing. Future studies may explore other sources driving such concerns. Future research directions may also explore novel factors influencing algorithm aversion—specifically, how factors like perceived managerial control, the source driving the AI implementation, or its perceived motivations can affect employees' preimplementation AI attitudes.

Third, using a "moderation-of-process" design, Study 5, not only provided stronger causal evidence for the social-based mechanism but also offered a potential effective managerial intervention. Future studies may explore the effectiveness of this stability intervention among employees facing actual AI implementations. Other interventions can be tested as well: for example, using self-affirmation interventions to counteract perceived personal threats (Cohen and Sherman 2014), such as writing about one's virtues or superiority in domains currently not supported by AI.

Fourth, given the current focus on augmentation, employees' fear of being replaced by AI, a realistic threat leading to negative attitudes (Kaplan and Haenlein 2020), was not explored. Indeed, given that AI aids are increasingly able to take over many tasks, they may be perceived by some employees (e.g., low-ranking employees) as able to take away their jobs. Examining whether these employees would be able to anticipate personal utility gains from such aids is an interesting avenue to explore. Moreover, although the focus here was on powerful AI aids within the focal task domain, the phenomenon observed is relevant to any technological advancements perceived as able to level the playing field on any self-relevant dimension. Yet, if such innovations are not expected to do so, or their power is not within a domain which poses a threat to the employee's identity, I expect the effect observed to be attenuated. In sum, the exploration of distinct attitudes by high- and low-ranked employees and for different levels of "innovative threat" presents a fruitful area for future research.

Finally, most studies were conducted among employees in the United States—an individualistic culture—and thus cannot be generalized to other cultures. Previous research found that higher collectivism scores were associated with a decreased desire to make downward comparisons (Chung and Mallery 1999). Thus, it may well be that in collectivistic cultures, employees' concerns with group utility would attenuate the effect observed.

Conclusion

Understanding how employees respond to powerful algorithmic aids is crucial for helping organizations leverage their power. Joining recent calls (Glikson and Woolley 2020, Wiesenfeld et al. 2022) to adopt a human-centred approach or a sociotechnical perspective to the study of AI in organizations, I hope this work would encourage scholars to identify other social factors involved in employee attitudes and interfaces with such powerful aids.

Acknowledgments

The author is grateful to Peter Bamberger, Daniel Heller, and Ella Glikson for their helpful comments and suggestions on an earlier draft of this manuscript. The author also thanks her research assistants, Chen Dabestany, Aseel Nadaf, and Olga Mondrus, for their assistance with this research.

Endnotes

[1] AI is used as an umbrella term that encompasses a variety of technologies enabling machines to mimic human intelligence (e.g., machine learning, deep learning, and natural language processing).

[2] Being a high performer does not necessarily imply domain expertise. Conversely, individuals may have deep knowledge, skill, and experience in a domain (i.e., domain expertise) without necessarily exhibiting a performance advantage compared with peers.

[3] For model simplicity, when assessing potential underlying mechanisms (Studies 3–4; see also Hypothesis 2), the focus is on the underlying cognitive mechanism (i.e., perceived potential loss of standing) rather than on the associated emotions of fear and concern with one's potential loss of standing.

[4] Photographs were taken from https://unsplash.com/, a free website for pictures. As mentioned by several participants, weight assessment based on these pictures, without being provided with individuals' heights, was not an easy task. Perceived difficulty of this task was assessed in Study 2.

[5] Note that Studies 2 and 3 used incentive-based settings within the same weight simulation.

[6] In fact, here and in Study 3, all participants were given an \$0.2 bonus at the end of data collection.

[7] Results of these robust analyses (here as well as in all the next studies) are available upon request.

[8] Based on Study 2's results (η_p^2 = 0.02), using a priori statistical power analysis (Faul et al. 2007) with effect size (f = 0.143) with power = 0.80, alpha of 0.05, for four conditions.

[9] Results of these analyses are available upon request. Syntax of key CFAs is provided in the OSF link.

[10] I also measured perceived AI usefulness to others (versus to self)—conceptually representing an indirect proxy of perceived potential loss of standing in that when one perceives AI usefulness to others as higher than to self, it reflects a potential loss of standing. This measure mediated the IV-DV effect but only when the more direct measure of perceived potential loss of standing (reported above) was *not* included in the analysis. For clarity, I report here the results using the direct measure (the one being used also in Study 4; see https://aspredicted.org/GYV_PGJ). Detailed findings and analyses related to the AI usefulness to others (versus to self) measure are fully provided in the SM.

[11] All the reported effects persisted here (and similarly in Study 4) when using model 15, which includes also the direct effect of the IV × moderator on the DV.

[12] A third reversed item "my own level of performance only (i.e., allocated on an absolute basis)" was dropped because it reduced reliability (a = 0.28, as opposed to 0.79 without it).

[13] All the reported effects persisted even when including only the "worsen" item as the mediator.

References

Allen RT, Choudhury P (2022) Algorithm-augmented work and domain experience: The countervailing forces of ability and aversion. *Organ. Sci.* 33(1):149–169.

Arkes HR, Dawes RM, Christensen C (1986) Factors influencing the use of a decision rule in a probabilistic task. *Organ. Behav. Human Decision Processes* 37(1):93–110.

Asgari E, Hunt RA, Lerner DA, Townsend DM, Hayward ML, Kiefer K (2021) Red giants or black holes? The antecedent conditions and multilevel impacts of star performers. *Acad. Management Ann.* 15(1):223–265.

Barr SH, Conlon EJ (1994) Effects of distribution of feedback in work groups. *Acad. Management J.* 37(3):641–655.

Belogolovsky E, Bamberger PA (2014) Signaling in secret: Pay for performance and the incentive and sorting effects of pay secrecy. *Acad. Management J.* 57(6):1706–1733.

Brown DJ, Ferris DL, Heller D, Keeping LM (2007) Antecedents and consequences of the frequency of upward and downward social comparisons at work. *Organ. Behav. Human Decision Processes* 102(1):59–75.

Brynjolfsson E, Li D, Raymond LR (2023) Generative AI at work. NBER Working Paper No. 31161, National Bureau of Economic Research, Cambridge, MA.

Burton JW, Stein MK, Jensen TB (2020) A systematic review of algorithm aversion in augmented decision making. *J. Behav. Decision Making* 33(2):220–239.

Buunk AP, Gibbons FX (2007) Social comparison: The end of a theory and the emergence of a field. *Organ. Behav. Human Decision Processes* 102(1):3–21.

Call ML, Nyberg AJ, Thatcher S (2015) Stargazing: An integrative conceptual review, theoretical reconciliation, and extension for star employee research. *J. Appl. Psych.* 100(3):623–640.

Campbell EM, Liao H, Chuang A, Zhou J, Dong Y (2017) Hot shots and cool reception? An expanded view of social consequences for high performers. *J. Appl. Psych.* 102(5):845–866.

Chernyak-Hai L, Davidai S (2022) "Do not teach them how to fish": The effect of zero-sum beliefs on help giving. *J. Experiment. Psych. General* 151(10):2466–2480.

Chiu YT, Zhu YQ, Corbett J (2021) In the hearts and minds of employees: A model of pre-adoptive appraisal toward artificial intelligence in organizations. *Internat. J. Inform. Management* 60:102379.

Chung T, Mallery P (1999) Social comparison, individualism-collectivism, and self-esteem in China and the United States. *Current Psych.* 18(4):340–352.

Cohen GL, Sherman DK (2014) The psychology of change: Self-affirmation and social psychological intervention. *Annual Rev. Psych.* 65(1):333–371.

Dell'Acqua F, McFowland E, Mollick ER, Lifshitz-Assaf H, Kellogg K, Rajendran S, Krayer L, Candelon F, Lakhani KR (2023) Navigating the jagged technological frontier: Field experimental evidence of the effects of AI on knowledge worker productivity and quality. Working Paper No. 24-013, Harvard Business School Technology & Operations Management Unit, Cambridge, MA.

Dietvorst BJ, Simmons JP, Massey C (2015) Algorithm aversion: People erroneously avoid algorithms after seeing them err. *J. Experiment. Psych. General* 144(1):114–126.

Dwivedi YK, Rana NP, Janssen M, Lal B, Williams MD, Clement M (2017) An empirical validation of a unified model of electronic government adoption (UMEGA). *Government Inform. Quart.* 34(2):211–230.

Ericsson KA, Towne TJ (2010) Expertise. *Wiley Interdisciplinary Rev. Cognitive Sci.* 1(3):404–416.

Faul F, Erdfelder E, Lang AG, Buchner A (2007) G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behav. Res. Methods* 39(2):175–191.

Garcia SM, Song H, Tesser A (2010) Tainted recommendations: The social comparison bias. *Organ. Behav. Human Decision Processes* 113(2):97–101.

Garcia SM, Tor A, Schiff TM (2013) The psychology of competition: A social comparison perspective. *Perspect. Psych. Sci.* 8(6):634–650.

Gerhart B, Rynes SL, Fulmer IS (2009) Pay and performance: Individuals, groups, and executives. *Acad. Management Ann.* 3(1): 251–315.

Glikson E, Woolley AW (2020) Human trust in artificial intelligence: Review of empirical research. *Acad. Management Ann.* 14(2): 627–660.

Goddard K, Roudsari A, Wyatt JC (2012) Automation bias: A systematic review of frequency, effect mediators, and mitigators. *J. Amer. Medical Inform. Assoc.* 19(1):121–127.

Goodman PS, Haisley E (2007) Social comparison processes in an organizational context: New directions. *Organ. Behav. Human Decision Processes* 102(1):109–125.

Grant H, Dweck CS (2003) Clarifying achievement goals and their impact. *J. Personality Soc. Psych.* 85(3):541–553.

Greenberg J, Ashton-James CE, Ashkanasy NM (2007) Social comparison processes in organizations. *Organ. Behav. Human Decision Processes* 102(1):22–41.

Grove WM, Zald DH, Lebow BS, Snitz BE, Nelson C (2000) Clinical vs. mechanical prediction: A meta-analysis. *Psych. Assessment* 12(1):19–30.

Hanemaayer A (2021) Don't touch my stuff: Historicising resistance to AI and algorithmic computer technologies in medicine. *Interdisciplinary Sci. Rev.* 46(1–2):126–137.

Hayes AF (2017) *Introduction to Mediation, Moderation, and Conditional Process Analysis: A Regression-Based Approach*, 2nd ed. (Guilford Publications, New York).

Heilman ME, Block CJ, Stathatos P (1997) The affirmative action stigma of incompetence: Effects of performance information ambiguity. *Acad. Management J.* 40(3):603–625.

Hendricks JL, Call ML, Campbell EM (2023) High performer peer effects: A review, synthesis, and agenda for future research. *J. Management* 49(6):1997–2029.

Herold DM, Farmer SM, Mobley MI (1995) Pre-implementation attitudes toward the introduction of robots in a unionized environment. *J. Engrg. Tech. Management* 12(3):155–173.

Hosny A, Parmar C, Quackenbush J, Schwartz LH, Aerts HJWL (2018) Artificial intelligence in radiology. *Nature Rev. Cancer* 18(8):500–510.

Huang MH, Rust RT (2018) Artificial intelligence in service. *J. Service Res.* 21(2):155–172.

Huijsmans I, Ma I, Micheli L, Civai C, Stallen M, Sanfey AG (2019) A scarcity mindset alters neural processing underlying consumer decision making. *Proc. Natl. Acad. Sci. USA* 116(24):11699–11704.

Jarrahi MH (2018) Artificial intelligence and the future of work: Human-AI symbiosis in organizational decision making. *Bus. Horizons* 61(4):577–586.

Johnson C (2012) Behavioral responses to threatening social comparisons: From dastardly deeds to rising above. *Soc. Personality Psych. Compass* 6(7):515–524.

Jussupow E, Spohrer K, Heinzl A (2022) Identity threats as a reason for resistance to artificial intelligence: Survey study with medical students and professionals. *JMIR Formative Res.* 6(3):e28750.

Kaplan A, Haenlein M (2020) Rulers of the world, unite! The challenges and opportunities of artificial intelligence. *Bus. Horizons* 63(1):37–50.

Kaplan AD, Kessler TT, Brill JC, Hancock PA (2021) Trust in artificial intelligence: Meta-analytic findings. *Human Factors* 65(2):337–359.

Kehoe RR, Lepak DP, Bentley FS (2018) Let's call a star a star: Task performance, external status, and exceptional contributors in organizations. *J. Management* 44(5):1848–1872.

Kellogg KC, Valentine MA, Christin A (2020) Algorithms at work: The new contested terrain of control. *Acad. Management Ann.* 14(1):366–410.

Kim HW, Kankanhalli A (2009) Investigating user resistance to information systems implementation: A status quo bias perspective. *MIS Quart.* 33(3):567–582.

Kuncel NR, Klieger DM, Connelly BS, Ones DS (2013) Mechanical vs. clinical data combination in selection and admissions decisions: A meta-analysis. *J. Appl. Psych.* 98(6):1060–1072.

Lazear EP, Oyer P (2013) Personnel economics. Gibbons R, Roberts J, eds. *The Handbook of Organizational Economics* (Princeton University Press, Princeton, NJ), 479–519.

Lebovitz S, Lifshitz-Assaf H, Levina N (2022) To engage or not to engage with AI for critical judgments: How professionals deal with opacity when using AI for medical diagnosis. *Organ. Sci.* 33(1):126–148.

Leonard-Barton D (1985) Experts as negative opinion leaders in the diffusion of a technological innovation. *J. Consumer Res.* 11(4): 914–926.

Lewis LK, Russ TL (2012) Soliciting and using input during organizational change initiatives: What are practitioners doing. *Management Comm. Quart.* 26(2):267–294.

Lichtenthaler U (2018) Substitute or synthesis: The interplay between human and artificial intelligence. *Res. Tech. Management* 61(5):12–14.

Lichtenthaler U (2020) Extremes of acceptance: Employee attitudes toward artificial intelligence. *J. Bus. Strategy* 41(5):39–45.

Logg JM, Minson JA, Moore DA (2019) Algorithm appreciation: People prefer algorithmic to human judgment. *Organ. Behav. Human Decision Processes* 151:90–103.

Magee JC, Galinsky AD (2008) Social hierarchy: The self-reinforcing nature of power and status. *Acad. Management Ann.* 2(1):351–398.

Mahmud H, Islam AKMN, Ahmed SI, Smolander K (2022) What influences algorithmic decision-making? A systematic literature review on algorithm aversion. *Tech. Forecasting Soc. Change* 175:121390.

Matta FK, Van Dyne L (2020) Understanding the disparate behavioral consequences of LMX differentiation: The role of social comparison emotions. *Acad. Management Rev.* 45(1): 154–180.

Mayfield P (2014) Engaging with stakeholders is critical when leading change. *Indust. Commercial Training* 46(2):68–72.

Metcalf L, Askay DA, Rosenberg LB (2019) Keeping humans in the loop: Pooling knowledge through artificial swarm intelligence to improve business decision making. *Calif. Management Rev.* 61(4):84–109.

Michlitsch JF (2000) High-performing, loyal employees: The real way to implement strategy. *Strategy Leadership* 28(6):28–33.

Mitchell RL, Bae KK, Case CR, Hays NA (2020) Drivers of desire for social rank. *Current Opinion Psych.* 33:189–195.

Mittone L, Savadori L (2009) The scarcity bias. *Appl. Psych.* 58(3): 453–468.

Molleman E, Nauta A, Buunk BP (2007) Social comparison-based thoughts in groups: Their associations with interpersonal trust and learning outcomes. *J. Appl. Soc. Psych.* 37(6):1163–1180.

Moore DA, Klein WM (2008) Use of absolute and comparative performance feedback in absolute and comparative judgments and decisions. *Organ. Behav. Human Decision Processes* 107(1):60–74.

Mussweiler T, Rüter K, Epstude K (2004) The ups and downs of social comparison: Mechanisms of assimilation and contrast. *J. Personality Soc. Psych.* 87(6):832–844.

Noy S, Zhang W (2023) Experimental evidence on the productivity effects of generative artificial intelligence. *Science* 381(6654):187–192.

Oppenheimer DM, Meyvis T, Davidenko N (2009) Instructional manipulation checks: Detecting satisficing to increase statistical power. *J. Experiment. Soc. Psych.* 45(4):867–872.

Prakash AV, Das S (2021) Medical practitioner's adoption of intelligent clinical diagnostic decision support systems: A mixed-methods study. *Inform. Management* 58(7):103524.

Reh S, Tröster C, Van Quaquebeke N (2018) Keeping (future) rivals down: Temporal social comparison predicts coworker social undermining via future status threat and envy. *J. Appl. Psych.* 103(4):399–415.

Rogers EM (1982) *Diffusion of Innovations*, 3rd ed. (The Free Press, New York).

Sánchez-Prieto JC, Cruz-Benito J, Therón R, García-Peñalvo F (2020) Assessed by machines: Development of a TAM-based tool to measure AI-based assessment acceptance among students. *Internat. J. Interactive Multimedia Artificial Intelligence* 6(4):80–86.

Shrestha YR, Krishna V, von Krogh G (2021) Augmenting organizational decision-making with deep learning algorithms: Principles, promises, and challenges. *J. Bus. Res.* 123:588–603.

Smith RH (2000) Assimilative and contrastive emotional reactions to upward and downward social comparisons. Suls J, Wheeler L, eds. *Handbook of Social Comparison. The Springer Series in Social Clinical Psychology* (Springer, Boston), 173–200.

Solberg E, Kaarstad M, Eitrheim MHR, Bisio R, Reegård K, Bloch M (2022) A conceptual model of trust, perceived risk, and reliance on AI decision aids. *Group Organ. Management* 47(2):187–222.

Spencer SJ, Zanna MP, Fong GT (2005) Establishing a causal chain: Why experiments are often more effective than mediational analyses in examining psychological processes. *J. Personality Soc. Psych.* 89(6):845–851.

Sutton RI (2007) The war for talent is back. *Harvard Bus. Rev.* (April 23), https://hbr.org/2007/04/the-war-for-talent-is-back.

Taylor S, Todd P (1995) Decomposition and crossover effects in the theory of planned behavior: A study of consumer adoption intentions. *Internat. J. Res. Marketing* 12(2):137–155.

Thompson ER (2007) Development and validation of an internationally reliable short-form of the Positive and Negative Affect Schedule (PANAS). *J. Cross Cultural Psych.* 38(2):227–242.

Venkatesh V, Bala H (2008) Technology acceptance model 3 and a research agenda on interventions. *Decision Sci.* 39(2): 273–315.

Von Krogh G (2018) Artificial intelligence in organizations: New opportunities for phenomenon-based theorizing. *Acad. Management Discoveries* 4(4):404–409.

Wiesenfeld BM, Aphinyanaphongs Y, Nov O (2022) AI model transferability in healthcare: A sociotechnical perspective. *Nature Machine Intelligence* 4(10):807–809.

Yen HR, Chuang SC (2008) The effect of incidental affect on preference for the status quo. *J. Acad. Marketing Sci.* 36(4):522–537.

Ilanit SimanTov-Nachlieli is an assistant professor of management specializing in organizational behavior at the Coller School of Management, Tel Aviv University. She received her master's in social psychology with honors from VU University in Amsterdam and her PhD from the School of Psychological Sciences at Tel Aviv University. Her research focuses on artificial intelligence in organizations, pay transparency, employee prosocial and counterproductive behaviors, and trust.