Perceived Opacity Increases Workplace Algorithm Aversion

Authors: Yijun Zhao, Xu Liying, Yu Feng, Jin Wanglong, XU Liying, Yu Feng

Date: 2023-12-22T00:00:00+00:00

Abstract

The use of algorithms as assistance and replacement for human decision-making in the workplace is increasingly common, yet people exhibit algorithm aversion. Through four progressive experiments conducted across various workplace application scenarios, this study compared individuals' attitudes toward decisions made by human versus algorithmic decision-makers, and examined the underlying psychological mechanisms and boundary conditions. The results revealed that in workplace contexts, individuals demonstrated lower acceptability, preference, and utilization willingness toward algorithmic decisions compared to those made by human decision-makers, thereby exhibiting "algorithm aversion." The underlying psychological mechanism of this phenomenon is that people perceive algorithmic decisions as less transparent than human decisions (Experiments 2~3). Further investigations revealed that when algorithms were endowed with anthropomorphic features, individuals' aversion to algorithmic decisions was reversed, and their acceptance attitudes were enhanced (Experiment 4). These findings contribute to a better understanding of people's reactions to algorithmic decision-making and provide implications for promoting intelligent social governance and guiding the ethical use of algorithms.

Full Text

Perceived Opacity Increases Algorithm Aversion in the Workplace

ZHAO Yijun, XU Liying, YU Feng, JIN Wanglong
(Department of Psychology, Wuhan University, Wuhan 430072, China)

Abstract

Algorithms are increasingly used as aids and replacements for human decision-making in the workplace, yet people exhibit algorithm aversion. Through four sequential experiments, this study compared attitudes toward decisions made by human versus algorithmic decision-makers across various workplace scenarios, exploring the underlying mechanisms and boundary conditions. Results revealed that in workplace contexts, people show lower permissibility, liking, and willingness to utilize algorithmic decisions compared to human decisions, demonstrating "algorithm aversion." The intrinsic psychological mechanism is that algorithmic decisions are perceived as more opaque than human decisions (Experiments 2–3). Further research found that when algorithms are anthropomorphized, this aversion is reversed and acceptance increases (Experiment 4). These findings enhance understanding of reactions to algorithmic decision-making and provide insights for promoting intelligent social governance and ethical algorithm usage.

Keywords: algorithm aversion, transparency, anthropomorphism, workplace
Classification: B849: C91

In August 2022, Meta, Facebook's parent company, used an algorithm to lay off 60 contract workers. Many employees reported receiving no explanation beyond being told they were "randomly" terminated by an algorithm (Hays, 2022). Meta is hardly alone in using algorithms for human resource management. As early as 2015, Amazon developed an AI efficiency monitoring and evaluation system for warehouse management to track employee performance in real-time and incorporate it into performance reviews (Upadhye, 2018). In August 2021, Russian online payment service company Xsolla similarly used algorithm-assessed "digital footprints" to fire 150 employees deemed disengaged and unproductive (McAloon, 2021). Algorithm use in the workplace has become a social hot topic. In recent years, more enterprises have applied algorithms to workplace environments, improving organizational efficiency while assuming and liberating HR managers from tedious tasks (Basu et al., 2023; Brynjolfsson & McAfee, 2014; Chalfin et al., 2016; Duggan et al., 2020; Garg et al., 2022). According to a survey by online recruitment platform CareerBuilder, approximately 55% of HR managers in the United States believe AI systems centered on algorithms will become a regular part of their work within the next few years (HR Daily Advisor Staff, 2017). Thus, intelligent decision-making systems represented by algorithms are gradually transitioning from concept to reality in workplace integration.

Consistent with the reasons algorithms have rapidly swept through medical diagnosis (Hao, 2020), financial investment (Ahmed et al., 2022), judicial sentencing (Hao, 2019), and autonomous driving (Badue et al., 2021), algorithms' rapid workplace adoption stems primarily from inherent limitations in traditional human decision-making. In conventional models, human decision-makers are highly susceptible to interference factors such as prior experience (Kahneman & Tversky, 1972) and emotions (Lerner et al., 2015), leading to decision errors and adverse consequences. However, with the vigorous development of big data, deep learning, and neural networks, algorithms have emerged as new decision-making agents. Algorithmic decision-making possesses powerful capabilities for processing vast information (Blair & Saffidine, 2019), rapid decision speed (Bonnefon et al., 2016), objectivity and impartiality (Andrews et al., 2006), and immunity to fatigue and emotional interference (Barnes et al., 2015). Consequently, algorithms are widely applied in providing recommendations, judgments, and predictions, forming the concept of algorithmic decision-making (Burton et al., 2020). This study focuses on attitudes toward algorithm use in the workplace and further explores potential underlying reasons and boundary conditions.

1.1 Algorithm Aversion in the Workplace

Despite algorithmic decision-making's numerous advantages and revolutionary improvements to social efficiency, people have not readily embraced this emerging technology due to its powerful computing capabilities and accurate predictions. Instead, they hold a pervasive pessimistic attitude (Jussupow et al., 2020; Liu et al., 2023; Mahmud et al., 2022). Some entrepreneurs, based on practical business experience, consider automation's rise humanity's "greatest existential threat" (McFarland, 2014). Tesla founder Elon Musk suggests AI is summoning a "demon" (Lemieux, 2017). Former Google CEO Eric Schmidt expresses concern that AI poses an "existential risk" to humanity (Kharpal, 2023). Philosophers harbor metaphysical concerns about algorithmic decision-making's rise, believing machine substitution for human decision-making could lead to outright catastrophe (Bostrom, 2014). This catastrophe might be existential—once AI develops self-consciousness, humanity would have created a technological "god" capable of destroying human history and negating human existence (赵汀阳, 2018; 2019). For the general public, despite algorithms typically outperforming humans in decision-making tasks, people still prefer human decisions over algorithmic recommendations, a tendency termed "algorithm aversion" (Dietvorst et al., 2015). This aversion represents a cognitive bias toward algorithms, manifesting comprehensively in cognition, emotion, and behavior (张语嫣等, 2022): negative attitudes cognitively (e.g., Prahl & van Swol, 2017), disgust emotionally (e.g., Lee, 2018), and rejection behaviorally (e.g., Filiz et al., 2021). Algorithm aversion also appears in art judgment (Millet et al., 2023), poetry aesthetics (Hitsuwari et al., 2023), medical decisions (Longoni et al., 2019), war and judicial decisions (Bigman & Gray, 2018), and daily movie and book recommendations (Longoni & Cian, 2022; Yeomans et al., 2019).

In the workplace, algorithms are believed capable of eliminating unconscious human bias (Cheng & Hackett, 2021), potentially making more equitable, interference-free decisions. Research finds that under normal workloads, algorithmic decisions outperform human manager decisions in over 80% of scenarios (Yu et al., 2017). Yet despite this advantage, people still view algorithmic management as a "tyranny," akin to "clock tyranny" that exploits and oppresses employees (Lehdonvirta, 2018). Increasingly automated jobs deprive people of control over their work (Holford, 2022), reducing autonomy and responsibility (Goods et al., 2019; Langer et al., 2021). Moreover, algorithmic management reduces well-being by diminishing autonomy and expectations of workplace rewards (Kinowska & Sienkiewicz, 2022). Regarding promotion, dismissal, and bonus allocation, people perceive algorithmic decisions as merely quantitative and decontextualized, unable to consider qualitative and environmental factors for comprehensive decisions, making them seem unfair compared to human decisions, even reducing organizational commitment (Newman et al., 2020). For recruitment, applicants trust human managers more than algorithms for resume screening and interviews because human selection is perceived as more controllable and fair (Langer et al., 2019). This phenomenon also appears when recruiters promise contracts: people more strongly reject algorithmic promises of growth and harmonious atmosphere compared to human recruiters offering substantial salaries and bonuses (Tomprou & Lee, 2022). Meanwhile, employees experience six burdens from automated HR algorithms—emotional, psychological, privacy, and social—which are alleviated when human decision-makers participate and provide oversight (Park et al., 2021).

Therefore, this study proposes Hypothesis H1: Compared to human HR managers (hereinafter referred to as HR), people tend to cognitively reject, emotionally dislike, and behaviorally avoid decisions made by algorithmic HR in workplace contexts.

1.2 Perceived Transparency

What factors cause different reactions to decisions made by different agents? Transparency is crucial. Algorithmic transparency means users can understand what an algorithmic system is doing and why (IEEE, 2017). It emphasizes that for general users, intelligent systems' decision processes and mechanisms must be knowable and explainable, with predictable and accountable consequences—decision processes must be easily understandable (Nefdt, 2020). Transparency matters because reactions to decision outcomes depend partly on information provided and its explainability (Dodge et al., 2019), and people want decisions they can understand (McNee et al., 2006; Herlocker et al., 2004). However, algorithmic systems are cognitive "black boxes" (Burrell, 2016; Nicholson Price, 2018)—people perceive them as mysterious, complex systems where only inputs and outputs are observable, not the processing mechanisms (Pasquale, 2016). This black box nature may stem from unpredictable operations or lack of relevant knowledge (Kroll et al., 2017).

This opacity hinders algorithm acceptance. Without autonomous explanation capabilities, algorithms may generate perceptions of unfairness and anxiety (Acikgoz et al., 2020; Langer et al., 2019). Research finds algorithmic transparency positively predicts satisfaction—people are more satisfied with algorithmic services when they perceive the decision process as transparent (Shin & Park, 2019; Shin, 2020). Furthermore, understanding system decision principles increases trust (Lee & Boynton, 2017). This suggests opening the "black box" may change inherent negative reactions.

Conversely, people exhibit overconfidence in understanding human decisions (Chen et al., 2023; Moore & Healy, 2008), overestimating their comprehension of others' mental processes. This bias may arise from intuitive perspective-taking (Nisbett & Wilson, 1977), though people actually rely more on intuition (Dane et al., 2012) and heuristics (Kahneman, 2003) when understanding others' decisions. Even in the specialized medical field, laypeople believe they understand how human doctors diagnose cancer (Cadario et al., 2021).

Thus, people perceive different transparency levels for different decision agents, and understanding decision processes greatly influences attitudes. Accordingly, this study proposes Hypothesis H2: Perceived transparency mediates the effect of workplace decision-making agents (human HR vs. algorithmic HR) on decision reactions.

1.3 Anthropomorphism

Anthropomorphism is the phenomenon of attributing human properties, characteristics, intentionality, and mental states to non-human entities (Epley et al., 2007; 许丽颖等, 2017). Viewing an entity as possessing mental capacities constitutes a key component of anthropomorphism (Waytz et al., 2010). The comprehensibility of thought processes is one such mental capacity; when we believe an entity's thinking is understandable and accessible, we essentially endow it with human-like mentality (Gray et al., 2007; Gray et al., 2012). The process of granting non-human algorithms human-unique perceived transparency naturally equals anthropomorphism. Thus, transparency constitutes a crucial difference between human and algorithmic resources, excluding the latter from entities with mental capacities. We can predict that anthropomorphizing algorithms through appearance (De Visser et al., 2016), voice (Adam et al., 2021), movement (Fraune, 2020), or mental capacities (Moussawi & Koufaris, 2019) may bridge the perceived gap between algorithms and humans, potentially yielding more positive attitudes. Previous research finds anthropomorphizing machines, algorithms, and AI reduces aversion. For instance, superficial anthropomorphism of autonomous vehicles (giving them human-like appearance or names) effectively increases trust (Wu et al., 2023). Compared to AI assistants with mechanical voices and cold responses, those with human-like appearance, anthropomorphic names, human-like voices, and warm emotional responses are perceived as psychologically closer and receive higher satisfaction ratings (Li & Sung, 2021). Even when robots make poor decisions, anthropomorphic robots are perceived as having higher experiential capacity, reducing negative evaluations of their mistakes and increasing forgiveness (Yam et al., 2021). Anthropomorphism thus prompts people to view algorithms as mindful agents, eliciting more human-like evaluations and reactions that influence algorithm aversion. For inherently black-box algorithms, anthropomorphism means granting them human characteristics, narrowing the perceived gap between algorithmic and human decision-makers, potentially making attitudes toward algorithmic decisions more similar to those toward human decisions, thereby reducing algorithm aversion (许丽颖等, 2022). In the workplace, algorithmic decisions are originally ambiguous, rejected, and unauthorized; if endowed with human traits to create more human-like anthropomorphic algorithms, people should react more positively when accepting their decisions.

Thus, this study proposes Hypothesis H3: Algorithm anthropomorphism moderates the effect of workplace decision-making agents (human HR vs. algorithmic HR) on decision attitudes. This means people more readily accept anthropomorphized algorithmic HR compared to non-anthropomorphized algorithms.

1.4 Overview of Studies

Based on the above, this study conducted four sequential experiments to examine whether attitudes toward algorithmic HR decisions differ from those toward human HR decisions in workplace contexts, and to explore psychological mechanisms and boundary conditions. The fundamental hypothesis is that compared to traditional human HR, people exhibit greater aversion toward algorithmic HR decisions, specifically showing lower cognitive permissibility, emotional dislike, and behavioral unwillingness to reuse algorithms for similar decisions. This effect is mediated by perceived transparency and moderated by anthropomorphism. Four scenario experiments tested these hypotheses across diverse HR decisions: recruitment and hiring (Experiment 1), year-end bonus allocation (Experiment 2), resume screening (Experiment 3), and performance appraisal (Experiment 4), with nationally representative samples and university student samples.

Specifically: Experiment 1 explored the main hypothesis—that people exhibit aversive reactions to algorithmic HR decisions in the workplace. Experiment 2 investigated the underlying mechanism, testing perceived transparency's mediating role in the effect of decision agents on decision attitudes. Experiment 3 manipulated algorithmic decision transparency to further examine whether perceived transparency causes aversion to algorithmic HR decisions. Experiment 4 explored boundary conditions, investigating anthropomorphism's moderating effect on decision agents' influence on decision attitudes.

Experiment 1

Experiment 1 preliminarily examined whether attitudes toward algorithmic HR decisions are more aversive than toward human HR decisions. Using a scenario experiment, we randomly assigned participants to human or algorithm conditions, where they read about decisions made by human HR or algorithmic HR and reported decision permissibility, liking, and utilization willingness, comparing attitudinal differences to verify workplace algorithm aversion.

2.1.1 Participants

We first used G*Power 3.1 software (Faul et al., 2009) to calculate required sample size. For independent samples t-tests with $\alpha = 0.05$ and medium effect size (d = 0.5), at least 172 participants were needed for 90% statistical power. To ensure sufficient data, we distributed 416 questionnaires via Wenjuanxing platform at a university. After excluding 72 participants who failed attention checks, 344 questionnaires remained (82.69% recovery rate). Excluding 41 who failed manipulation checks yielded 303 valid participants (72.84% total recovery rate), including 125 males (41.3%) and 178 females (58.7%), with mean age M = 20.80 years, SD = 1.61 years. Participants were randomly assigned to human (n = 163) or algorithm (n = 140) groups. All participants volunteered with informed consent and received compensation after completing attention checks.

2.1.2 Design and Procedure

Experiment 1 used a single-factor, two-level between-subjects design with human and algorithm groups. All participants were randomly assigned to read scenario materials about human HR or algorithmic HR making decisions.

The material read (with decision agent variations in bold): "In last year's autumn recruitment, Jiuri Group's HR Director Wang Peng was responsible for/Jiuri used the '330F' recruitment algorithm to evaluate resumes from fresh graduates. Throughout the recruitment process, Wang Peng's team/the '330F' algorithm independently completed, fully decided, and took responsibility for recruitment outcomes. Wang Peng's team is an industry veteran HR evaluation team/The '330F' recruitment algorithm is a new recruitment algorithm capable of carefully reviewing applicants' resumes and backgrounds, accurately predicting future employees who will meet job requirements and fit corporate culture, and identifying the most suitable candidates from massive applicant pools. According to reports, 50 fresh graduates participated in interviews, but after Wang Peng's team/the '330F' algorithm's decision, only 5 were hired (meeting Jiuri Company's required 10% acceptance rate). The list follows: Wang Shengyi, Huang Jiaoxu, Chen Zhenjiang, Xie Wenxiang, Xu Hanyun. Wang Peng's team/the '330F' algorithm immediately uploaded and published the hiring list on the recruitment website without further explanation (including evaluation score rankings, detailed criteria, etc.)."

Adapted from Newman et al. (2020), materials differed only in decision agent identity (human HR team led by Wang Peng vs. recruitment algorithm "330F"), with no other changes. To ensure careful reading and comprehension, participants answered attention check questions after reading: human group participants answered, "Was Jiuri Group's recruitment decision agent (the entity participating in, leading, and executing decisions, the soul and core of the decision system) the veteran HR Director Wang Peng and his team?" Algorithm group participants answered, "Was Jiuri Group's recruitment decision agent the new recruitment algorithm '330F'?" Participants responded "yes" or "no"; those answering "no" failed the attention check.

After reading materials and completing attention checks, participants responded to measures of decision permissibility, liking, and utilization willingness. Permissibility measurement adapted from Bigman and Gray (2018) included three items (with algorithm condition items in parentheses): "Was Wang Peng's (algorithm '330F's) decision appropriate?" (1–5 scale, 1 = completely inappropriate, 5 = completely appropriate); "Should Wang Peng (algorithm '330F') be allowed to make these decisions?" (1–5 scale, 1 = should not be allowed at all, 5 = should be completely allowed); "Should Wang Peng (algorithm '330F') be prohibited from making these decisions?" (1–5 scale, reverse-scored, 1 = should not be prohibited at all, 5 = should be completely prohibited). Items used a 5-point Likert scale; the third item was reverse-scored. Higher total scores indicated greater permissibility. Cronbach's $\alpha = 0.76$ for permissibility in Experiment 1.

Participants also responded to two items measuring decision liking, adapted from Jago (2019), using a 7-point Likert scale: "How much do you approve of Wang Peng's (algorithm '330F's) decision?" (1–7, 1 = completely disapprove, 7 = completely approve); "How much do you like Wang Peng's (algorithm '330F's) decision?" (1–7, 1 = completely dislike, 7 = completely like). Higher scores indicated greater liking. The correlation between the two items was $r = 0.701$, $p < 0.001$.

Additionally, participants reported utilization willingness: "If you were a company leader, to what extent would you hire Wang Peng and his team (intelligent algorithm 330F) to complete the above decision-making task?" Adapted from Cadario et al. (2021), this used a 7-point Likert scale (1 = completely unlikely, 7 = extremely likely), with higher scores indicating stronger utilization willingness.

Finally, participants reported demographic information (gender and age). Two random attention check items (e.g., "Please select 1 for this question") appeared during questionnaire completion to screen for careless responding.

2.2 Results

Independent samples t-tests showed human group participants' permissibility ratings ($M = 10.60$, $SD = 1.84$) were significantly higher than algorithm group ratings ($M = 9.25$, $SD = 2.38$), $t(301) = 5.56$, $p < 0.001$, Cohen's $d = 0.63$. Human group liking ratings ($M = 7.89$, $SD = 2.10$) were also significantly higher than algorithm group ratings ($M = 7.30$, $SD = 2.10$), $t(301) = 2.08$, $p = 0.038$, Cohen's $d = 0.24$. For utilization willingness, human group ratings for human HR ($M = 4.04$, $SD = 1.42$) were higher than algorithm group ratings for algorithmic HR ($M = 3.74$, $SD = 1.50$), $t(301) = 1.75$, $p = 0.081$, Cohen's $d = 0.21$, a non-significant difference but $p < 0.1$ with a small effect size (Cohen, 1969). Results appear in Figure 1 [Figure 1: see original paper].

A multivariate analysis of variance (MANOVA) with decision agent (human vs. algorithm) as the independent variable and permissibility, liking, and utilization willingness as dependent variables showed a significant main effect of decision agent, Wilks' $\lambda = 0.896$, $F(3,299) = 11.615$, $p < 0.001$, $^{2}p = 0.104$.

2.3 Discussion

Experiment 1 preliminarily verified that algorithmic decisions in recruitment are less permissible, less liked, and less likely to be reused compared to identical human decisions, supporting workplace algorithm aversion. However, Experiment 1 only covered one aspect of HR management—recruitment—and did not further explore mechanisms behind algorithm aversion. Therefore, Experiment 2 used year-end bonus allocation to enhance robustness and further explore perceived transparency's mediating role.

Experiment 2

Building on Experiment 1, Experiment 2 enriched workplace HR decision contexts by examining year-end bonus allocation while further exploring perceived transparency's potential mediating role.

3.1.1 Participants

Using Monte Carlo simulations and referencing Experiment 1's effect size (Cohen's $d = 0.63$), with 90% statistical power and a narrow corridor of stability width $w = 0.1$, the minimum required sample size was 150 (Schönbrodt & Perugini, 2013). We recruited participants via Credamo platform, real-time excluding those failing attention checks through rolling recruitment, ultimately obtaining 179 valid participants. Mean age was $M = 30.85$ years, $SD = 7.09$ years; 115 were female (64.2%) and 64 male (35.8%). Participants were randomly assigned to human ($n = 90$) or algorithm ($n = 89$) groups. All participants read experimental instructions and provided informed consent; those passing attention checks received compensation.

3.1.2 Design and Procedure

Like Experiment 1, Experiment 2 used a single-factor, two-level between-subjects design. Participants first read materials about HR decision-makers (human or algorithm) allocating year-end bonuses. Adapted from Newman et al. (2020), materials differed only in decision agent identity and related wording (differences marked in bold).

Human group read: "Xinyue Company just completed its year-end bonus distribution. To determine each employee's bonus amount, Xinyue Company relied on its veteran HR manager Zhang Yun's team, which considered various factors. After Zhang Yun's team conducted a series of deliberations, the specific bonus distribution model was determined."

Algorithm group read: "Xinyue Company just completed its year-end bonus distribution. To determine each employee's bonus amount, Xinyue Company relied on a reliable HR intelligent algorithm 'RTC,' which considered various factors. After the intelligent algorithm 'RTC' conducted a series of calculations, the specific bonus distribution model was determined."

After reading, participants reported permissibility, liking, and utilization willingness using the same measures as Experiment 1. Permissibility measurement showed internal consistency reliability Cronbach's $\alpha = 0.80$; the two liking items correlated significantly (r = 0.660, p < 0.001).

Next, we measured perceived transparency of the decision agent's activity using one item (adapted from Cadario et al., 2021): "To what extent do you understand how Zhang Yun (algorithm 'RTC') made the above decision?" Rated on a 7-point Likert scale (1 = completely cannot understand, 7 = completely can understand), with higher scores indicating greater perceived transparency.

Finally, participants answered the same attention check questions as Experiment 1, interspersed among measurement items for screening, and reported demographic information (gender and age).

3.2 Results

3.2.1 Effects of Decision Agent on Permissibility, Liking, and Utilization Willingness Independent samples t-tests revealed significant differences between human and algorithm groups across all three measures. For permissibility, human group ratings (M = 12.64, SD = 1.89) were significantly higher than algorithm group ratings (M = 11.76, SD = 2.38), t(177) = 2.74, p = 0.007, Cohen's d = 0.41. For liking, human group ratings (M = 11.13, SD = 1.98) were significantly higher than algorithm group ratings (M = 10.45, SD = 2.39), t(177) = 2.09, p = 0.038, Cohen's d = 0.31. For utilization willingness, human group ratings (M = 5.89, SD = 0.88) were significantly higher than algorithm group ratings (M = 5.27, SD = 1.36), t(177) = 3.62, p < 0.001, Cohen's d = 0.54.

A MANOVA with decision agent (human vs. algorithm) as independent variable, gender and age as covariates, and permissibility, liking, and utilization willingness as dependent variables showed a significant main effect of decision agent, Wilks' $\lambda = 0.925$, F(3,175) = 4.730, p = 0.003, ^{2}p = 0.075.

3.2.2 Mediating Effect of Perceived Transparency Independent samples t-tests showed decision agent significantly affected perceived transparency, with human group ratings (M = 5.44, SD = 1.15) significantly higher than algorithm group ratings (M = 5.02, SD = 1.31), t(177) = 2.29, p = 0.023, Cohen's d = 0.34. To further explore the psychological mechanism, we used Hayes's (2013) SPSS PROCESS macro (Model 4). With decision agent as independent variable (human = 1, algorithm = 2), perceived transparency as mediator, and permissibility, liking, and utilization willingness as separate dependent variables, we conducted mediation tests with 5,000 Bootstrap samples and bias-corrected 95% confidence intervals.

Results showed (see table below) that for permissibility, the indirect effect of perceived transparency was –0.50, 95% CI [–0.96, –0.08], excluding 0, indicating significant mediation. After controlling for the mediator, the direct effect of decision agent on permissibility was –0.38, 95% CI [–0.85, 0.09], including 0, indicating the direct effect became non-significant, confirming full mediation. For liking, the indirect effect was –0.56, 95% CI [–1.09, –0.09], excluding 0, with direct effect –0.12, 95% CI [–0.55, 0.32], including 0, confirming full mediation. For utilization willingness, the indirect effect was –0.25, 95% CI [–0.51, –0.04], excluding 0, with direct effect –0.37, 95% CI [–0.63, –0.10], excluding 0, indicating partial mediation.

Table 1 Bootstrap analysis and effect values for mediation significance testing in Experiment 2

Dependent Variable	Indirect Effect	95% Indirect Effect LLCI	95% Indirect Effect ULCI	95% Direct Effect LLCI	95% Direct Effect ULCI
Permissibility	–0.50	–0.96	–0.08	–0.85	0.09
Liking	–0.56	–1.09	–0.09	–0.55	0.32
Utilization Willingness	–0.25	–0.51	–0.04	–0.63	–0.10

To further verify mediation robustness, we also conducted traditional stepwise regression analysis (温忠麟等, 2004), with results shown in Figure 2 [Figure 2: see original paper].

Figure 2 Mediating role of perceived transparency
Note: $p < 0.05$, $\boldsymbol{p < 0.01}$, $p < 0.001$

3.3 Discussion

Consistent with Experiment 1, Experiment 2 replicated workplace algorithm aversion—people prefer human decisions, finding algorithmic decisions less permissible, less likable, and less reusable, even when decision content is identical. Moreover, we discovered perceived transparency's mediating role: people avert algorithmic HR decisions because they perceive them as more opaque and less understandable. Experiments 1 and 2 robustly demonstrate workplace algorithm aversion and provide initial exploration of perceived transparency's mediation. To deepen understanding of this mechanism, Experiment 3 manipulates perceived transparency of algorithmic decisions, hypothesizing that higher perceived transparency will increase permissibility, liking, and willingness to utilize algorithmic HR decisions.

Experiment 3: Manipulating Perceived Transparency

Experiment 3 compared preferences for algorithmic decisions with different transparency levels, focusing on resume screening—a critical HR function—to enhance robustness and further discuss perceived transparency's mediating role.

4.1.1 Participants

Using G*Power 3.1 software (Faul et al., 2009) for independent samples t-tests with medium effect size d = 0.5, $\alpha = 0.05$, at least 172 participants were needed for 90% power. Via Credamo platform, we real-time excluded participants failing attention checks through rolling recruitment, ultimately obtaining 180 valid participants. Mean age was 31.17 ± 7.41 years; 115 were female (63.9%) and 65 male (36.1%). Participants were randomly assigned to low-transparency (n = 88) or high-transparency (n = 92) groups. All participants read experimental instructions and provided informed consent; valid participants received compensation.

4.1.2 Design and Procedure

This single-factor, two-level between-subjects design randomly assigned participants to high- or low-transparency algorithm groups. Both groups first read scenario materials about algorithmic resume screening adapted from Newman et al. (2020).

Low-transparency group read: "In last year's autumn recruitment, Bingquan Company used an intelligent recruitment algorithm 'IRA-N' to evaluate graduates' resumes. The intelligent recruitment algorithm 'IRA-N' conducts intelligent resume screening based on backpropagation neural networks. Bingquan Company released a unified electronic resume template created by the intelligent recruitment algorithm through its official website, requiring all applicants to use this template. The entire resume evaluation process was independently completed, fully decided, and responsibility-assumed by intelligent algorithm 'IRA-N.' Resume screening evaluation is fundamental and important for interview work."

High-transparency group read: "In last year's autumn recruitment, Bingquan Company used an intelligent recruitment algorithm 'IRA-N' to evaluate graduates' resumes. The intelligent recruitment algorithm 'IRA-N' conducts intelligent resume screening based on backpropagation neural networks. Intelligent algorithm 'IRA-N' divides applicants' comprehensive qualities into five modules: basic indicators (including education level, foreign language proficiency, awards, etc.), personality traits (responsibility, confidence, tolerance, etc.), motivation (achievement motivation, learning motivation, big-picture thinking, etc.), knowledge and skills (disciplinary expertise, research achievements, etc.), and competency (job experience, teamwork, innovation, tool utilization ability, etc.). Bingquan Company released a unified electronic resume template created by the intelligent recruitment algorithm through its official website, requiring all applicants to use this template. Intelligent recruitment algorithm 'IRA-N' has undergone deep learning through big data BP neural networks, passed internal testing, and is nearly mature, capable of independently and perfectly completing electronic resume screening. The entire resume evaluation process was independently completed, fully decided, and responsibility-assumed by intelligent algorithm 'IRA-N.' Resume screening evaluation is fundamental and important for interview work."

Both groups read about algorithm 'IRA-N' facing the same task with equal capabilities and unknown decision outcomes, but the high-transparency group received detailed information about the algorithm's decision criteria (marked in bold). This manipulation operationalized transparency levels—high transparency provided additional explanation of algorithmic decision processes and computational capabilities.

Subsequently, participants completed the perceived transparency item from Experiment 2 to verify manipulation success, followed by permissibility, liking, and utilization willingness measures. In this experiment, permissibility showed internal consistency reliability Cronbach's $\alpha = 0.896$; the two liking items correlated significantly ($r = 0.854$, $p < 0.001$).

Finally, all participants answered the same attention check questions as previous experiments and reported demographic information (gender and age).

4.2 Results

Independent samples t-tests showed high-transparency group perceived transparency scores ($M = 5.68$, $SD = 1.19$) were significantly higher than low-transparency group scores ($M = 4.84$, $SD = 1.68$), $t(178) = 3.90$, $p < 0.001$, Cohen's $d = 0.58$, confirming successful manipulation.

Independent samples t-tests revealed high-transparency group permissibility ratings ($M = 12.25$, $SD = 2.53$) were significantly higher than low-transparency group ratings ($M = 10.66$, $SD = 3.42$), $t(178) = 3.56$, $p < 0.001$, Cohen's $d = 0.53$. High-transparency group liking ratings ($M = 10.83$, $SD = 2.64$) were significantly higher than low-transparency group ratings ($M = 9.05$, $SD = 3.10$), $t(178) = 4.16$, $p < 0.001$, Cohen's $d = 0.62$. High-transparency group utilization willingness ratings ($M = 5.57$, $SD = 1.33$) were significantly higher than low-transparency group ratings ($M = 4.68$, $SD = 1.66$), $t(178) = 3.94$, $p < 0.001$, Cohen's $d = 0.59$. Results appear in Figure 3 [Figure 3: see original paper].

A MANOVA with transparency as independent variable, gender and age as covariates, and permissibility, liking, and utilization willingness as dependent variables showed a significant main effect of transparency, Wilks' $\lambda = 0.907$, $F(3,176) = 6.035$, $p = 0.001$, ${}^2p = 0.093$.

4.3 Discussion

Experiment 3 directly manipulated perceived transparency of algorithmic decisions, finding that increasing perceived transparency significantly improved permissibility, liking, and utilization willingness toward algorithmic decisions, further validating perceived transparency's mediating role in workplace algorithm aversion. Since people avert algorithms because they perceive them as black boxes, does anthropomorphizing algorithms—making non-human algorithms more human-like—moderate this aversion? To address this and explore boundary conditions, Experiment 4 tested anthropomorphism's potential moderating effect.

Experiment 4: The Moderating Role of Algorithm Anthropomorphism

To examine boundary conditions for workplace algorithm aversion demonstrated in previous experiments, Experiment 4 investigated anthropomorphism's moderating effect. Additionally, focusing on performance appraisal enriched the scope of HR work covered.

5.1.1 Participants

For Experiment 4's single-factor, three-level between-subjects design, G*Power 3.1 software (Faul et al., 2009) calculated a minimum required sample of 207 participants for 90% power at $\alpha = 0.05$ with medium effect size (f = 0.25). To ensure robust data, we recruited participants via Credamo and PowerCX platforms, randomly assigning them to non-anthropomorphic algorithm, anthropomorphic algorithm, or human groups. Platforms real-time excluded participants failing attention checks, leaving 549 valid participants (327 female, 59.6%). Mean age was M = 31.4 years, SD = 7.88 years. The non-anthropomorphic algorithm group had 181 participants, the anthropomorphic algorithm group had 185, and the human group had 183. All participants read experimental instructions and provided informed consent; valid participants received compensation.

5.1.2 Design and Procedure

Experiment 4 used a single-factor, three-level between-subjects design with human, anthropomorphic algorithm, and non-anthropomorphic algorithm groups. Participants first read scenario materials about different decision agents conducting performance appraisals.

Human group read: "Hello, I am Huang Fei, HR manager of Chuhe Group. I am responsible for this employee performance appraisal. I will evaluate each employee's performance data and assess employees' 5-minute self-statement videos. After evaluating performance data and videos, I will make final decisions on this year's employee performance appraisals, independently completing and fully assuming responsibility. These results may affect employees' salaries, bonuses, promotion eligibility, and in some cases, termination."

Anthropomorphic algorithm group read: "Hi there! I'm Archie, your intelligent HR assistant. I'm responsible for this employee performance appraisal. I will evaluate each employee's performance data and assess employees' 5-minute self-statement videos. After evaluating performance data and videos, I will make final decisions on this year's employee performance appraisals, independently completing and fully assuming responsibility. These results may affect employees' salaries, bonuses, promotion eligibility, and in some cases, termination."

Non-anthropomorphic algorithm group read: " 'HRA' is a new HR intelligent algorithm. This employee performance appraisal is conducted by algorithm 'HRA.' Algorithm 'HRA' will evaluate each employee's performance data and assess employees' 5-minute self-statement videos. After evaluating performance data and videos, HR intelligent algorithm 'HRA' will make final decisions on this year's employee performance appraisals, independently completing and fully assuming responsibility. These results may affect employees' salaries, bonuses, promotion eligibility, and in some cases, termination."

As an anthropomorphism manipulation, the anthropomorphic and non-anthropomorphic algorithm groups differed in narrative perspective: the anthropomorphic group used first-person subjective statements, while the non-anthropomorphic group used third-person objective statements. This manipulation referenced established paradigms (Hur et al., 2015; May & Monga, 2014) where giving non-human agents a human name and first-person description effectively increases anthropomorphism. Otherwise, algorithm descriptions were identical. To verify manipulation effectiveness, anthropomorphic and non-anthropomorphic algorithm group participants rated the algorithm's anthropomorphism level before other questions: "To what extent does HR intelligent assistant 'Archie'/intelligent algorithm 'HRA' remind you of human characteristics?" (7-point Likert scale, 1 = not at all, 7 = very much), adapted from Hur et al. (2015).

After reading materials and completing manipulation checks, all three groups reported permissibility, liking, and utilization willingness toward the decision agent's decisions using the same measures as previous experiments.

Finally, participants answered attention check questions as in previous experiments and reported demographic information: gender, age, algorithm familiarity ("How familiar are you with AI algorithms?" 1 = completely unfamiliar, 7 = very familiar), and algorithm knowledge ("How much do you know about AI algorithms?" 1 = know nothing, 7 = know very much), both adapted from Leo and Huh (2020).

5.2 Results

Independent samples t-tests showed anthropomorphic algorithm anthropomorphism ratings (M = 5.46, SD = 1.23) were significantly higher than non-anthropomorphic algorithm ratings (M = 4.87, SD = 1.45), t(364) = –4.32, $p < 0.001$, Cohen's d = 0.44, confirming successful anthropomorphism manipulation.

One-way ANOVA on permissibility revealed a significant main effect of decision agent, F(2,546) = 3.15, p = 0.044, ^{2}p = 0.11. ANOVA on liking showed a non-significant main effect, F(2,546) = 2.39, p = 0.093, ^{2}p = 0.09. ANOVA on utilization willingness also showed a non-significant main effect, F(2,546) = 0.40, p = 0.668.

Planned contrast analyses showed that for permissibility, human group scores (M = 10.54, SD = 2.81, p = 0.021, Cohen's d = 0.25) were significantly higher than non-anthropomorphic algorithm scores (M = 9.81, SD = 3.11), while anthropomorphic algorithm scores (M = 10.42, SD = 2.94) were numerically higher than non-anthropomorphic algorithm scores (M = 9.81, SD = 3.11, p = 0.055, Cohen's d = 0.20), a non-significant difference but approaching statistical significance with a small effect size (Cohen, 1969). Human and anthropomorphic algorithm groups did not differ, p = 0.705. For liking, human group scores (M = 9.13, SD = 3.16) were higher than non-anthropomorphic algorithm scores (M = 8.68, SD = 3.43) but not statistically significant, p = 0.192, showing no statistically meaningful algorithm aversion. However, anthropomorphic algorithm scores (M = 9.40, SD = 2.94, p = 0.032, Cohen's d = 0.23) were significantly higher than non-anthropomorphic algorithm scores (M = 8.68, SD = 3.43), indicating anthropomorphism effectively increased liking of algorithmic decisions. Results appear in Figure 4 [Figure 4: see original paper].

A MANOVA with decision agent (human vs. anthropomorphic algorithm vs. non-anthropomorphic algorithm) as independent variable, algorithm familiarity and knowledge as covariates, and permissibility, liking, and utilization willingness as dependent variables showed a significant main effect of decision agent, Wilks' $\lambda = 0.965$, F(6,1084) = 3.251, p = 0.004, ^{2}p = 0.018.

With algorithm familiarity and knowledge as covariates, ANOVA on permissibility showed significant effects of familiarity, F(1,544) = 4.16, p = 0.042, ^{2}p = 0.008, and knowledge, F(1,544) = 5.64, p = 0.018, ^{2}p = 0.010, but a non-significant group effect that approached significance with small effect size, F(2,544) = 2.98, p = 0.052, ^{2}p = 0.011. With the same covariates, ANOVA on liking showed significant effects of familiarity, F(1,544) = 9.45, p = 0.002, ^{2}p = 0.017, and knowledge, F(1,544) = 9.67, p = 0.002, ^{2}p = 0.017, with group effects also approaching significance, F(2,544) = 2.62, p = 0.074, ^{2}p = 0.01. Thus, prior familiarity and knowledge about algorithms strongly influence attitudes toward algorithmic management, consistent with previous research (Ireland, 2020; Komatsu, 2016). This suggests that greater familiarity and understanding lead to greater acceptance of algorithmic workplace decisions and more legitimacy granted to algorithmic decision-making. In other words, improving public algorithm awareness and literacy may help reduce aversion.

In summary, algorithm anthropomorphism shows some moderating effect on decision agent influences on permissibility and liking. Specifically, anthropomorphizing algorithms significantly increased liking of algorithmic decisions, expressing less algorithm aversion. However, anthropomorphism did not effectively increase utilization willingness toward algorithmic decision-makers.

General Discussion

This study explored whether attitudes differ toward human versus algorithmic decision-makers in workplace applications and examined underlying mechanisms and boundary conditions. Through four sequential experiments, we found that algorithmic decision-makers elicit harsher evaluations than human decision-makers because algorithmic decisions (vs. human decisions) are perceived as more opaque, with this difference moderated by anthropomorphism. Specifically, by presenting participants with identical HR decisions made by humans or algorithms and measuring their attitudes, we found people found algorithmic decisions less permissible, less likable, and less reusable, showing rejection across cognitive, emotional, and behavioral dimensions with considerable robustness (Experiments 1–4). By measuring perceived decision transparency (Experiment 2) and manipulating algorithmic decision explainability (Experiment 3), we further identified perceived transparency as the psychological mechanism causing different attitudes toward decision agents (human vs. algorithm)—algorithmic decisions are perceived as more opaque and less understandable, causing avoidance. By manipulating algorithm anthropomorphism levels (Experiment 4), we found decision agent effects on attitudes were moderated by anthropomorphism—when algorithms were highly anthropomorphized, attitudes became more favorable. This suggests anthropomorphic algorithms are an effective pathway to reduce workplace algorithm aversion. In analyzing control variables, we found individual differences in algorithm familiarity and knowledge positively predicted positive attitudes toward algorithmic management, suggesting that improving public algorithm literacy may also facilitate intelligent management adoption.

The study examined diverse workplace application scenarios: recruitment and hiring (Experiment 1), bonus allocation (Experiment 2), resume screening (Experiment 3), and performance appraisal (Experiment 4), with samples including nationally representative participants (Experiments 2–4) and university students (Experiment 1). This diversity in scenarios and samples ensures robust findings.

6.1 Rejected Algorithmic Management

As Industry 4.0 advances, algorithmic systems empowered by information technology increasingly penetrate all aspects of social life, revolutionizing human-machine relationships from "user-tool" master-slave paradigms to more equal "partner" paradigms, and now emerging "subordinate-leader" relationships that fundamentally transform understanding of human-machine relations (Wesche & Sonderegger, 2019). The classic Computers as Social Actors (CASA) paradigm posits that people treat computers and advanced IT as independent social entities, interacting according to human social rules rather than viewing them as rigid programmed presentations (Nass et al., 1994). The derived "computers as leaders" paradigm emphasizes that automated algorithms will become middle-level managers, responsible for upward and downward communication between senior management and frontline employees (Wesche & Sonderegger, 2019). In this context, research on attitudes toward algorithmic management is highly beneficial. This study found relatively stable rejection and aversion toward algorithm use in workplace decisions, consistent with previous findings that both general public and managers feel uneasy about autonomous algorithmic decisions (Acikgoz et al., 2020; Diab et al., 2011; Fischer & Peterson, 2018; Haesevoets et al., 2021; Nørskov et al., 2020). We found that even when humans and algorithms make identical HR decisions, people still resist algorithmic management more. This conclusion provides new evidence for workplace algorithm aversion and reminds researchers and managers to reconsider relationships between algorithmic leaders and humans, contemplating how to smooth the psychological transition of algorithmic management for ordinary employees and how to truly accept algorithmic HR as an independent organizational entity.

Notably, this study constructed a "three-dimension, two-perspective" dependent variable model to measure workplace algorithm aversion, expanding on previous research's narrow operationalizations. Past studies typically measured single response indicators like trust (Hoff & Bashir, 2015) or purchase intention (Wien & Peluso, 2021). However, this is insufficient, as attitudes can show cognition-behavior inconsistency—people may recognize algorithmic superiority while rejecting recommendation systems (Yeomans et al., 2019). Integrating understanding of human psychology, we argue algorithm aversion must be understood across cognitive, emotional, and behavioral (intention) dimensions (referencing the ABC theory of attitudes, Breckler, 1984). Therefore, we selected permissibility, liking, and utilization willingness to represent these dimensions. Permissibility reflects cognitive legitimacy judgments about decision agents' capability, rights, and qualifications to make decisions. Liking is a positive emotional response and intuitive judgment about external objects (Bartneck et al., 2009), reflecting emotional tendencies. Utilization willingness examines behavioral intentions from a hypothetical manager perspective. Additionally, our dependent variables captured both "employee-employer" workplace perspectives (Cummins, 1998; Gigerenzer & Hug, 1992). Permissibility and liking reflect employee perspectives on algorithms, while utilization willingness asks participants to imagine themselves as company leaders. This perspective difference may explain why algorithm aversion in utilization willingness was not replicated in Experiments 1 and 4, possibly because the predominantly young student sample lacked experience as company leaders or even interactions with them, making it difficult to imagine this role's choice tendencies.

Admittedly, our conclusions demonstrate pervasive workplace algorithm aversion, but this does not mean enterprises should abandon digital transformation. Originally designed to enhance employee welfare and organizational performance

(Benlian et al., 2022), algorithmic technology unfortunately became a super tool that enslaves rather than liberates due to its opaque nature, ultimately inviting suspicion and rejection (Jussupow et al., 2020). However, this does not justify abandoning progress. In optimizing algorithmic management, researchers and managers must consider not only how to design more acceptable, achievement-enhancing algorithmic decision systems (Simth & Shum, 2018) but also how human-machine and human-human relationships change under algorithmic management. In essence, this research sounds an alarm: facing algorithmic technology that drives management innovation, we must correctly examine its double-edged nature, maximize strengths while minimizing weaknesses, and unleash its full potential.

6.2 Paths to Intelligent Management

From a practical standpoint, this study explores key factors for accepting algorithmic management to promote automated, intelligent HR decision-making. Based on four sub-experiments, we propose three viable paths to intelligent management.

First, enhance algorithmic decision transparency and develop Explainable Artificial Intelligence. Experiment 2 found that attitudinal differences toward human versus algorithmic decision agents stem from perceiving algorithmic decisions as more opaque and less understandable. Experiment 3 showed that opening the "black box" and improving explainability effectively ameliorates negative attitudes. In summary, transparency mediates algorithm aversion, suggesting a feasible approach to enhancing intelligent management effectiveness. Generally, algorithmic transparency involves two aspects: (1) the algorithm as decision agent presenting its decision processes and rationales, making unknowable processes knowable, and (2) showing observers the underlying operational rules and logic (Confalonieri et al., 2021; Leichtmann et al., 2023). Therefore, enhancing transparency requires developing Explainable Artificial Intelligence (XAI) that enables users to understand why AI algorithms make decisions, transforming the opaque "black box" into a transparent "glass box" (Rai, 2019).

Second, design anthropomorphic management algorithms. Experiment 4 found that superficial anthropomorphic manipulation—giving algorithms names and human-like expression styles—effectively improved aversion attitudes. Compared to mechanical-named, third-person algorithms, people more readily permit and like anthropomorphic algorithms with human names and first-person communication styles for performance appraisals. This aligns with previous research on anthropomorphism's positive effects (Han, 2021; Natarajan & Gombolay, 2020; Yuan & Dennis, 2019). Based on these findings, we advocate that algorithm management system designers adopt anthropomorphic forms to "warm up" cold algorithms and increase acceptance for decision-making roles.

Third, improve public algorithm literacy. Experiment 4 found that control variables measuring algorithm familiarity and knowledge positively influenced dependent variables—more familiar and knowledgeable participants held more positive attitudes toward algorithmic management. This can be explained by the mere exposure effect: repeated exposure to information increases liking (Zajonc, 1968). This suggests that improving public acceptance of algorithmic management may be achieved by enhancing familiarity and understanding. More broadly, this requires improving algorithmic literacy—users' awareness, knowledge, imagination, strategies, and skills surrounding algorithms (Swart, 2021). Living in an automated, information-based, intelligent society, people need enhanced knowledge to cope with ubiquitous algorithms. Managers should also cultivate employees' algorithm awareness and algorithm identity (i.e., IT identity, Craig et al., 2019) to better promote intelligent modern management.

6.3 Limitations and Future Directions

This study has several limitations. First, although dependent variables covered cognitive, emotional, and behavioral dimensions from both manager-employee perspectives, each dimension was represented by only one indicator. Many other important factors remain for future measurement: trust (Logg et al., 2019), fairness perceptions (Schoeffer et al., 2022), creepiness (Mende et al., 2019), moral blame (Malle et al., 2016), punitive behavior (Lokhorst & van den Hoven, 2011), etc. Additionally, utilization willingness only represents self-reported behavioral intentions to reuse the decision agent, lacking objective, accurate measurement and ecological validity. People may suffer from saying-doing inconsistency, reacting differently in real management scenarios than in experimental contexts. Future research should test these effects in real organizational settings, using field observations to record authentic reactions to different manager types (human vs. various algorithmic forms).

Second, this study only demonstrated the effectiveness of the most superficial anthropomorphism method—giving algorithms human names and anthropomorphic language (e.g., first-person perspective). This is because linguistic expression and name usage are the simplest anthropomorphism methods in practical applications (e.g., Apple's Siri and Xiaomi's Xiao Ai are already widely used) and most readily generalizable. Deeper anthropomorphism through appearance and movement has more complex effects on attitudes, potentially causing counterproductive interference like uncanny valley effects (Laakasuo et al., 2021; Mori, 1970; Mori et al., 2012) or identity threat (Yogeeswaran et al., 2016; Zlotowski et al., 2017). Future research should more thoroughly examine how different types and degrees of anthropomorphism affect workplace algorithm attitudes.

Finally, other mechanisms and boundary conditions may explain workplace algorithm aversion. While we focused on perceived transparency's mediating role, human cognition is complex and multidimensional; other factors may also mediate. For example, reduced free will (许丽颖等, 2022), lower mind perception

(Bigman & Gray, 2018), and uniqueness threat (Ferrari et al., 2016) are potential causes. Future research should more carefully examine these variables, compare their influences comprehensively, and more thoroughly understand the complex mechanisms underlying differential reactions to algorithms versus humans. Similarly, anthropomorphism is not the only boundary condition affecting algorithm acceptance. Explorations of machine acceptance and efforts to reduce algorithm aversion can be understood from three aspects: human traits, machine properties, and human-machine interaction patterns (许丽颖, 喻丰, 2020). Therefore, algorithm user individual differences, different psychological perceptions caused by algorithms and their embodiments (e.g., warmth/competence, Fiske et al., 2002), and human-algorithm collaboration weighting modes may all potentially influence algorithm management acceptance, awaiting future research.

Conclusion

This study yields three main conclusions: First, across various workplace application scenarios, people show lower permissibility, liking, and utilization of algorithmic versus human decisions, demonstrating stable algorithm aversion. Second, the underlying psychological mechanism is that algorithmic (vs. human) decisions are perceived as less understandable and more opaque. Third, the more anthropomorphic the algorithm, the weaker the algorithm aversion.

References

Acikgoz, Y., Davison, K. H., Compagnone, M., & Laske, M. (2020). Justice perceptions of artificial intelligence in selection. *International Journal of Selection and Assessment, 28*(4), 399–416.

Adam, M., Wessel, M., & Benlian, A. (2021). AI–based chatbots in customer service and their effects on user compliance. *Electronic Markets, 31*(2), 427–445.

Ahmed, S., Alshater, M., Ammari, A., & Hammami, H. (2022). Artificial intelligence and machine learning in finance: A bibliometric review. *Research in International Business and Finance, 61*, 101646.

Andrews, D., Bonta, J., & Wormith, J. (2006). The recent past and near future of risk and/or need assessment. *Crime and Delinquency, 52*(1), 7–27.

Badue, C., Guidolini, R., Carneiro, R., Azevedo, P., Cardoso, V., Forechi, A., Jesus, L., Berriel, R., Paixão, T. M., Mutz, F., de Paula Veronese, L., Oliveira-Santos, T., & De Souza, A. (2021). Self–driving cars: A survey. *Expert Systems with Applications, 165*, 113816.

Barnes, C. M., Lucianetti, L., Bhave, D. P., & Christian, M. S. (2015). "You wouldn't like me when I'm sleepy": Leaders' sleep, daily abusive supervision, and work unit engagement. *Academy of Management Journal, 58*(5), 1419–1437.

Bartneck, C., Kulic, D., Croft, E., & Zoghbi, S. (2009). Measurement instruments for the anthropomorphism, animacy, likeability, perceived intelligence, and perceived safety of robots. *International Journal of Social Robotics, 1*(1), 71–81.

Basu, S., Majumdar, B., Mukherjee, K., Munjal, S., & Palaksha, C. (2023). The role of artificial intelligence in HRM: A systematic review and future research direction. *Human Resource Management Review, 33*(1), 100893.

Benlian, A., Wiener, M., Cram, W., Krasnova, H., Maedche, A., Möhlmann, M., Recker, J., & Remus, U. (2022). Algorithmic management: Bright and dark sides, practical implications, and research opportunities. *Business & Information Systems Engineering, 64*(6), 825–839.

Bigman, Y. E., & Gray, K. (2018). People are averse to machines making moral decisions. *Cognition, 181*, 21–34.

Blair, A., & Saffidine, A. (2019). AI surpasses humans at six–player poker. *Science, 365*(6456), 864–865.

Bonnefon, J. F., Shariff, A., & Rahwan, I. (2016). The social dilemma of autonomous vehicles. *Science, 352*(6293), 1573–1576.

Bostrom, N. (2014). *Superintelligence.* New York: Oxford University Press.

Breckler, S. J. (1984). Empirical validation of affect, behavior, and cognition as distinct components of attitude. *Journal of Personality and Social Psychology, 47*(6), 1191–1205.

Brynjolfsson, E., & McAfee, A. (2014). *The second machine age: Work, progress, and prosperity in a time of brilliant technologies.* New York: W.W. Norton & Company.

Burrell, J. (2016). How the machine 'thinks': Understanding opacity in machine learning algorithms. *Big Data & Society, 3*(1), 205395171562251.

Burton, J., Stein, M., & Jensen, T. (2020). A systematic review of algorithm aversion in augmented decision making. *Journal of Behavioral Decision Making, 33*(2), 220–239.

Cadario, R., Longoni, C., & Morewedge, C. K. (2021). Understanding, explaining, and utilizing medical artificial intelligence. *Nature Human Behaviour, 5*(12), 1636–1642.

Chalfin, A., Danieli, O., Hillis, A., Jelveh, Z., Luca, M., Ludwig, J., & Mullainathan, S. (2016). Productivity and selection of human capital with machine learning. *The American Economic Review, 106*(5), 124–127.

Chen, Z., Liu, Y., Meng, J., & Wang, Z. (2023). What's in a face? An experiment on facial information and loan-approval decision. *Management Science, 69*(4), 2263–2283.

Cheng, M., & Hackett, R. (2021). A critical review of algorithms in HRM: Definition, theory, and practice. *Human Resource Management Review, 31*(1), 100698.

Cohen, J. (1969). *Statistical power analysis for the behavioral sciences.* New York: Academic Press.

Confalonieri, R., Coba, L., Wagner, B., & Besold, T. R. (2021). A historical perspective of explainable Artificial Intelligence. *Wiley Interdisciplinary Reviews. Data Mining and Knowledge Discovery, 11*(1), e1391.

Craig, K., Thatcher, J. B., Grover, V. (2019). The IT identity threat: A conceptual definition and operational measure. *Journal of Management Information Systems, 36*(1), 259–288.

Cummins, D. (1998). Social norms and other minds: The evolutionary roots of higher cognition. In D. D. Cummins & C. Allen (Eds), *The evolution of mind* (pp. 30–50). New York: Oxford University Press.

Dane, E., Rockmann, K. W., & Pratt, M. (2012). When should I trust my gut? Linking domain expertise to intuitive decision–making effectiveness. *Organizational Behavior and Human Decision Processes, 119*(2), 187–194.

De Visser, E., Monfort, S., McKendrick, R., Smith, M., McKnight, P., Krueger, F., & Parasuraman, R. (2016). Almost human: Anthropomorphism increases trust resilience in cognitive agents. *Journal of Experimental Psychology: Applied, 22*(3), 331–349.

Diab, D. L., Pui, S. Y., Yankelevich, M., & Highhouse, S. (2011). Lay perceptions of selection decision aids in US and non–US samples. *International Journal of Selection and Assessment, 19*(2), 209–216.

Dietvorst, B. J., Simmons, J. P., & Massey, C. (2015). Algorithm aversion: people erroneously avoid algorithms after seeing them err. *Journal of Experimental Psychology: General, 144*(1), 114–126.

Dodge, J., Liao, Q., Zhang, Y., Bellamy, R., & Dugan, C. (2019, March). Explaining models: An empirical study of how explanations impact fairness judgment. Paper presented at the meeting of 24th ACM International Conference on Intelligent User Interfaces, Marina del Ray, United State.

Duggan, J., Sherman, U., Carbery, R., & McDonnell, A. (2020). Algorithmic management and app–work in the gig economy: A research agenda for employment relations and HRM. *Human Resource Management Journal, 30*(1), 114–132.

Epley, N., Waytz, A., & Cacioppo, J. T. (2007). On seeing human: A three–factor theory of anthropomorphism. *Psychological Review, 114*(4), 864–886.

Faul, F., Erdfelder, E., Buchner, A., & Lang, A-G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses.* Behavior Research Methods, 41*(4), 1149–1160.

Ferrari, F., Paladino, M., & Jetten, J. (2016). Blurring human–machine distinctions: Anthropomorphic appearance in social robots as a threat to human distinctiveness. *International Journal of Social Robotics, 8*(2), 287–302.

Filiz, I., Judek, J., Lorenz, M., & Spiwoks, M. (2021). Reducing algorithm aversion through experience. *Journal of Behavioral and Experimental Finance, 31*, 100524.

Fischer, S., & Peterson, T. (2018). *Was Deutschland über Algorithmen weiß und denkt: Ergebnisse einer repräsentativen Bevölkerungsumfrage* [What Germany knows and think about algorithms: Results of a representative survey]. Gütersloh, Germany: Bertelsmann Stiftung.

Fiske, S. T., Cuddy, A. J. C., Glick, P., Xu, J. (2002). A model of (often mixed) stereotype content: Competence and warmth respectively follow from perceived status and competition. *Journal of Personality and Social Psychology, 82*(6), 878–902.

Fraune, M. R. (2020). Our robots, our team: Robot anthropomorphism moderates group effects in human–robot teams. *Frontiers in Psychology, 11*, Article 1275.

Garg, S., Sinha, S., Kar, A., & Mani, M. (2022). A review of machine learning applications in human resource management. *International Journal of Productivity and Performance Management, 71*(5), 1590–1610.

Gigerenzer, G., & Hug, K. (1992). Domain-specific reasoning: Social contracts, cheating and perspective change. *Cognition, 43*(2), 127–171.

Goods, C., Veen, A., & Barratt, T. (2019). "Is your gig any good?" Analysing job quality in the Australian platform–based food–delivery sector. *Journal of Industrial Relations, 61*(4), 502–527.

Gray, H. M., Gray, K., & Wegner, D. M. (2007). Dimensions of mind perception. *Science, 315*(5812), 619.

Gray, K., Young, L., & Waytz, A. (2012). Mind perception is the essence of morality. *Psychological Inquiry, 23*(2), 101–124.

Haesevoets, T., De Cremer, D., Dierckx, K., & Van Hiel, A. (2021). Human-machine collaboration in managerial decision making. *Computers in Human Behavior, 119*, 106730.

Han, M. (2021). The impact of anthropomorphism on consumers' purchase decision in chatbot commerce. *Journal of Internet Commerce, 20*(1), 46–65.

Hao, K. (2019). AI is sending people to jail—and getting it wrong. *MIT Technology Review*, Retrieved January 21, 2019, from https://www.technologyreview.com/2019/01/21/137783/algorith criminal-justice-ai/

Hao, K. (2020). Doctors are using AI to triage covid–19 patients. The tools may be here to stay. *MIT Technology Review*, Retrieved April 2020, https://www.technologyreview.com/2020/04/23/1000410/ai-triage-covid-19-patients-health-care/

Hays, K. (2022). Facebook contractors learned they lost work with the company through a video call with anonymous representatives who said an 'algorithm' chose random people to cut, workers say. *Business Insider*. Retrieved August 19,2022,from https://www.businessinsider.com/facebook-contract-workers-accenture-austin-lost-jobs-2022-8

Hayes, A. F. (2013). *Introduction to mediation, moderation, and conditional process analysis: A regression–based approach*. New York: Guilford Press.

Herlocker, J., Konstan, J., Terveen, L., & Riedl, J. (2004). Evaluating collaborative filtering recommender systems. *ACM Transactions on Information Systems, 22*(1), 5–53.

Hitsuwari, J., Ueda, Y., Yun, W., & Nomura, M. (2023). Does human–AI collaboration lead to more creative art? Aesthetic evaluation of human–made and AI–generated haiku poetry. *Computers in Human Behavior, 139*, 107539.

Hoff, K. A., & Bashir, M. (2015). Trust in automation: Integrating empirical evidence on factors that influence trust. *Human Factors: The Journal of the Human Factors and Ergonomics Society, 57*(3), 407–434.

Holford, W. (2022). An ethical inquiry of the effect of cockpit automation on the responsibilities of airline pilots: Dissonance or meaningful control? *Journal of Business Ethics, 176*(1), 141–157.

HR Daily Advisor Staff. (2017). Artificial Intelligence will become a regular part of HR in next 5 years. *HR Daily Advisor*. Retrieved 2017, https://hrdailyadvisor.blr.com/2017/06/08/artificial-intelligence-will-become-regular-part-hr-next-5-years/

Hur, J. D., Koo, M., & Hofmann, M. (2015). When temptations come alive: How anthropomorphism undermines self–control. *Journal of Consumer Research, 42*(2), 340–358.

Institute of Electrical and Electronics Engineers (IEEE). (2017). *Ethically aligned design: A vision for prioritizing human wellbeing with artificial intelligence and autonomous systems*. Author.

Ireland, L. (2020). Who errs? Algorithm aversion, the source of judicial error, and public support for self-help behaviors. *Journal of Crime and Justice, 43*(2), 174–192.

Jago, A. S. (2019). Algorithms and authenticity. *Academy of Management Discoveries, 5*(1), 38–56.

Jussupow, E., Benbasat, I., & Heinzl, A. (2020, June). Why are we averse towards algorithms? A comprehensive literature review on algorithm aversion.

Paper presented at the meeting of the Proceedings of the 28th European Conference on Information Systems (ECIS), An Online AIS Conference, Marrakech, Morocco.

Kahneman, D. (2003). Maps of bounded rationality: Psychology for behavioral economics. *The American Economic Review, 93*(5), 1449–1475.

Kahneman, D., & Tversky, A. (1972). Subjective probability: A judgment of representativeness. *Cognitive Psychology, 3*(3), 430–454.

Kharpal,A. (2023). A.I. poses existential risk of people being 'harmed or killed,' ex-Google CEO Eric Schmidt says. *Consumer Business Channel*. Retrieved May 2023, https://www.cnbc.com/2023/05/24/ai-poses-existential-risk-former-google-ceo-eric-schmidt-says.html

Kinowska, H., & Sienkiewicz, Ł. J. (2022). Influence of algorithmic management practices on workplace well-being–evidence from European organisations. *Information Technology & People (West Linn, Or.), 36*(8), 1–25.

Komatsu, T. (2016, March). Japanese students apply same moral norms to humans and robot agents: Considering a moral HRI in terms of different cultural and academic backgrounds. Paper presented at the meeting of 2016 11th ACM/IEEE International Conference on Human-Robot Interaction (HRI), Christchurch, New Zealand.

Kroll, J. A., Huey, J., Barocas, S., Felten, E. W., Reidenberg, J. R., Robinson, D. G., & Yu, H. (2017). Accountable algorithms. *University of Pennsylvania Law Review, 165*(3), 633–705.

Laakasuo, M., Palomäki, J., & Köbis, N. (2021). Moral uncanny valley: A robot's appearance moderates how its decisions are judged. *International Journal of Social Robotics, 13*(7), 1679–1688.

Langer, M., König, C. J., & Papathanasiou, M. (2019). Highly automated job interviews: Acceptance under the influence of stakes. *International Journal of Selection and Assessment, 27*(3), 217–234.

Langer, M., König, C., & Busch, V. (2021). Changing the means of managerial work: Effects of automated decision support systems on personnel selection tasks. *Journal of Business Psychology, 36*(5), 751–769.

Lee, M. K. (2018). Understanding perception of algorithmic decisions: Fairness, trust, and emotion in response to algorithmic management. *Big Data & Society, 5*(1), 2053951718756684.

Lee, T., & Boynton, L. A. (2017). Conceptualizing transparency: Propositions for the integration of situational factors and stakeholders' perspectives. *Public Relations Inquiry, 6*(3), 233–251.

Lehdonvirta, V. (2018). Flexibility in the gig economy: Managing time on three online piecework platforms. *New Technology, Work, and Employment, 33*(1), 13–29.

Leichtmann, B., Humer, C., Hinterreiter, A., Streit, M., & Mara, M. (2023). Effects of explainable artificial intelligence on trust and human behavior in a high-risk decision task. *Computers in Human Behavior, 139*, 107604.

Lemieux, P. (2017). Rise of the machines? *Regulation, 40*(4), 4–5.

Leo, X., & Huh, Y. (2020). Who gets the blame for service failures? Attribution of responsibility toward robot versus human service providers and service firms. *Computers in Human Behavior, 113*, 106520.

Lerner, J., Li, Y., Valdesolo, P., & Kassam, K. (2015). Emotion and decision making. *Annual Review of Psychology, 66*(1), 799–823.

Li, X., & Sung, Y. (2021). Anthropomorphism brings us closer: The mediating role of psychological distance in User–AI assistant interactions. *Computers in Human Behavior, 118*, 106680.

Liu, N., Kirshner, S., & Lim, E. (2023). Is algorithm aversion WEIRD? A cross-country comparison of individual-differences and algorithm aversion. *Journal of Retailing and Consumer Services, 72*, 103259.

Logg, J. M., Minson, J. A., & Moore, D. A. (2019). Algorithm appreciation: People prefer algorithmic to human judgment. *Organizational Behavior and Human Decision Processes,151*, 90–103.

Lokhorst, G. J., & van den Hoven, J. (2011). Responsibility for military robots. In P. Lin, K. Abeney, & George A. Bekey (Eds.). *Robot ethics: the ethical and social implications of robotics* (pp. 145–156). Cambridge, MA: The MIT Press.

Longoni, C., Bonezzi, A., & Morewedge, C. K. (2019). Resistance to medical artificial intelligence. *Journal of Consumer Research, 46*(4), 629–650.

Longoni, C., & Cian, L. (2022). Artificial intelligence in utilitarian vs. hedonic contexts: The "word–of–machine" effect. *Journal of Marketing, 86*(1), 91–108.

Mahmud, H., Islam, A., Ahmed, S., & Smolander, K. (2022). What influences algorithmic decision-making? A systematic literature review on algorithm aversion. *Technological Forecasting & Social Change, 175*, 121390.

Malle, B. F., Scheutz, M., Forlizzi, J., & Voiklis, J. (2016, March). Which robot am I thinking about? The impact of action and appearance on people's evaluations of a moral robot. Paper presented at the meeting of 2016 11th ACM/IEEE International Conference on Human-Robot Interaction (HRI), Christchurch, New Zealand.

May, F., & Monga, A. (2014). When time has a will of its own, the powerless don't have the will to wait: Anthropomorphism of time can decrease patience. *Journal of Consumer Research, 40*(5), 924–942.

McAloon, A. (2021). Xsolla lays off 150 after an algorithm ruled staff 'unengaged and unproductive'. *Game Developer.* Retrieved August

2021, https://www.gamedeveloper.com/business/xsolla-lays-off-150-after-an-algorithm-ruled-staff-unengaged-and-u nproductive-#close-modal

McFarland, M. (2014). Elon Musk: 'With artificial intelligence we are summoning the demon.' *The Washington Post*. Retrieved October 2014, https://www.washingtonpost.com/news/innovations/wp/2014/10/24/elon-musk-with-artificial-intelligence-we -are-summoning-the-demon/

McNee, S., Riedl, J., & Konstan, J. (2006, April). Being accurate is not enough: How accuracy metrics have hurt recommender systems. Paper presented at the meeting of CHI '06 Extended Abstracts on Human Factors in Computing Systems, Montréal, Québec, Canada.

Mende, M., Scott, M., Van Doorn, J., Grewal, D., & Shanks, I. (2019). Service robots rising. *Journal of Marketing Research, 56*(4), 535–556.

Millet, K., Buehler, F., Du, G., & Kokkoris, M. (2023). Defending humankind: Anthropocentric bias in the appreciation of AI art. *Computers in Human Behavior, 143*, 107707.

Moore, D. A., & Healy, P. J. (2008). The trouble with overconfidence. *Psychological Review, 115*(2), 502–517.

Mori, M. (1970). The uncanny valley. *Energy, 7*, 33–35.

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

Moussawi, S., & Koufaris, M. (2019, January). Perceived intelligence and perceived anthropomorphism of personal intelligent agents: Scale development and validation. Paper presented at the meeting of Proceedings of the 52nd Annual Hawaii in International Conference on System Sciences, Hawaii, United State.

Nass, C., Steuer, J., & Tauber, E. R. (1994, April). Computers are social actors. Paper presented at the meeting of Conference on Human Factors in Computing Systems, Boston, Massachusetts, United State.

Natarajan, M., & Gombolay, M. (2020, March). Effects of anthropomorphism and accountability on trust in human robot interaction. Paper presented at the meeting of 2020 15th ACM/IEEE International Conference on Human–Robot Interaction (HRI), Cambridge, United Kingdom.

Nefdt, R. (2020). A puzzle concerning compositionality in machines. *Minds and Machines (Dordrecht), 30*(1), 1–21.

Newman, D., Fast, N., & Harmon, D. (2020). When eliminating bias isn't fair: Algorithmic reductionism and procedural justice in human resource decisions. *Organizational Behavior and Human Decision Processes, 160*, 149–167.

Nicholson Price, W. (2018). Big data and black-box medical algorithms. *Science Translational Medicine, 10*(471), Article aao5333.

Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review, 84*(3), 231–259.

Nørskov, S., Damholdt, M., Ulhøi, J., Jensen, M., Ess, C., & Seibt, J. (2020). Applicant fairness perceptions of a robot–mediated job interview: A video vignette–based experimental survey. *Frontiers in Robotics and AI, 7*, Article 71.

Park, H., Ahn, D., Hosanagar, K., & Lee, J. (2021, May). Human–AI interaction in human resource management: Understanding why employees resist algorithmic evaluation at workplaces and how to mitigate burdens. Paper presented at the meeting of the 2021 CHI Conference on Human Factors in Computing Systems, Yokohama, Japan.

Pasquale, F. (2015). *The Black Box Society: The Secret Algorithms That Control Money and Information.* Boston: Harvard University Press.

Prahl, A., & van Swol, L. (2017). Understanding algorithm aversion: When is advice from automation discounted? *Journal of Forecasting, 36*(6), 691–702.

Rai, A. (2020). Explainable AI: from black box to glass box. *Journal of the Academy of Marketing Science, 48*(1), 137–141.

Schoeffer, J., Machowski, Y., & Kuehl, N. (2022). Perceptions of fairness and trustworthiness based on explanations in human vs. automated decision-making. Paper presented at the meeting of the Annual Hawaii International Conference on System Sciences, .Hawaii, United State.

Schönbrodt, F. D., & Perugini, M. (2013). At what sample size do correlations stabilize? *Journal of Research in Personality, 47*(5), 609-612.

Shin, D. (2020). User perceptions of algorithmic decisions in the personalized AI system: Perceptual evaluation of fairness, accountability, transparency, and explainability. *Journal of Broadcasting & Electronic Media, 64*(4), 541–565.

Shin, D., & Park, Y. (2019). Role of fairness, accountability, and transparency in algorithmic affordance. *Computers in Human Behavior, 98*, 277–284.

Simth, B., & Shum, H. (2018). *The future computed: Artificial intelligence and its role in society.* Independently Published By Microsoft.

Swart, J. (2021). Experiencing algorithms: How young people understand, feel about, and engage with algorithmic news selection on social media. *Social Media Society, 7*(2), 205630512110088.

Tomprou, M., & Lee, M. (2022). Employment relationships in algorithmic management: A psychological contract perspective. *Computers in Human Behavior, 126*, 106997.

Upadhye, C. (2018). How algorithms run Amazon's warehouse. *Medium.* Retrieved December 27, 2018, from https://charviupadhye.medium.com/how-algorithms-run-amazons-warehouse-61e620ad27a7

Waytz, A., Cacioppo, J., & Epley, N. (2010). Who sees human? The stability and importance of individual differences in anthropomorphism. *Perspectives on Psychological Science, 5*(3), 219–232.

Wen, Z., Zhang, L., Hou, J., & Liu, H. (2004). Testing and application of the mediating effects. *Acta Psychological Sinica, 36*(5), 614–620. [温忠麟, 张雷, 侯杰泰, 刘红云. (2004). 中介效应检验程序及其应用. 心理学报, 36(5), 614–620.]

Wesche, J. S., & Sonderegger, A. (2019). When computers take the lead: The automation of leadership. *Computers in Human Behavior, 101*, 197–209

Wien, A., & Peluso, A. (2021). Influence of human versus AI recommenders: The roles of product type and cognitive processes. *Journal of Business Research, 137*, 13–27.

Wu, M., Wang, N., & Yuen, K. (2023). Deep versus superficial anthropomorphism: Exploring their effects on human trust in shared autonomous vehicles. *Computers in Human Behavior, 141*, 107614.

Xu, L., & Yu, F. (2020). Factors that influence robot acceptance. *Chinese Science Bulletin, 65*(6), 496–510. [许丽颖, 喻丰. (2020). 机器人接受度的影响因素. 科学通报, 65(6), 496–510.]

Xu, L., Yu, F., & Peng, K. (2022). Algorithmic discrimination causes less desire for moral punishment that human discrimination. *Acta Psychological Sinica, 54*(9), 1076–1092. [许丽颖, 喻丰, 彭凯平. (2022). 算法歧视比人类歧视引起更少道德惩罚欲. 心理学报, 54(9), 1076–1092.]

Xu, L., Yu, F., Wu, J., Han, T., & Zhao., L. (2017). Anthropomorphism: Antecedents and consequences. *Advances in Psychological Science, 25*(11), 1942–1954. [许丽颖, 喻丰, 邬家骅, 韩婷婷, 赵靓. (2017). 拟人化: 从“它”到“他”. 心理科学进展, 25(11), 1942–1954.]

Yam, K., Bigman, Y., Tang, P., Ilies, R., De Cremer, D., Soh, H., & Gray, K. (2021). Robots at work: People prefer–and forgive–service robots with perceived feelings. *Journal of Applied Psychology, 106*(10), 1557–1572.

Yeomans, M., Shah, A., Mullainathan, S., & Kleinberg, J. (2019). Making sense of recommendations. *Journal of Behavioral Decision Making, 32*(4), 403–414.

Yogeeswaran, K., Złotowski, J., Livingstone, M., Bartneck, C., Sumioka, H., & Ishiguro, H. (2016). The interactive effects of robot anthropomorphism and robot ability on perceived threat and support for robotics research. *Journal of Human–Robot Interaction, 5*(2), 29—47.

Yu, H., Miao, C., Chen, Y., Fauvel, S., Li, X., & Lesser, V. (2017). Algorithmic management for improving collective productivity in crowdsourcing. *Scientific Reports, 7*(1), 12541.

Yuan, L., & Dennis, A. (2019). Acting like humans? Anthropomorphism and consumer's willingness to pay in electronic commerce. *Journal of Management Information Systems, 36*(2), 450–477.

Zajonc, R. B. (1968). Attitude effects of mere exposure. *Journal of Personality and Social Psychology, 9*(2), 1–27.

Zhang, Y., Xu, L., Yu, F., Ding, X., Wu, J., & Zhao, L. (2022). A three-dimensional motivation model of algorithm aversion. *Advances in Psychological Science, 30*(5), 1093–1105. [张语嫣, 许丽颖, 喻丰, 丁晓军, 邬家骅, 赵靓. (2022). 算法拒绝的三维动机理论. 心理科学进展, 30(5), 1093-1105.]

Zhao, T. (2018). Comprehensive considerations for the "revolution" of artificial intelligence: An ethical and ontological analysis. *Philosophical Trends,* (4), 5–12. [赵汀阳. (2018). 人工智能"革命"的"近忧"和"远虑"——一种伦理学和存在论的分析. 哲学动态, (4), 5–12.]

Zhao, T. (2019). How could AI develop its self-consciousness? *Journal of Dialectics of Nature, 41*(1), 1–8. [赵汀阳. (2019). 人工智能的自我意识何以可能? 自然辩证法通讯, 41(1), 1–8.]

Złotowski, J., Yogeeswaran, K., & Bartneck, C. (2017). Can we control it? Autonomous robots threaten human identity, uniqueness, safety, and resources. *International Journal of Human–Computer Studies, 100*, 48—54.

Note: Figure translations are in progress. See original paper for figures.