International Journal of Human–Computer Interaction

ISSN: (Print) (Online) Journal homepage: www.tandfonline.com/journals/hihc20

When AI is Perceived to Be Fairer than a Human: Understanding Perceptions of Algorithmic Decisions in a Job Application Context

Hyesun Choung, John S. Seberger & Prabu David

To cite this article: Hyesun Choung, John S. Seberger & Prabu David (15 Oct 2023): When AI is Perceived to Be Fairer than a Human: Understanding Perceptions of Algorithmic Decisions in a Job Application Context, International Journal of Human–Computer Interaction, DOI: 10.1080/10447318.2023.2266244

To link to this article: https://doi.org/10.1080/10447318.2023.2266244

INTERNATIONAL JOURNAL OF HUMAN–COMPUTER INTERACTION
https://doi.org/10.1080/10447318.2023.2266244

Taylor & Francis
Taylor & Francis Group

OPEN ACCESS Check for updates

When AI is Perceived to Be Fairer than a Human: Understanding Perceptions of Algorithmic Decisions in a Job Application Context

Hyesun Choung[a], John S. Seberger[b], and Prabu David[c]

[a]Department of Communication, College of Communication Arts and Sciences, Michigan State University, East Lansing, MI, USA; [b]Department of Information Science, College of Computing and Informatics, Drexel University, Philadelphia, PA, USA; [c]Rochester Institute of Technology, Rochester, NY, USA

ABSTRACT

This study investigates people's perceptions of AI decision-making as compared to human decision-making within the job application context. It takes into account both favorable and unfavorable outcomes, employing a 2×2 experimental design (decision-making agent: AI algorithm vs. human; outcome: favorable vs. unfavorable). Upon evaluating a job seeker's suitability for a position, participants viewed algorithmic decisions as fairer, more competent, more trustworthy, and more useful than those made by humans. Interestingly, when a candidate was deemed unsuitable for hiring, people reacted more negatively to the verdict given by a human than to the same judgment offered by AI. Moreover, participants credited algorithmic decisions with greater sensitivity to both quantitative and qualitative qualifications, thus indicating algorithmic appreciation. Our findings shed light on the psychological basis of perceptions surrounding Algorithmic Decision-Making (ADM) and the responses to the decisions rendered by ADM systems.

KEYWORDS
Artificial Intelligence (AI); Algorithmic Decision-Making (ADM); Human-Decision-Making (HDM); outcome favorability bias; algorithmic appreciation; algorithmic aversion

1. Introduction

Ushered in by the revolution in artificial intelligence (AI), algorithmic decision-making (ADM) systems, or systems that use data and statistical models to make context-specific decisions, have become intricately involved in our lives.[1] For example, ADM is routinely used in college admissions (Marcinkowski et al., 2020), news recommendations (Thurman et al., 2019), and a vast array of contexts including policy, health, finance, and employment (Araujo et al., 2020; Rahwan et al., 2019). Across such contexts, ADM influences decisions once considered the exclusive domain of humans. Such influence raises concerns about bias, social justice, ethics, and human autonomy (Choung et al., 2022a; Krafft et al., 2022), tempering the promise of ADM with skepticism that is grounded in the fundamental belief that humans are better at making complex, context-sensitive, and nuanced human decisions. Despite such concerns, experts predict that algorithms will eventually outperform humans in many contexts (Grace et al., 2018).

Among various applications, the use of AI algorithms in hiring processes has gained much attention from human resource (HR) practitioners and researchers (Engler, 2021). In order to design and deploy trustworthy human-centered ADM in hiring processes, empirical understanding of when and why people perceive ADM to be fair is essential (Starke et al., 2021). While job applicants may perceive of ADM as a better option than HDM because algorithms are consistent in the application of rules and (ostensibly) not prone to human biases, these same strengths can be perceived as weaknesses (Araujo et al., 2020). Such strengths may double as weaknesses because algorithms are limitedly sensitive to context and reductive in their approach, reducing the rich qualitative aspects of human performance into quantifiable metrics (Newman et al., 2020).

Currently, studies on perceptions about AI-enabled technology used in employment processes report mixed findings (e.g., van Esch et al., 2021; Wesche & Sonderegger, 2021). The present study contributes to building consensus by describing the results of an experiment in which perceptions about ADM and HDM were compared on fairness, competency, trust, and usefulness for favorable and unfavorable outcomes in the context of a job candidate evaluation using a 2×2 (decision-making agent: AI algorithm vs. human; outcome: favorable vs. unfavorable) between-subjects design.

1.1. ADM in the hiring process

Many companies have already implemented ADM processes in the review and screening of applicants, analysis of interviews, performance reviews, and determination of

CONTACT Hyesun Choung choungh@msu.edu Department of Communication, College of Communication Arts and Sciences, Michigan State University, East Lansing, MI, USA

Supplemental data for this article can be accessed online at https://doi.org/10.1080/10447318.2023.2266244.

promotions and raises (Hunkenschroer & Luetge, 2022). Employers such as Google (Kuehn, 2018) claim that ADM outperforms HDM in hiring processes. Yet, such claims are not without contestation or controversy rooted in continued, context-specific reliance on HDM and the potential discriminatory outcomes of ADM hiring processes (Dastin, 2018; Köchling & Wehner, 2020). Ultimately, the perceived fairness of ADM in hiring processes resides with the applicant affected by the decision (Langer & Landers, 2021); as does the ultimate acceptance of such systems (Lee & Cha, 2023). Hence, ADM applications must be implemented with an abundance of caution, with a deep under of fairness and trust in algorithms alongside their usefulness and competence.

As ADM is widely used – even celebrated – by organizations in the tech sector and beyond, researchers and practitioners are increasingly paying attention to the public's perceptions of algorithmic decisions and factors that influence acceptance of ADM systems (Helberger et al., 2020). While most prior studies focus on the use of ADM by companies or organizations and job candidates' reactions to such automated processes (van Esch et al., 2021), little is known about job candidates' view of ADM systems that directly provide service *for them*. For example, reading job advertisements and evaluating fit for an advertised position is the first task of job seekers. In this initial state of the application process, job seekers typically know little about the organizations or jobs for which they could apply. Thus, in the absence of direct information, they may find an ADM system that provides fit information useful. Job seekers might then use such a system as an informational cue and signal to apply (Wesche & Sonderegger, 2021). Next, we consider key concepts and theories associated with perceptions of algorithmic decisions in fit evaluations for a job. Notably, we do so relative to the earliest phase of the job-seeking process: the initial assessment of fit between an applicant and an advertised position.

1.1.1. Perceptions of fairness, competence, trust, and usefulness in ADM

In considering concepts and theories associated with ADM in job-seeking contexts, this study is grounded in knowledge about how and why people use and continue to use algorithmic technologies. More specifically, we consider people's perceptions of the fairness, competence, trust, and usefulness of algorithmic decisions.

Studies have shown that fairness is one of the key attributes that shapes perceptions of, and experiences with, algorithms (Lee et al., 2019). The issue of fairness, along with potential biases, is attracting more scrutiny, particularly in the context of ADM in HR (Köchling & Wehner, 2020). *Fairness* is often assessed with equality (i.e., procedural fairness, see Leventhal, 1980) and equity (i.e., distributive fairness, see Deutsch, 1975) in treating and evaluating other people. It is often thought that algorithms apply the same rules and procedures with each iteration, thus enabling consistent interpretation. Therefore, algorithmic decisions could be perceived to be fairer than human decisions for HR tasks. However, some studies report contrasting findings. For instance, AI-supported selection tools are perceived to be less fair than traditional human-based section processes (Köchling & Wehner, 2023). This perception may stem from concerns about potential bias in AI-tools (Cheng & Hackett, 2021) and unfamiliarity with the technology (Gonzalez et al., 2022).

Furthermore, users' perceptions and expectations regarding ADM systems may be influenced by doubts regarding the capabilities of AI-driven technology broadly construed. Such doubts specifically relate to the competence to AI make holistic decisions about human potential. *Competence* refers to the functional capability of a technology (Mcknight et al., 2011), identifying whether applicants have the knowledge, skills, abilities, and other characteristics needed to perform effectively in particular jobs, resulting in fitness assessment for a job.

However, fairness and competence are only two parts of a bigger puzzle. As ADM systems continue to evolve and proliferate, they need to be trusted (Alexander et al., 2018; Fenneman et al., 2021). Traditionally, trust is tied to relationships *between people* and is essential to build mutuality and interdependence in human communication. Trust serves as a fundamental human mechanism to cope with vulnerability, uncertainty, complexity, and ambiguity, which collectively constitute risks. Increasingly, researchers are addressing social and psychological trust in ADM and/or the broad suite of technologies contained in the set "AI." Glikson and Woolley (2020) identified factors that predict cognitive and emotional trust in AI, while Choung et al. (2022b) distinguished human-like trust and functionality trust in AI. Prior work suggests that trust is a key predictor in the adoption of AI technologies (Shin, 2021), and that trust in algorithms increases when others' use of the algorithm is disclosed (Alexander et al., 2018). Similarly, algorithm adherence depends on the perceived efficacy of an algorithm and the trust that people have in that algorithm (Fenneman et al., 2021).

Finally, the perceived *usefulness* of a given technology is another core determinant of acceptance of a technology, as the Technology Acceptance Model (TAM) and its subsequent iterations suggest (Davis, 1989; Venkatesh & Bala, 2008; Venkatesh & Davis, 2000). TAM has been widely used to understand adoption of new technologies (McLean & Osei-Frimpong, 2019) including recent AI-driven technologies (Choung et al., 2022b). In sum, based on prominence in the literature and relationships with one another, we chose fairness, competence, trust, and usefulness as key outcomes in this study, which are examined in our first research question:

> **RQ1:** How are ADM and HDM perceived in an evaluative task in a job application scenario for the following outcomes: (a) fairness, (b) competence, (c) trust, and (d) usefulness?

To further frame this question, we presently engage in a review of works related to algorithm aversion and appreciation and their contextuality; outcome favorability bias; and the attribution of quantitative and qualitative sensitivity to algorithmic decisions.

1.2. Algorithm aversion and appreciation

In a broader terms, algorithm aversion describes the behavior where individuals downplay decisions made by algorithms compared to those made by humans. This tendency can manifest despite an individual's conscious awareness of the high-performance of algorithms, and it may originate from an unconscious, fundamental distrust towards these systems (Mahmud et al., 2022, p. 3). Earlier research also pointed to specific instances of this aversion, such as negative perceptions that algorithms are reductive (Newman et al., 2020) and lack "the human touch," or contextual sensitivity (Dietvorst et al., 2015). These studies indicate that people tend to discount algorithmic decisions relative to human decisions, particularly when an algorithm is observed to have made an error (Dietvorst et al., 2015; Prahl & Van Swol, 2017). Following an experience in which the algorithm had erred, people's aversion to the algorithm persisted even in the face of evidence that the algorithm outperforms humans on average (Dietvorst et al., 2015).

Paradoxically, there is also evidence that people trust algorithmic advice more than human advice. *Algorithm appreciation* refers to positive perceptions that machines are safer and more trustworthy than humans (Logg et al., 2019). Through a series of experiments, Logg et al. (2019) demonstrated that laypeople tend to rely more on algorithmic advice than human advice for various estimation tasks, such as weight estimation, song popularity prediction, and romantic matchmaking tasks. However, algorithm appreciation declined when individuals had to choose between the algorithm and their own judgment or when the individuals themselves were experts in areas in where ADM was employed. Other studies have reported findings of preference for AI or algorithms than human experts in contexts like health and justice (Araujo et al., 2020) and news recommendation (Thurman et al., 2019). Krügel et al. (2022) found that users readily trust ethical advice from algorithms, even when they lack information about the training data or when the data is presumed to be biased. These findings suggest that people tend to be more receptive of or less discerning about the algorithms than anticipated.

In the field of HR, empirical evidence about job seekers' reactions to ADM is mixed, exhibiting both algorithm aversion and appreciation at different stages of the application process. For example, Wesche and Sonderegger (2021) found that people demonstrate strong aversion to interviews being conducted by AI-driven technologies, but found only weaker negative effects when application screening involved ADM systems. The aversion to algorithms used in human evaluations might be rooted in the complexities of quantifying human potential and making decisions out of context. This algorithm-driven decision-making process might seem more challenging to comprehend compared to the intuitive processes that typically underpin human decision-making (Yeomans et al., 2019). In some studies, participants have indicated that they would rather take their chances with a subjective human than with an ostensibly objective ADM process (Newman et al., 2020).

To better understand when people appreciate or are averse to recommendations from algorithms, researchers have paid attention to contextual factors that influence perceptions. A systematic review (Mahmud et al., 2022) of the literature identifies four themes in ADM research: algorithm, individual, task, and high-level factors. The theme of *algorithm* focuses on the decision characteristics and the design of the algorithm, such as explainability, accuracy, and decision outlook. *Individual factors* are related to personality, demography, and familiarity with algorithms. *Task factors* focus on the nature of the task itself, such as subjective vs. objective evaluation, technical vs. human task, complex vs. simple evaluation, and whether the task involves moral judgment. *High-level factors* refer to macro-level factors such as organizational, societal, and cultural contexts. Mahmud et al. (2022) conclude that the preponderance of research has examined algorithms and individual factors and call for more research on tasks and macro-level factors.

1.3. Outcome favorability bias in ADM

One approach to task evaluation involves assessing the match between a decision rendered by AI and the expectations of the recipient of the decision. While outcomes matter critically in decision-making, it has not been explored extensively in the context of ADM (Mahmud et al., 2022; Wang et al., 2020). In Wang et al.'s (2020) study, outcome favorability in ADM was defined as "whether an algorithm's prediction or decision is favorable to specific individuals or groups" (p. 684). They found that when the outcome is (i) unbiased across different demographic groups *and* (ii) favorable to the individual affected by the decision, people are more likely to evaluate the decision-making process as fair. Furthermore, their study highlighted users' perceptions of algorithmic fairness and conformity of outcome favorability bias, explained as a tendency to be self-serving, accepting responsibility for positive outcomes and denying responsibility for negative outcomes (Schroth & Shah, 2000; Shepperd et al., 2008).

The experimental study by Lünich and Kieslich (2022) further contextualizes findings discussed above by exploring the effects of the decision-making agent and social favorability conditions where they examined resource allocation to a favored group (teachers) versus a less favored group (prisoners). Lünich and Kieslich's findings revealed that decisions favoring early vaccination of a less-preferred group were deemed less legitimate compared to decisions favoring early vaccination of a generally preferred group. In the light of these insights, outcome favorability extends beyond mere confirmation of one's judgment. It can be better understood as the degree to which a decision outcome aligns with the recipient's expectations, values, and social or moral beliefs. In our study, we conceptualize a decision as having a favorable outcome when it aligns with the recipient's expectations, and as unfavorable when it does not. Drawing on the existing literature on outcome favorability bias, we propose our first hypothesis:

H1: Compared to an unfavorable outcome, when a favorable outcome is offered by a decision-maker, that decision-maker will be evaluated more favorably on (a) fairness, (b) competence, (c) trust and (d) usefulness.

However, it is possible that the effect of outcome favorability may not be uniform across ADM and HDM conditions. Recent research has shown differences in the impressions of humans and robots based on behavior valence and choice (Edwards & Edwards, 2022). People attributed unpopular robot behaviors to a robot's dispositions and were more confident in their attributions of a robot than a human. Furthermore, a study by Yalcin et al. (2022) found that consumers react less positively to favorable decisions made by algorithms as compared to those made by humans, while the difference is less pronounced for unfavorable decisions. They attributed this to the ease of internalizing a positive decision from a human and externalizing a negative decision regardless of the decision-maker. Thus, it is plausible that different cognitive processes are involved in the evaluation of ADM and HDM, particularly in motivated processing and social comparisons.

As social creatures, people continuously calibrate behaviors and expectations in relation to other people and routinely engage in comparison to assess their self-worth. Comparison to a superior counterpart can lead to inspiration or threat (Gerber et al., 2018). Yet, the degree of such perceived inspiration or threat (i.e., the outcome of social comparison) may be different in human-algorithm interactions (Spatola & Normand, 2021) than in human-human interactions.

Although algorithms are, indeed, *others*, ADM systems are not *other people*. This distinction may lead to differences in how individuals perceive and react to outcomes provided by ADMs compared to those provided by humans. When faced with an unfavorable outcome, individuals may be more inclined to question the human decision-maker's motives or competence, while they may attribute the negative result from an algorithmic agent to its impartiality and objectivity. Conversely, when receiving favorable outcomes, people might view both humans and algorithmic decision-makers positively, appreciating their competence and fairness in making the decision. In short, the way people perceive and respond to outcomes from ADM or HDM systems may vary based on the favorability of outcome, which is the essence of our second research question:

> **RQ2:** Is there an interaction effect between the decision-making agent (AI algorithm vs. human) and outcome favorability (favorable vs. unfavorable) on perceived (a) fairness, (b) competency, (c) trust, and (d) usefulness?

1.4. Consideration of quantitative and qualitative attributes of job applicants

Extant literature on tasks shows that people are more willing to trust algorithmic decisions if the decision-making involves objective rather than subjective evaluation, given that algorithms are seemingly less fallible to objective biases in judgment (Castelo et al., 2019; Lee, 2018). People also rely more on algorithmic advice than human advice for difficult intellective tasks (Bogert et al., 2021). But for creative works such as art and music (Epstein et al., 2020; Tigre Moura & Maw, 2021) and in moral decision-making (Jago, 2019; Niszczota & Kaszás, 2020), people generally show algorithm aversion and prefer human decisions. These studies suggest that ADM will be appreciated if the tasks involve nonmoral, logically complex, and objective tasks (Mahmud et al., 2022). Interestingly, Krügel et al. (2022) found that individuals often trust ethical advice from AI-driven ADM systems, irrespective of their understanding of the algorithm's training data or even when they are aware of biases in the training data. This observation implies that the concept of algorithm appreciation may extend beyond nonmoral domains. Such evidence suggests that people's perceptions of ADM are continuously evolving, seemingly influenced by their ongoing interactions and experiences with algorithms and AI technologies.

Potential perceptual biases toward ADM in the job candidate evaluation context can be explained using two underlying considerations. In general, people appreciate algorithms because they hold positive stereotypes of algorithms, thinking that algorithms are more objective and superior at dealing with quantifiable metrics than humans (Sundar & Kim, 2019). On the other hand, the reason to perceive of algorithmic decisions as less fair than human decisions is predicated on the belief that algorithms do not take qualitative information or context into account (Starke et al., 2021).

Newman et al. (2020) introduced the idea of "reductionism" as a possible explanation for fairness or unfairness of algorithms in HR decisions. Reductionism refers to the process of quantification, which can potentially enhance the efficiency and fairness of organizational decisions by reducing human biases. However, organizations that rapidly adopt algorithms as a solution to human decision-making biases might overlook the qualitative aspects of human nature and the specific circumstances in which they occur (Newman et al., 2020, p. 151).

Such reductionism exists in opposition to the richness of subjectivity, complicating the attribution of quantitative (i.e., objective) and qualitative (i.e., subjective) sensitivity to a candidate's qualifications. Newman et al. (2020) found that decisions made by algorithms are perceived to be less fair than decisions made by humans when individuals perceive the algorithmic decision-making process as fundamentally reductionistic. As quantification is a core aspect of reductionism, in their studies, reductionism was directly manipulated through an explicit emphasis on quantitative factors to see if its effect would vary between human and algorithm-driven procedures.

Based on findings by Newman et al. (2020), we expected that people may believe that algorithms reduce the attributes of job candidates to a few numbers that results in the loss of rich qualitative information. However, this belief can be updated as people increasingly interact with AI-enhanced technology and become more aware of its capability to process qualitative and to some extent context-sensitive information.

When a decision-making task involves more than mechanical calculations, evaluations of the decision depend on the extent to which the agent takes account of quantitative (i.e., objective) and qualitative (i.e., subjective) information.

Particularly, when reviewing a job applicant, a holistic assessment should include *both* quantifiable metrics and qualitative impressions. While hard skills (e.g., degree and prior experience) are more easily quantifiable, soft skills (e.g., communication skills and creativity) are more subjective and qualitative. The former skills are readily "machine-readable;" the latter are perhaps best *interpreted* by an agent that possesses the same embodied mode of subjectivity as the candidate. Hence, we examine perceptions of the reliance on quantitative and qualitative attributes of ADM and HDM for favorable and unfavorable outcomes:

> **RQ3**: Is there a difference in the perceived levels of considerations of (a) quantitative attributes and (b) qualitative attributes by the decision-making agent and outcome favorability?

In addition, using a mediation analysis, we examined whether these attributes could shed light on the underlying psychological mechanisms of perceptions of algorithmic decisions and human decisions.

> **RQ4**: How are perceptions of decision-making agents mediated by the perceived levels of the job seeker's (a) quantitative and (b) qualitative attributes considered in the evaluation?

2. Method

2.1. Overview

The study employed a 2×2 (decision-making agent: AI algorithm vs. human; outcome: favorable vs. unfavorable) between-subjects online experiment, which was approved by a university Institutional Review Board. An online experiment was chosen as the most suitable format for our study for two reasons: first, such a format allows specific and controlled focus on those variables relevant to the research questions and hypotheses stated above; second, it complements related online experimental work about people's perceptions about algorithmic decision-making (e.g., Lee, 2018; Lee et al., 2019). Participants were recruited through Amazon Mechanical Turk (MTurk) between October and November of 2021. All participants were based in the U.S., with an MTurk record of at least 50 completed tasks and approval rate of 90% or higher. Participants received \$1.50 for their participation (i.e., a rate of roughly \$10/hr).

2.2. Participants

A total of 266 participants completed the experiment and 31 respondents who gave wrong answers to both attention check questions were excluded from the sample.[2] Participants who answered either or both attention checks correctly were included in the analysis. We carried out the analyses with a final sample size of 235. More than one-half of the participants were men (54%), with women (45%) and other (1%) making up slightly less than half of the sample. The average age of participants was 49.25 ($SD = 11.20$). Most participants identified as White (77.87%), with Hispanic or Latino/a/x (15.74%), Black or African American (4.2%), Asian (1.7%), and bi- or multi-racial (<1%) making up the rest of the sample. Most participants had a bachelor's degree or a higher level of education (72.8%), with median income range between \$40,000 and \$50,000.[3] The mean duration of the experiment was nine minutes.

2.3. Procedure

Participants completed a consent form and were then introduced to a fictitious online job recommendation platform called "JobConnect" which offered an evaluation of a job seeker's qualification and fit for a job. Participants were assigned randomly to one of the four 2×2 conditions: ADM-favorable ($n = 56$), ADM-unfavorable ($n = 59$), HDM-favorable ($n = 60$), HDM-unfavorable ($n = 60$). All participants read a job candidate's profile and a job description for a social media internship. They were asked to review the candidate's fit for the position and later directed to view the JobConnect agent's evaluation of the candidate. Participants assigned to the ADM condition were informed that the JobConnect platform relied on an AI algorithm to generate the evaluation. Participants in the HDM condition were informed that the evaluation was made by a human job coordinator.

Participants who were assigned to the favorable condition saw an agent's favorable evaluation of the job candidate with a high fit score of ("90/100") and a recommendation to apply for the position. In the unfavorable condition, participants viewed an agent's negative evaluation of the job candidate's fit to the advertised position with a low fit score of ("40/100").[4] Stimuli are presented in Appendix. We opted for scores of 90/100 and 40/100 to represent high and low fit, respectively, using the standard 100-point academic grading scale with a score of 90 considered an A-grade or excellent performance and 40 as a failing grade or poor performance. This scale is familiar to U.S. participants and provides a clear distinction between a favorable and unfavorable outcome.

2.4. Measures

After reviewing the agent's recommendation, participants responded to questions about their perceptions of JobConnect and its decision-making process. Perceived fairness was measured with one item (Lee, 2018): "To what extent do you think the evaluation of the applicant by JobConnect was fair?" *(1 = Very Unfair, 7 = Very Fair)*. Perceived competency was measured with an item that we created: "In your opinion, how well did JobConnect evaluate the applicant's fit to the advertised position?" *(1 = Performed Poorly, 7 = Performed Very Well)*. Perceived trust (Lee, 2018) was measured by asking "How much do you trust the fit assessment offered by JobConnect?" *(1 = No Trust, 7 = Strongly Trust)*. Perceived usefulness was measured with three items, adapted from Davis (1989) ($\alpha = 0.92$): "JobConnect enables applicants to find fitting jobs," "JobConnect improves efficiency in finding jobs," "JobConnect is useful to find jobs that fit users interests." Items were rated on a 7-point (*1 = Strongly Disagree, 7 = Strongly Agree*). Participants were also asked to estimate the six evaluation criteria: "Please indicate the extent to which you think each of these criteria was considered by JobConnect to evaluate the fit of the applicant." The criteria were drawn from the job description. Three of the six criteria—degree and courses taken, work experience, and social media experience—were combined to create an index of *quantitative attributes* ($\alpha = 0.82$). Three other criteria—motivations and life story, communication skills, and creativity—were combined to create an index of *qualitative attributes* ($\alpha = 0.82$). These items, adapted from Newman et al. (2020), were rated on 7-point scale *(1 = Not Considered, 7 = Strongly Considered)*. See Table 1 for the correlations, means, and standard deviations for all variables used in the analyses. The screenshot of the post-exposure questionnaire can be found in the Supplemental Materials.

2.5. Manipulation check

The results of the manipulation checks are summarized in Table 2. Two questions were asked to check the agent manipulation: To what extent do you agree or disagree with the following statement about JobConnect? ("The profile of the user was assessed by an AI algorithm" and "The profile of the user was assessed by a human coordinator"), which were rated on a 7-point scale (*1 = Strongly Disagree, 7 = Strongly Agree*). Participants in the ADM condition expressed higher agreement that the profile was evaluated by an algorithm. Similarly, participants in the HDM condition expressed higher agreement that the profile was evaluated by a human. These findings suggest that the manipulations had the intended effect of conveying the identity of the agent offering the evaluation, although the manipulation was stronger in the human agent condition than in the algorithmic agent condition. For the outcome favorability manipulation, respondents were asked to recall a fit score generated by the agent. Three possible answers were given: "90/100" (favorable), "40/100" (unfavorable), and "I don't know." Nearly all participants in both the favorable and unfavorable conditions chose the correct score. Their accurate recall suggests that they paid attention to the outcome provided by the decision-making agent and correctly perceived it as either favorable or unfavorable. Thus, we concluded that the favorability manipulation was successful.

Table 1. Intercorrelations, means, and standard deviations of measured variables.

	1	2	3	4	5	6
1. Perceived Fairness	–					
2. Perceived Competency	0.86	–				
3. Perceived Trust	0.75	0.75	–			
4. Perceived Usefulness	0.78	0.83	0.78	–		
5. Quantitative Attributes	0.71	0.69	0.62	0.64	–	
6. Qualitative Attributes	0.63	0.65	0.59	0.58	0.67	–
M	5.16	5.12	5.03	5.45	5.47	5.09
SD	1.74	1.81	1.57	1.31	1.36	1.34

N = 235. All correlations significant at $p < 0.01$.

2.6. Data analysis

Analysis was carried out in three steps. First, using two-way ANOVA, we examined the main effects and the interaction of the experimental conditions on the four key dependent variables (fairness, competence, trust, and usefulness). In the second step, using the same analytic approach, we examined the role of quantitative and qualitative attributes in candidate evaluation. In the third step, we tested moderated mediation models using a PROCESS macro (Hayes, 2018) to explore the mediating role of the consideration of quantitative and qualitative attributes. Moderated mediation depicted in Figure 1 was examined with PROCESS Model 8.[5]

3. Results

3.1. Effects of agent and outcome favorability on perceptions

First, we present findings on fairness, competence, trust and usefulness perceptions, each analyzed using a 2×2 between-subjects ANOVA. Overall, both main effects and the interaction were statistically significant at $p < 0.001$ for fairness, competence, and trust. For usefulness, the main effects were significant at $p < 0.05$, but the interaction was tending toward significance ($p = 0.053$). The means of each dependent variable by experimental condition, along with statistical significance, are summarized in Table 3. In the following sections, we present detailed findings of each outcome, and address relevant research questions and hypotheses.

3.1.1. Perceived fairness

As indicated in the first entry of Table 3, main effects of agent, $F\ (1, 231) = 19.29$, $p < 0.001$, $\eta_p{}^2 = 0.08$, and outcome, $F\ (1, 231) = 73.72$, $p < 0.001$, $\eta_p{}^2 = 0.24$, and the interaction, $F\ (1, 231) = 23.68$, $p < 0.001$, $\eta_p{}^2 = 0.09$, were significant.

Table 2. Agent and outcome favorability manipulation checks.

	Manipulation item	M (SD) ADM	M (SD) HDM	*F*	*Df*	*p*
Agent manipulation	"The profile of the user was assessed by an AI"	5.56 (1.28)	2.71 (2.35)	156.58	1233	<0.001
	"The profile of the user was assessed by a human coordinator"	5.04 (1.69)	5.80 (2.03)	14.36	1233	<0.001
		% Favorable	% Unfavorable	χ^2	*Df*	*p*
Outcome manipulation	"What was the fit score generated by the Job Connect's agent?"	Favorable = 98.3% Unfavorable = 1.7%	Favorable = 0% Unfavorable = 100%	227.13	1	<0.001

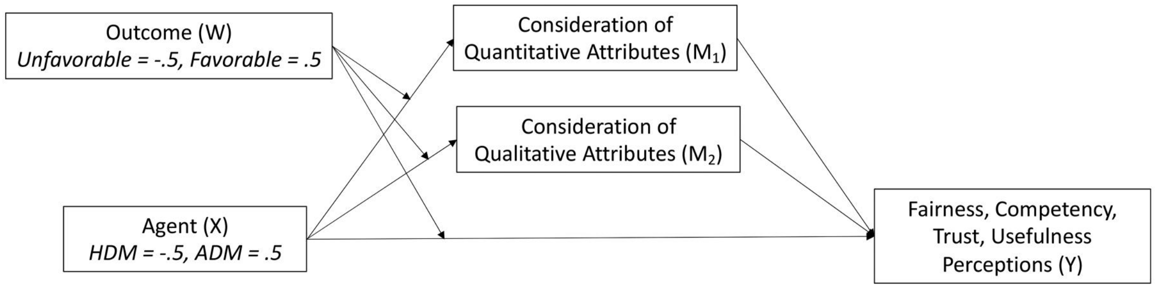

Figure 1. Moderated mediation Model: Effects of agent and outcome on perceived appreciation through Mediators.

Table 3. Observed means and standard deviations by experimental conditions.

	ADM		HDM			
	Favorable Outcome	Unfavorable Outcome	Favorable Outcome	Unfavorable Outcome	F (1, 231)	η_p^2
1. Perceived Fairness	5.93 (.83)	5.24 (1.50)	6.02 (1.31)	3.52 (1.84)	23.68***	.09
2. Perceived Competency	5.84 (1.02)	5.46 (1.50)	5.80 (1.46)	3.43 (1.94)	24.90***	.10
3. Perceived Trust	5.63 (1.12)	5.32 (1.44)	5.38 (1.37)	3.83 (1.63)	11.55**	.05
4. Perceived Usefulness	5.71 (1.06)	5.36 (1.58)	4.90 (1.59)	3.07 (1.75)	3.78+	.02
5. Quantitative Attributes	5.94 (.73)	5.48 (1.18)	6.01 (.92)	4.48 (1.78)	11.25**	.05
6. Qualitative Attributes	5.93 (.68)	5.46 (1.11)	5.03 (1.22)	3.98 (1.40)	3.79+	.02

Standard deviations are presented in parentheses.
***p < 0.001,.**p < 0.01,.*p < 0.05;.+p < 0.10.
F statistic is for the 2 × 2 interaction between decision-making agent and outcome favorability.

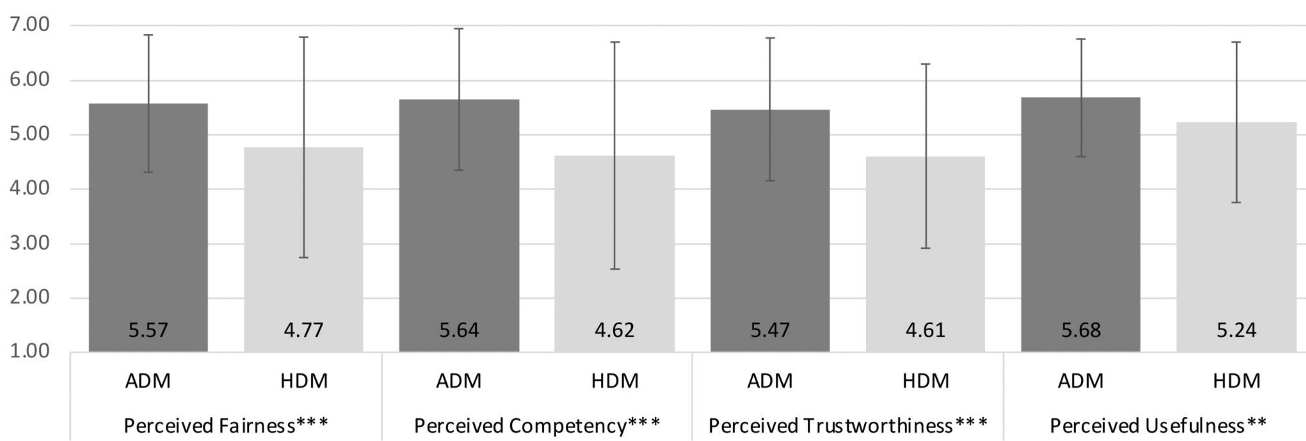

Figure 2. Mean differences between ADM vs. HDM conditions. ***p < 0.001, **p < 0.01.

3.1.2. Perceived competency

The findings related to perceived competency were like the findings for fairness (see second entry of Table 3). The main effects of agent (F [1, 231] = 26.91, p < 0.001, η_p^2 = 0.10) and outcome favorability (F [1, 231] = 47.72, p < 0.001, η_p^2 = 0.17), and the interaction (F [1, 231] = 24.90, p < 0.001, η_p^2 = 0.10) were statistically significant.

3.1.3. Perceived trust

Findings about perceived trust also followed the pattern observed for fairness and competency (see third entry of Table 3), with significant main effects for agent, F (1, 231) = 22.23, p < 0.001, η_p^2 = 0.09, and outcome favorability, F (1, 231) = 11.55, p < 0.01, η_p^2 = 0.05, and a significant interaction, F (1, 231) = 11.55, p < 0.01, η_p^2 = 0.05.

3.1.4. Perceived usefulness

A similar pattern was observed for usefulness, as indicated in the fourth entry of Table 3, though the interaction was only tending toward significance. The main effects of agent (F [1, 231] = 7.92, p < 0.01, η_p^2 = 0.03) and outcome favorability (F [1, 231] = 24.50, p < 0.001, η_p^2 = 0.10) were statistically significant, while the interaction effect was tending F (1, 231) = 3.78, p = 0.053, η_p^2 = 0.02.

The above findings provide answers to RQ1, H1, and RQ2. RQ1 asks whether algorithmic decision or human decision is preferred for job candidate's fit evaluation. Our findings show that an AI algorithm was more positively evaluated than a human coordinator in the fit assessment of job candidates (see Figure 2). People who saw the evaluation from an AI agent expressed greater levels of perceived fairness (RQ1a), competency (RQ1b), trust (RQ1c), and usefulness (RQ1d) than participants who saw an evaluation generated by a human job coordinator.

We predicted that when a favorable outcome is offered, it will lead to positive attributions of the decision-maker (H1). We find strong evidence in support of H1. Across all measures, the favorable outcome, compared to the unfavorable

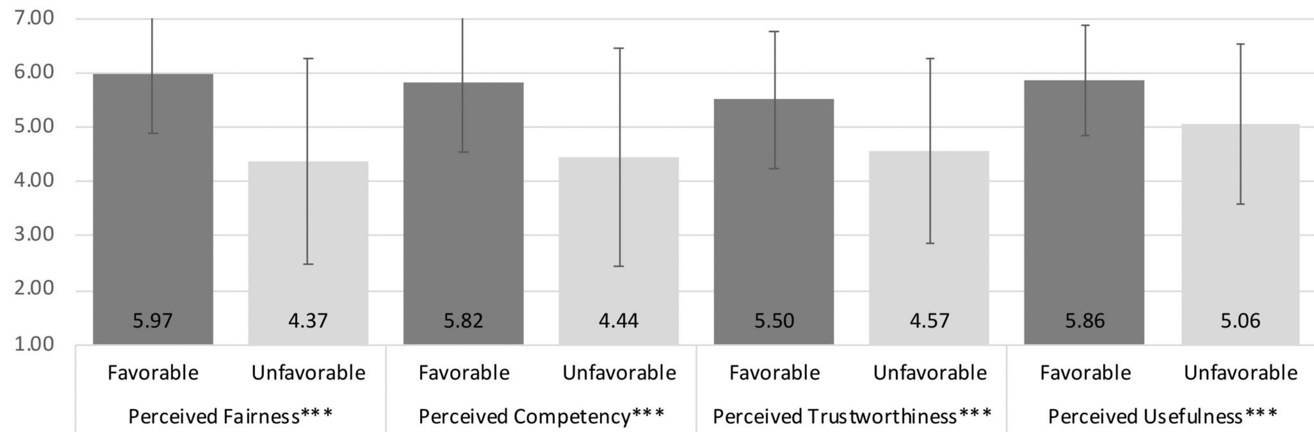

Figure 3. Mean differences between favorable vs. Unfavorable outcome conditions. ***$p < 0.001$.

outcome, led to more favorable assessments of decision-making agent. When the outcome was favorable to a job candidate, people perceived the decision-making agent as fairer (H1a), more competent (H1b), more trustworthy (H1c), and more useful (H1d) than when the outcome was unfavorable (see Figure 3).

RQ2 focused on the interaction between agent and outcome favorability on perceptions. Statistically significant interaction effects were observed for three of the four dependent variables (fairness, competency, trust), and the interaction was trending toward significance at $p = 0.053$ for usefulness. Given the significant interaction, the effects of agent (RQ1) and outcome favorability (H1) were examined by the four experimental conditions. A closer examination revealed that the appreciation of algorithmic decisions over human decisions occurred only when the outcome was unfavorable. The strong effect of algorithm appreciation in the unfavorable condition seemed to drive the observed main effect of agent.

As expected, the follow-up analysis revealed that algorithm appreciation was only significant in the unfavorable outcome condition ($p < 0.001$ for fairness, competency, and trust). This means that when the decision-making outcome is unfavorable, people are more likely to appreciate algorithmic decision-makers than human decision-makers. When the outcome is favorable, algorithmic decisions and human decisions are equally appreciated. Put differently, people exhibited more reactance to an unfavorable decision made by another human than to the same outcome offered by an algorithm.

In sum, our findings demonstrate that the preference toward algorithms over humans (i.e., algorithmic appreciation) depends on the outcome of the decision and is more pronounced when the outcome is unfavorable.

3.2. Considerations of quantitative and qualitative attributes

In this section, we examine considerations of quantitative and qualitative attributes that could potentially explain the pattern of findings (i.e., algorithmic appreciation) presented above. Quantitative consideration was obtained by averaging ratings of education, experience and social media experience, and qualitative was obtained by averaging ratings of motivation and life story, communication and creativity.

3.2.1. Perceived level of quantitative attributes assessment

For quantitative attributes, as indicated in the fifth entry of Table 3, main effects for both decision-making agent (F [1, 231] = 8.68, $p < 0.01$, $\eta_p^2 = 0.04$) and outcome favorability (F [1, 231] = 38.37, $p < 0.001$, $\eta_p^2 = 0.22$) were significant. The interaction also was statistically significant, F (1, 231) = 11.25, $p < 0.01$, $\eta_p^2 = 0.05$.

Next, the means were examined to characterize the interaction. In the favorable outcome condition, quantitative attributes contributed equally to the decisions of both AI algorithm ($M = 5.94$, $SD = 0.73$) and human ($M = 6.01$, $SD = 0.92$). In the unfavorable condition, quantitative attributes contributed more to the ADM condition ($M = 5.48$, $SD = 1.18$) than to the HDM ($M = 4.48$, $SD = 1.78$) condition. This shows that overall, participants agreed that the AI algorithm considered the job applicant's quantitative attributes (e.g., education, years of experience, social media experience) to a greater extent, and such perception was less influenced by the favorability of the decision outcome. In contrast, in the HDM condition, participants agreed that a human coordinator took into consideration quantitative attributes only when the outcome was favorable but not when it was unfavorable (RQ3a).

3.2.2. Perceived level of qualitative attributes assessment

For qualitative attributes, as summarized in the sixth entry of Table 3, a similar pattern was found on the main effects of decision-making agent (F [1, 231] = 64.14, $p < 0.001$, $\eta_p^2 = 0.22$) and outcome favorability (F [1, 231] = 26.18, $p < 0.001$, $\eta_p^2 = 0.10$). However, the interaction effect was not statistically significant, but tending toward significance, F (1, 231) = 3.79, $p = 0.053$, $\eta_p^2 = 0.02$.

Participants reported that the AI algorithm ($M = 5.69$, $SD = 0.95$) considered job candidate's qualitative attributes to a greater degree than the human coordinator ($M = 4.51$, $SD = 1.41$) across both outcomes. In other words, participants believed that the AI evaluated qualitative attributes of job candidates more thoroughly (RQ3b). In addition, consistent with the previous findings on outcome favorability, consideration of qualitative attributes were rated positively in the favorable ($M = 5.47$, $SD = 1.09$) condition compared to the unfavorable ($M = 4.72$, $SD = 1.46$) condition.

3.3. Explaining perceptions: Moderated mediation analysis

To answer RQ4, we examined quantitative and qualitative attributes as mediators using a PROCESS macro (Hayes, 2018). Moderated mediation was examined with PROCESS Model 8 for each of the dependent variables (fairness, competency, trust, and usefulness). A 95% CI was calculated using 5000 bootstrapping samples in the present study. The full results of the moderated mediation analysis are available in the Supplemental Materials.

Figure 4(a–d) summarizes the results of moderated mediation analyses. The effect of agent on quantitative ($\beta = 0.47$, $p < 0.05$) and qualitative ($\beta = 1.19$, $p < 0.001$) attributes was statistically significant, as well as the influence of decision outcome on quantitative ($\beta = 0.1.00$, $p < 0.001$) and qualitative attributes ($\beta = 0.76$, $p < 0.001$). The interaction between the agent and outcome significantly influence quantitative attributes ($\beta = -1.07$, $p < 0.001$), but not qualitative attributes ($\beta = -0.58$, $p = 0.05$), though it was tending towards significance. In essence, this part of the analysis replicates of the findings from the previous ANOVA (see the fifth and sixth entries in Table 3). Next, we examined the association among the quantitative and qualitative considerations and the four outcome variables. Both quantitative and qualitative were positively associated with four outcomes, with coefficients for quantitative (fairness [$\beta = 0.53$, $p < 0.001$], competency [$\beta = 0.51$, $p < 0.001$], trust [$\beta = 0.44$, $p < 0.001$], usefulness [$\beta = 0.40$, $p < 0.001$]) greater than the coefficients for qualitative ([$\beta = 0.53$, $p < 0.001$], competency [$\beta = 0.51$, $p < 0.001$], trust [$\beta = 0.44$, $p < 0.001$], usefulness [$\beta = 0.40$, $p < 0.001$]) attributes were positively associated with the four outcome variables. These positive associations suggest that higher ratings of quantitative and qualitative attributes of the candidates led to higher ratings of the perceived outcomes of the evaluation of the candidate.

Next, we examined the indirect effects of agent condition (human vs. AI) mediated by quantitative and qualitative attributes, and the difference between the two outcome conditions (unfavorable vs. favorable) on these indirect effects. Specifically, we focused on understanding the mediating roles of both quantitative and qualitative attributes in relation to the effects of the experimental conditions on the outcome variables. This analysis was conducted for all four outcomes.

A similar pattern was evident across all dependent variables, showing robustness of findings. When the outcome was unfavorable, the effects of agent on perceived fairness, competence, trust, and usefulness were mediated by the perceived levels of the agent's considerations of the job candidate's quantitative and qualitative attributes. However, when the outcome was favorable, the effect of agent was only mediated by the degree of qualitative attributes consideration.

Specifically, the confidence intervals for the conditional indirect effects on all four dependent variables (fairness, competency, trust, usefulness) via quantitative attributes in the unfavorable outcome condition did not include zero (fairness: 0.54, CI [0.23, 0.88]; competency: 0.51, [0.21, 0.90]; trust: 0.44, CI [0.17, 0.79]; usefulness: 0.40, CI [0.17, 0.71]) whereas for the favorable outcome condition, the confidence intervals included zero (fairness: −0.03, CI [−0.20, 0.12]; competency: −0.03, [−0.20, 0.12]; trust: −0.03, CI [−0.16, 0.11]; usefulness: −0.03, CI [−0.16, 0.10]). This means that the level of quantitative attributes consideration mediates the effect of agent on perception when the decision outcome is deemed unfavorable but not when the outcome is favorable. A formal test of moderated mediation (i.e., index of moderated mediations) indicated that these conditional indirect effects were significantly differ from each other (fairness: −0.57, CI [−0.96, −0.23]; competency: −0.55, CI [−0.97, −0.20]; trust: −0.47, CI [−0.85, −0.17]; usefulness: −0.43, CI [−0.77, −0.15]).

Next, we assessed the conditional indirect effects via qualitative attributes, which were found to be statistically significant for both unfavorable outcome conditions across all dependent variables. In the unfavorable outcome condition, the effects were as follows: fairness (0.48, CI [0.18, 0.80]), competency (0.59, CI [0.24, 0.97]), trust (0.44, CI [0.14, 0.78]), usefulness (0.44, CI [0.18, 0.73]). For the favorable outcome condition, the results were: fairness (0.30, CI [0.10, 0.52]), competency (0.36 [0.13, 0.65]), trust (0.27, CI [0.08, 0.50]), usefulness (0.27, CI [0.09, 0.50]), This indicates the degree of qualitative attributes consideration mediates the effect of agent (human vs. AI) on perceptions under both outcome conditions. However, the index of moderated mediation was statistically significant only for the fairness perception (−0.19, CI [−0.43, −0.01]). This suggests that while the indirect effects of agent on competency (−0.23, CI [−0.51, 0.01]), trust (−0.17, CI [−0.40, 0.00]), and usefulness (−0.17, CI [−0.36, 0.00]) through the qualitative attributes consideration are present, these effects do not significantly differ across two outcome conditions.

The results show that when decision-making outcome was favorable, the effect of decision-making agent on various perceptions was mediated through perceived level of consideration of qualitative attributes of the job applicant (RQ4a). Further, ADM had indirect positive effects mediated through qualitative attributes. In brief, participants perceived an algorithmic agent fairer, more trustworthy, competent, and useful than a human agent because they perceived the algorithmic agent considered qualitative attributes in the decision process, more than the human agent.

In the unfavorable condition, as the indirect effects were significant on both mediators, the effects of decision-making agent on perceptions were mediated through decision-makers consideration of both quantitative and qualitative attributes considerations (RQ4b). The significant indirect effects mean that fairness, trust, competency, and

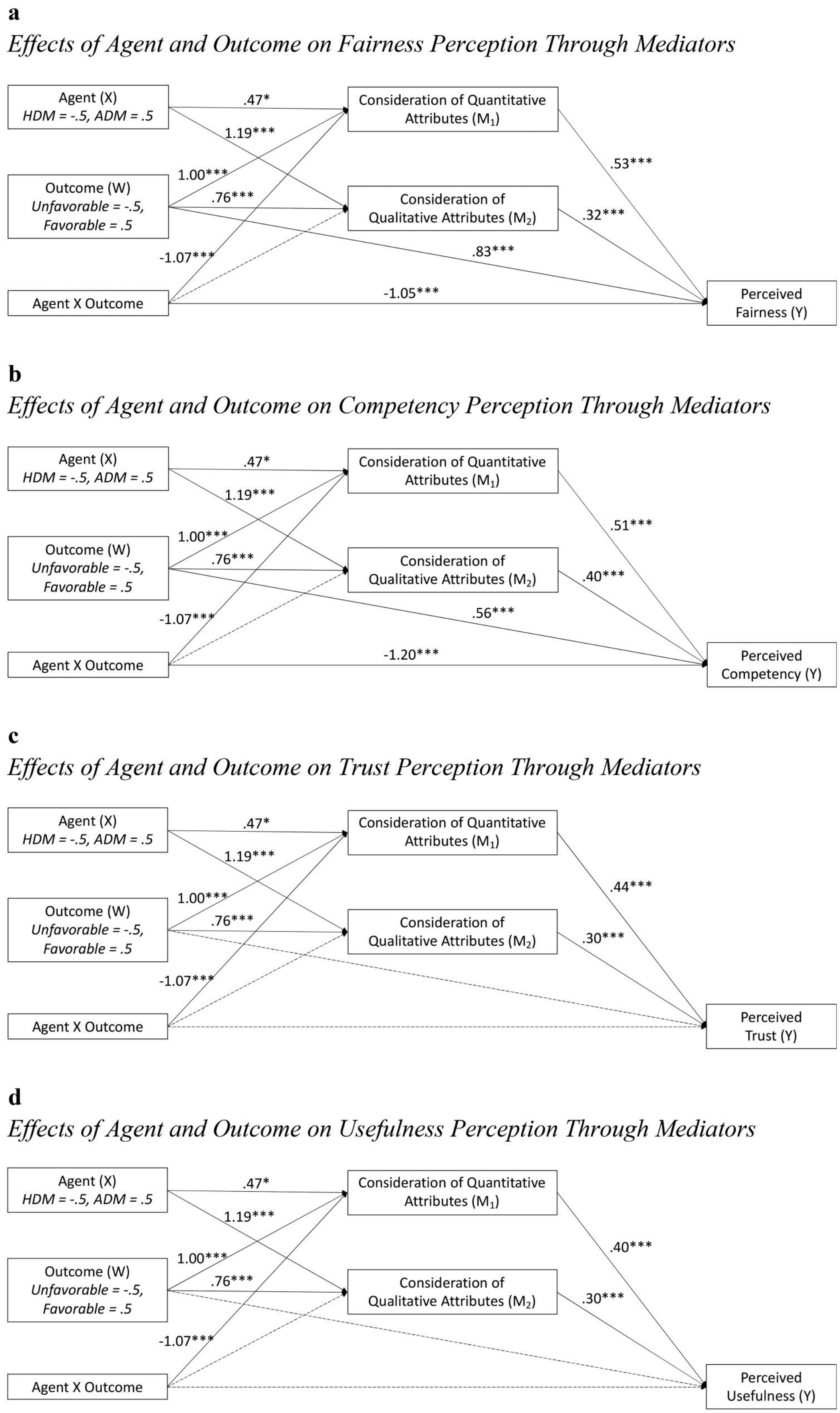

Figure 4. (a) Effects of agent and outcome on fairness perception through Mediators. (b) Effects of agent and outcome on competency perception through Mediators. (c) Effects of agent and outcome on trust perception through Mediators. (d) Effects of agent and outcome on usefulness perception through Mediators. $^*p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$, non-significant paths are depicted with dashed lines.

usefulness perceptions increased in the ADM condition, and this increase was mediated by consideration of both quantitative and qualitative attributes by the agent during the evaluation.

Combined, our findings support the algorithmic appreciation account and provide evidence of stereotypical thinking that algorithms have a great potential in making use of large amount of data that are quantifiable. However, the basis of algorithmic aversion, perceived lack of qualitative assessment in ADM, was not observed in our study. Contrary to expectations, people perceived that an AI agent engages in a qualitative assessment of a job candidate to a greater extent than a human agent, which resulted in more appreciative perceptions toward AI-powered decisions.

4. Discussion

The present study investigates perceptions of algorithmic decisions in job evaluation outcomes. In doing so, it addresses questions about the widespread deployment of AI in evaluating the potential of other humans. Specifically, algorithmic decisions and human decisions were compared on perceptions of fairness, trust, competency, and usefulness in an experiment involving a hypothetical job fit evaluation service. Furthermore, this study examined the perception of the agent based on the favorability of the outcome. To explain the pattern of findings, we explored the role of two potential mediators, quantitative (i.e., objective) and qualitative (i.e., subjective) indicators of a candidate's qualifications. Our study responds to urgent calls for research on the conditionality of perceptual biases toward ADM and explores the possible cognitive underpinnings of algorithmic appreciation and aversion in employment processes (Köchling & Wehner, 2020).

4.1. Summary of findings

Overall, the results were in line with studies in algorithmic appreciation rather than algorithm aversion in job application decision-making. Specifically, algorithmic decisions received higher ratings of perceived fairness, trust, competency, and usefulness in comparison to human decisions. These observations resonate with the generally optimistic public sentiment toward AI as reported in recent surveys. For instance, Zhang and Dafoe (2019) found that more Americans support AI-driven technologies than oppose them, while Kieslich et al. (2023) observed minimal concerns about AI-driven technologies among the majority of the German population. Similarly positive perceptions of AI have been reported across various countries, but the degree of appreciation can vary based on demographic factors, knowledge levels about AI, and cultural contexts (Fietta et al., 2022; Kelley et al., 2021).

The appreciation of ADM could be attributed to the growing popularity of AI. Prior research suggests that increased familiarity with technology can reduce related apprehensions (Gefen, 2000). Furthermore, while Lee (2018) demonstrated a tendency for favoring human decision-makers in tasks requiring human skills, our findings indicate a possible transition in this perception. This divergence may potentially be attributed to the advancements and growing ubiquity of AI technologies since 2018, which could have influenced people's perceptions of AI-driven technologies' abilities to manage tasks traditionally considered to be the exclusive domain of humans. Of particular relevance is the focus of our study on fit assessments, which might be viewed as more objective and quantifiable, thereby making them potentially more compatible with algorithmic evaluation, as compared to the resume review tasks explored in the work of Lee (2018). In addition, it is important to note that perceptions of AI-driven technologies are not universally positive. In Bao et al. (2022)'s survey, approximately one-third of respondents, categorized as the "negative attitudes" group, perceived more risks than benefits associated with AI applications.

Our findings can be interpreted within the context of algorithmic trustworthiness, which often hinges on transparency and objectivity. In our study, we offered participants information about how the algorithms were trained and how they made decisions. This transparency might have fostered positive impressions of ADM. Participants may perceive algorithms as being free of the personal biases that humans may inadvertently introduce in decision-making processes—particularly relevant in our study's context, where decisions about individuals were made. Given that algorithms operate based on pre-set rules and data, ideally devoid of subjective interpretations, participants may perceive them as more equitable and fairer, leading to higher ratings in these categories. However, trust and positive perceptions can be jeopardized by a single encounter with an errant algorithm (Dietvorst et al., 2015). These findings underscore the importance of ongoing research into public perceptions of AI and ADM, topics that are constantly evolving in response to changing experiences and advancements in technology. For example, the recent development of generative AI models, which demonstrate remarkable capabilities in handing qualitative information, could potentially influence public perceptions of AI, thereby adding a new layer of complexity to our understanding of how people perceive and react to ADM.

Next, we found a robust effect of outcome favorability on both ADM and HDM conditions. When the outcome was favorable to a job candidate, people perceived the agent as fairer, more trustworthy, competent, and useful. The outcome favorability bias was more notable in the HDM condition than the ADM condition; people exhibited greater reactance to the unfavorable outcome from the human coordinator than the AI agent in the job application scenario. When the agent generated an unfavorable outcome, cognitive discomfort may have occurred. Under such circumstances, people appeared to disqualify the human coordinator, evaluating them as less capable of thorough quantitative and qualitative assessments. Yet, when AI delivered the same negative outcome, it only resulted in a small discounting of the agent's evaluative capability.

The minimal reactance to an unfavorable outcome offered by an algorithm points to an interesting conditional difference between ADM and HDM. One possible explanation for this difference could be the perceived objectivity and impartiality of AI algorithms. When an outcome is unfavorable, people might attribute it to the human decision-maker's subjectivity or personal biases, leading to more negative evaluations of the human coordinator. In contrast, AI-driven algorithms may be perceived as more consistent and unbiased, resulting in less negative evaluation even when the outcome is unfavorable. However, when the outcome is favorable, both human and AI decision-makers receive high evaluations, which might be attributed to a ceiling effect as participants in both conditions scored near the possible upper end of the scale. Interestingly, our finding contrasts with that of Yalcin et al. (2022), who found that consumers react less positively to favorable decisions made by algorithms as compared to those made by humans. Our study, however, found no difference in the positive reactions towards favorable decisions made by either decision-maker. This discrepancy calls for further investigation into the factors that influence consumers' perceptions of algorithmic decisions, especially in the context of favorable outcomes.

Finally, the mediation analyses demonstrated the significant role of quantitative and qualitative assessments in explaining the effect of the decision-making agent. The appreciative perceptions of the job evaluation was mediated by the levels of quantitative and qualitative assessments of the job candidate, while conditional to the outcome of the evaluation.

4.2. Theoretical implications

Our results suggest that appreciation of algorithmic advice is more prominent than algorithmic aversion in the job applicant evaluation context. With the vast strides in AI, as well as ADM's proliferation in our daily life, our experience and perceptions of algorithmic decisions have constantly been evolving. Unlike what past studies found, participants of the current study did not necessarily view ADM as reductionistic and inferior in qualitative assessments. As familiarity contributes to experiential trust, people's accumulated experience with algorithms may have led to high trust in its capabilities and appreciation in AI-powered decisions. Yet, this can raise a normative concern related to promoting an appropriate level of trust. When trust exceeds the real capabilities of the system (i.e., overtrust), it can lead to misuse and overreliance on the system (Lee & See, 2004). Indeed, the risks associated with overtrust in AI systems are highlighted in recent studies. Krügel et al.(2022) stressed the necessity of improving digital literacy to prevent decision-making from being manipulated by flawed algorithms. A similar concern was echoed by Robinette et al. (2016), whose experiment revealed that participants in an emergency situation followed a robot's directions despite it demonstrating poor performance prior. These findings underscore the importance of promoting an appropriate level of trust in AI systems to avoid misuse and overreliance.

We also found that the degree of impact of outcome favorability bias was not equal. When the outcome of decision-making was deemed positive or favorable, there was not much difference in how people perceive algorithmic decisions and human decisions, while unfavorable outcomes elicited different responses. This can be due to the fundamental differences in motivation when we interact with humans and machines. Human-human communication is guided by the actors' motivations and intentionality. However, machines do not have consciousness. Thus, the anthropocentric theory of mind that we employ to interpret and make sense of human action is not directly applicable to human-machine communication. In particular, the theories of human communication that are closely connected to motivated cognition in a social context (e.g., social desirability bias) may not guide human interactions with machines or AI. For instance, mechanisms related to self-presentation and self-protection might be less prominent in human-AI communication. This highlights the significance of developing unique theories of human-machine interactions (Logg et al., 2019) – and the theory needs to be expanded and updated with the development in sociotechnical features of the machine.

4.3. Practical implications

With AI-driven technologies becoming more pervasive in our daily lives, this study offers practical implications for designing and deploying AI-driven decision-making in the HR context. Yet, while our study was conducted specifically in the context of hiring processes and therefore speaks most directly to that context, the relevance of our findings is not necessarily limited to such a narrow context. In futures wherein ADM is pervasive, the findings we present here may well speak to people's perceptions of algorithmic decisions and their relationships to human decisions in general.

Empirical comparison between an HDM agent and an AI-driven ADM agent allowed us to identify conditions where ADM is more appreciated than HDM. From a human-centered perspective, deeper knowledge of these conditions can help practitioners to decide which decision-making processes could be responsibly allocated to AI. That is, to decide when AI-powered decisions might be beneficial for end-users who are loath to accept criticism or negative outcomes from other humans. For example, our findings indicate when the decision-outcome is negative, people might be more accepting of feedback from AI than a human. Thus, ADM can be useful in situations where reducing unnecessary reactance results in better affective user outcomes.

But the willingness to accept an unfavorable outcome or judgement from an algorithm requires careful examination. There is the potential danger of placing too much faith in an algorithm without critical evaluation of the decision. Acquiescing to a decision by AI could be akin to digital resignation in the privacy domain (Draper & Turow, 2019), that has allowed corporations to exploit personal data. Further, it may mirror the normalization of affective discomfort in technology use (Seberger et al., 2021, 2022). As ADM becomes pervasive, the public must be trained to be discriminating in their reliance on algorithms and to be able to tell the difference between algorithm appreciation and the learned helplessness of algorithmic resignation or acquiescence. At the same time, willingness to heed the advice of an algorithm can be beneficial in certain circumstances. Future research, outreach, and policy must focus on training the public to be discerning consumers of algorithmic decisions.

The traditional understanding that ADM is reductionistic and therefore fails to adequately consider broader contexts was not evident in this study. With the increasing penetration of AI-driven technology in all aspects of life and its use of natural language processing, AI-driven ADM was believed to account for the soft skills of job applicants to a greater degree for the fit judgment. Especially, given the recent advancements in generative AI, the understanding of ADM being reductionistic and often failing to adequately consider broader contexts is being challenged. Again, while this may be a positive sign of algorithm appreciation, given the challenges of AI in understanding context and nuance, we may be rushing to place our faith in technology before it has truly reached that level of maturity.

Thus, the design of ADM agents should focus on people's perceptions of such agents *in addition to* improving such agents' performance in accounting for contextual information. Our findings suggest that deeper and more nuanced achievement of functional requirements will likely influence the perception and adoption of AI-powered ADM systems. These findings also reinforce the need for further research to identify tasks to which ADM systems are best suited versus tasks that are ultimately better suited to HDM.

4.4. Limitations and directions for future research

The findings of the present study come with limitations, some of which can be addressed in future research. First, although conducting an experiment allowed us to infer causality with manipulations, it is possible that an online experiment limits the external validity, and therefore generalizability, of our findings. As with any socially oriented experiment, the subsequent application of findings should be undertaken with care, particularly regarding alignment between sample populations and the role that online environments play in creating the contexts of algorithmic decision-making in hiring processes. The current study focuses on a relatively low-risk context that did not directly affect the individual. The results could be different in high-stakes decision-making. Also, given the variety of decision-making contexts where ADM can be adopted, the findings should be replicated in future studies that involve various applications of ADM.

Second, we included only two decision-making agent conditions in our study. Several methodological adjustments could be made in future iterations. In our manipulation check for the ADM condition, we observed that aside from a significant number of participants who believed that the decision was made by AI, many also felt that there was human involvement in the decision-making process. In contrast, in the HDM condition, the majority of participants did not believe AI was involved (refer to Table 2). Given the high level of agreement among participants of human involvement in the ADM condition, a stronger manipulation might be necessary in future work. Similarly, future work could explore a collaborative state in which algorithms and humans co-engage to make a decision. Moreover, in our study, the human agent was someone people did not have previous experience with, so they may have relied heavily on the outcome cue to infer the agent's capability. This could result in a preference toward an objective algorithm rather than an uncertain, subjective human. If the human agent is deemed a credible source of information or experts, the effect of agent can be different. Thus, future studies should diversify the types and characteristics of the decision-making agent. In addition, people's reactions toward algorithmic decisions might differ depending on who is/are subject to the decision. Suppose study participants are subjected to a favorable or unfavorable decision made by AI (e.g., Wang et al., 2020). In that case, they may engage in deeper cognitive reasoning to make sense of the outcome compared to when the same decision was made for others. Thus, the subject(s) of algorithmic decision will be an interesting factor to explore in future research.

Third, we explored only quantitative and qualitative attributes as two possible explanations or mediators of algorithmic appreciation. Future studies should explore boundary conditions of optimism or skepticism of AI in various tasks. Moreover, other potential underlying cognitive mechanisms or emotions could be in play in shaping perceptions around algorithmic decisions that are worth investigating.

Fourth, while our study has provided valuable insights into the influence of agent type and outcome favorability on fairness, competency, trust, and usefulness, a limitation lies in our treatment of these variables as discrete entities. Our research did uncover interdependencies among these constructs, with significant positive correlations observed: between fairness and competence ($r = 0.86$, p < 0.001), fairness and trust ($r = 0.78$, $p < 0.001$), fairness and usefulness ($r = 0.75$, p < 0.001), competence and trust ($r = 0.83$, p < 0.001), competence and usefulness ($r = 0.65$, p < 0.001), trust and usefulness ($r = 0.78$, p < 0.001). These results indicate a strong interconnectedness among fairness, competence, trust, and usefulness, as confirmed by a high Cronbach's alpha of 0.93. Among the variables we examined, trust has been particularly emphasized in recent research as a key outcome. Numerous studies have highlighted its significance (e.g., Choung et al., 2022b), pointing to various factors that influence trust in AI technologies (Lockey et al., 2021; Rheu et al., 2021). However, the intricate relationships between these variables remain to be fully explored. For instance, our study found a significant relationship between usefulness and trust. In contrast, Pitardi and Marriott (2021) suggested that usefulness might not be the primary driver of trust in AI voice assistants. They instead pointed to other social attributes, notably social cognition, as more influential. With these collective insights in view, future research should consider further theorizing on the intricate relationships among the variables we have highlighted in this manuscript.

In addition to these directions, as generative models such as OpenAI's ChatGPT and Google's Bard are becoming more popular, the impacts of generative AI on people's perceptions of AI's ability to handle qualitative and contextual information warrant further investigation. Future studies could explore how the capabilities of these advanced models influence perceptions of fairness, trust, and competency in ADM, as well as the potential for a shift from mere algorithm appreciation to algorithm reliance. In sum, this rapidly evolving landscape of AI technologies offers a wealth of opportunities for future research.

5. Conclusion

This paper presents the results of an online experiment that investigated people's perceptions of ADM and HDM in a job application/hiring process. Using a conditional approach, differences were examined for favorable and unfavorable outcomes. When the outcome was favorable, there were few differences between ADM and HDM. However, when the outcome was unfavorable, participants expressed disapproval of human decisions, but were less disapproving of the same decisions offered by AI. These findings point to algorithmic appreciation rather than algorithmic aversion. While positive perceptions of algorithms bode well for the continued proliferation of AI, acceptance of algorithms must be tempered through literacy and ongoing discussion to find the right balance between human and machine agency. Our findings highlight the need for contextually specific new theories of human-machine interaction, in which both the human and AI are endowed with agency.

Notes

1. AI refers to "a machine-based system that can, for a given set of human-defined objectives, make predictions, recommendations, or decisions influencing real or virtual environments (OECD, 2019, pp. 23–24)."
2. In our study, the exclusion criteria were determined balancing data quality and the need for an adequate sample size. Although excluding all participants who failed any attention test could improve data quality, it could also decrease our sample size, possibly impacting the statistical power of our analyses. We accounted for the potential for human error - the possibility that even attentive participants might occasionally misread or misclick, leading to an error in a single attention test. Therefore, we decided to exclude only those participants who failed both attention tests, indicative of a greater degree of inattentiveness. Out of the total number of participants, 71 failed only one attention test.
3. In comparison with the 2019 US Census data, our sample exhibited certain disparities. Notably, Black participants were underrepresented in our study, constituting only 4.2% of our sample as compared to 13% in the US Census data. Additionally, our sample was more educated, with a significantly higher proportion of participants holding a bachelor's degree or higher (72.8%), far exceeding the equivalent US census figure of 42%. Lastly, the median household income range in our sample was lower ($40,000–$50,000) compared to the national median income according to census data ($69,021).
4. Given that most of the participants evaluated that the candidate was a good fit for the advertised position ($M = 78.08$ $SD = 17.34$), associating high and low fit to scores of 90 and 40 seemed appropriate.
5. Model 8 is a moderated mediation model that permits us to assess whether the indirect effects of our independent variable on the dependent variables, via a mediator, alter across different levels of a moderator. This model was chosen for our analysis as it enables us to investigate not only the mediation effect, but also the potential variations in this effect contingent on the decision-making agent and the favorability of the outcome.

Disclosure statement

No potential conflict of interest was reported by the author(s).

ORCID

Hyesun Choung http://orcid.org/0000-0001-9464-0399
John S. Seberger http://orcid.org/0000-0002-3101-3686
Prabu David http://orcid.org/0000-0002-5096-1016

Data availability statement

The data that support the findings of this study are available from the corresponding author, HC, upon request.

References

Alexander, V., Blinder, C., & Zak, P. J. (2018). Why trust an algorithm? Performance, cognition, and neurophysiology. *Computers in Human Behavior*, *89*(2), 279–288. https://doi.org/10.1016/j.chb.2018.07.026

Araujo, T., Helberger, N., Kruikemeier, S., & de Vreese, C. H. (2020). In AI we trust? Perceptions about automated decision-making by artificial intelligence. *AI & Society*, *35*(3), 611–623. https://doi.org/10.1007/s00146-019-00931-w

Bao, L., Krause, N. M., Calice, M. N., Scheufele, D. A., Wirz, C. D., Brossard, D., Newman, T. P., & Xenos, M. A. (2022). Whose AI? How different publics think about AI and its social impacts. *Computers in Human Behavior*, *130*(4), 107182. https://doi.org/10.1016/j.chb.2022.107182

Bogert, E., Schecter, A., & Watson, R. T. (2021). Humans rely more on algorithms than social influence as a task becomes more difficult. *Scientific Reports*, *11*(1), 8028. https://doi.org/10.1038/s41598-021-87480-9

Castelo, N., Bos, M. W., & Lehmann, D. R. (2019). Task-dependent algorithm aversion. *Journal of Marketing Research*, *56*(5), 809–825. https://doi.org/10.1177/0022243719851788

Cheng, M. M., & Hackett, R. D. (2021). A critical review of algorithms in HRM: Definition, theory, and practice. *Human Resource Management Review*, *31*(1), 100698. https://doi.org/10.1016/j.hrmr.2019.100698

Choung, H., David, P., & Ross, A. (2022a). Trust and ethics in AI. *AI & Society*, *38*(2), 733–745. https://doi.org/10.1007/s00146-022-01473-4

Choung, H., David, P., & Ross, A. (2022b). Trust in AI and its role in the acceptance of AI technologies. *International Journal of Human–Computer Interaction*, *39*(9), 1727–1739. https://doi.org/10.1080/10447318.2022.2050543

Dastin, J. (2018, October 10). Amazon scraps secret AI recruiting tool that showed bias against women. *Reuters*. https://www.reuters.com/article/us-amazon-com-jobs-automation-insight-idUSKCN1MK08G

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, *13*(3), 319. https://doi.org/10.2307/249008

Deutsch, M. (1975). Equity, equality, and need: What determines which value will be used as the basis of distributive justice? *Journal of Social Issues*, *31*(3), 137–149. https://doi.org/10.1111/j.1540-4560.1975.tb01000.x

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. https://doi.org/10.1037/xge0000033

Draper, N. A., & Turow, J. (2019). The corporate cultivation of digital resignation. *New Media & Society*, *21*(8), 1824–1839. https://doi.org/10.1177/1461444819833331

Edwards, A., & Edwards, C. (2022). Does the correspondence bias apply to social robots?: Dispositional and situational attributions of human versus robot behavior. *Frontiers in Robotics and AI*, *8*, 788242. https://doi.org/10.3389/frobt.2021.788242

Engler, A. (2021). We're using AI more and more for hiring. *How can we ensure it's fair?* World Economic Forum. https://www.weforum.org/agenda/2021/06/auditing-employment-algorithms-discrimination-fairness/

Epstein, Z., Levine, S., Rand, D. G., & Rahwan, I. (2020). Who gets credit for AI-generated art? *iScience*, *23*(9), 101515. https://doi.org/10.1016/j.isci.2020.101515

Fenneman, A., Sickmann, J., Pitz, T., & Sanfey, A. G. (2021). Two distinct and separable processes underlie individual differences in algorithm adherence: Differences in predictions and differences in trust thresholds. *PLOS One*, *16*(2), e0247084. https://doi.org/10.1371/journal.pone.0247084

Fietta, V., Zecchinato, F., Stasi, B. D., Polato, M., & Monaro, M. (2022). Dissociation between users' explicit and implicit attitudes toward artificial intelligence: An experimental study. *IEEE Transactions on Human-Machine Systems*, *52*(3), 481–489. https://doi.org/10.1109/THMS.2021.3125280

Gefen, D. (2000). E-commerce: The role of familiarity and trust. *Omega*, *28*(6), 725–737. https://doi.org/10.1016/S0305-0483(00)00021-9

Gerber, J. P., Wheeler, L., & Suls, J. (2018). A social comparison theory meta-analysis 60+ years on. *Psychological Bulletin*, *144*(2), 177–197. https://doi.org/10.1037/bul0000127

Glikson, E., & Woolley, A. W. (2020). Human trust in artificial intelligence: Review of empirical research. *Academy of Management Annals*, *14*(2), 627–660. https://doi.org/10.5465/annals.2018.0057

Gonzalez, M. F., Liu, W., Shirase, L., Tomczak, D. L., Lobbe, C. E., Justenhoven, R., & Martin, N. R. (2022). Allying with AI? Reactions toward human-based, AI/ML-based, and augmented hiring processes. *Computers in Human Behavior*, *130*(4), 107179. https://doi.org/10.1016/j.chb.2022.107179

Grace, K., Salvatier, J., Dafoe, A., Zhang, B., & Evans, O. (2018). Viewpoint: When Will AI exceed human performance? Evidence from AI experts. *Journal of Artificial Intelligence Research*, *62*, 729–754. https://doi.org/10.1613/jair.1.11222

Hayes, A. F. (2018). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach.* (2nd ed.). Guilford Press.

Helberger, N., Araujo, T., & de Vreese, C. H. (2020). Who is the fairest of them all? Public attitudes and expectations regarding automated decision-making. *Computer Law & Security Review*, *39*(1), 105456. https://doi.org/10.1016/j.clsr.2020.105456

Hunkenschroer, A. L., & Luetge, C. (2022). Ethics of AI-enabled recruiting and selection: A review and research agenda. *Journal of Business Ethics*, *178*(4), 977–1007. https://doi.org/10.1007/s10551-022-05049-6

Jago, A. S. (2019). Algorithms and authenticity. *Academy of Management Discoveries*, *5*(1), 38–56. https://doi.org/10.5465/amd.2017.0002

Kelley, P. G., Yang, Y., Heldreth, C., Moessner, C., Sedley, A., Kramm, A., Newman, D. T., & Woodruff, A. (2021). *Exciting, Useful, Worrying, Futuristic: Public Perception of Artificial Intelligence in 8 Countries* [Paper presentation]. Proceedings of the 2021 AAAI/ACM Conference on AI, Ethics, and Society (pp. 627–637). ACM. https://doi.org/10.1145/3461702.3462605

Kieslich, K., Lünich, M., & Došenović, P. (2023). Ever heard of ethical AI? Investigating the salience of ethical AI issues among the German population. *International Journal of Human–Computer Interaction*, 1–14. https://doi.org/10.1080/10447318.2023.2178612

Köchling, A., & Wehner, M. C. (2020). Discriminated by an algorithm: A systematic review of discrimination and fairness by algorithmic decision-making in the context of HR recruitment and HR development. *Business Research*, *13*(3), 795–848. https://doi.org/10.1007/s40685-020-00134-w

Köchling, A., & Wehner, M. C. (2023). Better explaining the benefits why AI? Analyzing the impact of explaining the benefits of AI-supported selection on applicant responses. *International Journal of Selection and Assessment*, *31*(1), 45–62. https://doi.org/10.1111/ijsa.12412

Krafft, T. D., Zweig, K. A., & König, P. D. (2022). How to regulate algorithmic decision-making: A framework of regulatory requirements for different applications. *Regulation & Governance, Rego*, *16*(1), 119–136. https://doi.org/10.1111/rego.12369

Krügel, S., Ostermaier, A., & Uhl, M. (2022). Zombies in the loop? Humans trust untrustworthy AI-advisors for ethical decisions. *Philosophy & Technology*, *35*(1), 17. https://doi.org/10.1007/s13347-022-00511-9

Kuehn, K. (2018, September 6). *re:Work—Google's algorithm-powered internal job marketplace.* https://rework.withgoogle.com/blog/googles-algorithm-powered-internal-job-marketplace/

Langer, M., & Landers, R. N. (2021). The future of artificial intelligence at work: A review on effects of decision automation and augmentation on workers targeted by algorithms and third-party observers. *Computers in Human Behavior, 123*, 106878. https://doi.org/10.1016/j.chb.2021.106878

Lee, C., & Cha, K. (2023). FAT-CAT—Explainability and augmentation for an AI system: A case study on AI recruitment-system adoption. *International Journal of Human-Computer Studies*, *171*(C), 102976. https://doi.org/10.1016/j.ijhcs.2022.102976

Lee, J. D., & See, K. A. (2004). Trust in automation: Designing for appropriate reliance. *Human Factors*, *46*(1), 50–80. https://doi.org/10.1518/hfes.46.1.50_30392

Lee, M. K. (2018). Understanding perception of algorithmic decisions: Fairness, trust, and emotion in response to algorithmic management. *Big Data & Society*, *5*(1), 205395171875668. https://doi.org/10.1177/2053951718756684

Lee, M. K., Jain, A., Cha, H. J., Ojha, S., & Kusbit, D. (2019). Procedural justice in algorithmic fairness: Leveraging transparency and outcome control for fair algorithmic mediation. *Proceedings of the ACM on Human-Computer Interaction*, *3*(CSCW), 1–26. https://doi.org/10.1145/3359284

Leventhal, G. S. (1980). What should be done with equity theory? In K. J. Gergen, M. S. Greenberg, & R. H. Willis (Eds.), *Social exchange* (pp. 27–55). Springer US. https://doi.org/10.1007/978-1-4613-3087-5_2

Lockey, S., Gillespie, N., Holm, D., & Someh, I. A. (2021). *A review of trust in artificial intelligence: Challenges, vulnerabilities and future directions* [Paper presentation]. Hawaii International Conference on System Sciences. https://doi.org/10.24251/HICSS.2021.664

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*(10), 90–103. https://doi.org/10.1016/j.obhdp.2018.12.005

Lünich, M., & Kieslich, K. (2022). Exploring the roles of trust and social group preference on the legitimacy of algorithmic decision-making vs. Human decision-making for allocating COVID-19 vaccinations. *AI & SOCIETY*. https://doi.org/10.1007/s00146-022-01412-3

Mahmud, H., Islam, A. K. M. N., Ahmed, S. I., & Smolander, K. (2022). What influences algorithmic decision-making? A systematic literature review on algorithm aversion. *Technological Forecasting and Social Change*, *175*(49), 121390. https://doi.org/10.1016/j.techfore.2021.121390

Marcinkowski, F., Kieslich, K., Starke, C., & Lünich, M. (2020). *Implications of AI (un-)fairness in higher education admissions: The effects of perceived AI (un-)fairness on exit, voice and organizational reputation* [Paper presentation]. Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency (pp. 122–130). https://doi.org/10.1145/3351095.3372867

Mcknight, D. H., Carter, M., Thatcher, J. B., & Clay, P. F. (2011). Trust in a specific technology: An investigation of its components and measures. *ACM Transactions on Management Information Systems*, *2*(2), 1–25. https://doi.org/10.1145/1985347.1985353

McLean, G., & Osei-Frimpong, K. (2019). Hey Alexa … examine the variables influencing the use of artificial intelligent in-home voice assistants. *Computers in Human Behavior*, *99*, 28–37. https://doi.org/10.1016/j.chb.2019.05.009

Newman, D. T., Fast, N. J., & Harmon, D. J. (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. https://doi.org/10.1016/j.obhdp.2020.03.008

Niszczota, P., & Kaszás, D. (2020). Robo-investment aversion. *PLOS One*, *15*(9), e0239277. https://doi.org/10.1371/journal.pone.0239277

OECD. (2019). *Artificial intelligence in society*. OECD Publishing. https://doi.org/10.1787/eedfee77-en

Pitardi, V., & Marriott, H. R. (2021). Alexa, she's not human but… Unveiling the drivers of consumers' trust in voice-based artificial intelligence. *Psychology & Marketing*, *38*(4), 626–642. https://doi.org/10.1002/mar.21457

Prahl, A., & Van Swol, L. (2017). Understanding algorithm aversion: When is advice from automation discounted? *Journal of Forecasting*, *36*(6), 691–702. https://doi.org/10.1002/for.2464

Rahwan, I., Cebrian, M., Obradovich, N., Bongard, J., Bonnefon, J.-F., Breazeal, C., Crandall, J. W., Christakis, N. A., Couzin, I. D., Jackson, M. O., Jennings, N. R., Kamar, E., Kloumann, I. M., Larochelle, H., Lazer, D., McElreath, R., Mislove, A., Parkes, D. C., Pentland, A. '., … Wellman, M. (2019). Machine behaviour. *Nature*, *568*(7753), 477–486. https://doi.org/10.1038/s41586-019-1138-y

Rheu, M., Shin, J. Y., Peng, W., & Huh-Yoo, J. (2021). Systematic review: Trust-building factors and implications for conversational agent design. *International Journal of Human–Computer Interaction*, *37*(1), 81–96. https://doi.org/10.1080/10447318.2020.1807710

Robinette, P., Li, W., Allen, R., Howard, A. M., & Wagner, A. R. (2016). *Overtrust of robots in emergency evacuation scenarios* [Paper presentation]. 2016 11th ACM/IEEE International Conference on Human-Robot Interaction (HRI) (pp. 101–108). IEEE. https://doi.org/10.1109/HRI.2016.7451740

Schroth, H. A., & Shah, P. P. (2000). Procedures: Do we really want to know them? An examination of the effects of procedural justice on self-esteem. *The Journal of Applied Psychology*, *85*(3), 462–471. https://doi.org/10.1037/0021-9010.85.3.462

Seberger, J. S., Llavore, M., Wyant, N. N., Shklovski, I., & Patil, S. (2021). *Empowering resignation: There's an app for that* [Paper presentation]. Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems, (pp. 1–18). CHI. https://doi.org/10.1145/3411764.3445293

Seberger, J. S., Shklovski, I., Swaiatek, E., & Patil, S. (2022). *Still creepy after all these years: The normalization of affective discomfort in app use. P* [Paper presentation]. Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems (pp. 1–19). CHI. https://doi.org/10.1145/3491102.3502112

Shepperd, J., Malone, W., & Sweeny, K. (2008). Exploring causes of the self-serving bias. *Social and Personality Psychology Compass*, *2*(2), 895–908. https://doi.org/10.1111/j.1751-9004.2008.00078.x

Shin, D. (2021). The effects of explainability and causability on perception, trust, and acceptance: Implications for explainable AI. *International Journal of Human-Computer Studies*, *146*(83), 102551. https://doi.org/10.1016/j.ijhcs.2020.102551

Spatola, N., & Normand, A. (2021). Human vs. machine: The psychological and behavioral consequences of being compared to an outperforming artificial agent. *Psychological Research*, *85*(3), 915–925. https://doi.org/10.1007/s00426-020-01317-0

Starke, C., Baleis, J., Keller, B., Marcinkowski, F. (2021). Fairness perceptions of algorithmic decision-making: A systematic review of the empirical literature. *ArXiv:2103.12016* [*Cs*]. http://arxiv.org/abs/2103.12016

Sundar, S. S., & Kim, J. (2019). *Machine heuristic: When we trust computers more than humans with our personal information* [Paper presentation]. Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems – CHI '19, (pp. 1–9). https://doi.org/10.1145/3290605.3300768

Thurman, N., Moeller, J., Helberger, N., & Trilling, D. (2019). My friends, editors, algorithms, and I: Examining audience attitudes to news selection. *Digital Journalism*, *7*(4), 447–469. https://doi.org/10.1080/21670811.2018.1493936

Tigre Moura, F., & Maw, C. (2021). Artificial intelligence became Beethoven: How do listeners and music professionals perceive artificially composed music? *Journal of Consumer Marketing*, *38*(2), 137–146. https://doi.org/10.1108/JCM-02-2020-3671

van Esch, P., Black, J. S., & Arli, D. (2021). Job candidates' reactions to AI-Enabled job application processes. *AI and Ethics*, *1*(2), 119–130. https://doi.org/10.1007/s43681-020-00025-0

Venkatesh, V., & Bala, H. (2008). Technology acceptance model 3 and a research agenda on interventions. *Decision Sciences, 39*(2), 273–315. https://doi.org/10.1111/j.1540-5915.2008.00192.x

Venkatesh, V., & Davis, F. D. (2000). A theoretical extension of the Technology Acceptance Model: Four longitudinal field studies. *Management Science, 46*(2), 186–204. https://doi.org/10.1287/mnsc.46.2.186.11926

Wang, R., Harper, F. M., & Zhu, H. (2020). *Factors influencing perceived fairness in algorithmic decision-making: Algorithm outcomes, development procedures, and individual differences* [Paper presentation]. CHI '20: Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems (pp. 1–14). https://doi.org/10.1145/3313831.3376813

Wesche, J., & Sonderegger, A. (2021). Repelled at first sight? Expectations and intentions of job-seekers reading about AI selection in job advertisements. *Computers in Human Behavior*, *125*(4), 106931. https://doi.org/10.1016/j.chb.2021.106931

Yalcin, G., Lim, S., Puntoni, S., & Van Osselaer, S. M. J. (2022). Thumbs up or down: Consumer reactions to decisions by algorithms versus humans. *Journal of Marketing Research*, *59*(4), 696–717. https://doi.org/10.1177/00222437211070016

Yeomans, M., Shah, A., Mullainathan, S., & Kleinberg, J. (2019). Making sense of recommendations. *Journal of Behavioral Decision Making*, *32*(4), 403–414. https://doi.org/10.1002/bdm.2118

Zhang, B., & Dafoe, A. (2019). Artificial Intelligence: American Attitudes and Trends. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3312874

About the authors

Hyesun Choung is an Assistant Professor in the Department of Communication at Michigan State University. Her research explores the psychological underpinnings of how individuals perceive and respond to decisions made by AI systems. Her work addresses the social, ethical, and cognitive facets of emerging technologies.

John S. Seberger is an Assistant Professor in Information Science at Drexel University. He conducts interdisciplinary research in humanistic HCI and social informatics. His research specifically addresses questions about privacy, dignity, and agency in the context of the (increasingly digital) infrastructures of daily life.

Prabu David is a Professor of Communication at the Rochester Institute of Technology. He studies human-centered AI, currently focusing on trust and ethics in AI.

Appendix. Experiment materials

1. Job Description and Candidate CV

Job Description

Social Media Summer Internship

We are a multinational corporation looking for a creative Social Media Intern to join our team. As a Social Media Intern, you will be responsible for developing and implementing our social media strategy in order to increase online presence and improve communication efforts. **You will also need to create content as needed and monitor content success and progress.** You will be working closely with our social media department.

Duties and Responsibilities:

- Assist in the development and writing of social media posts.
- Engage with followers on behalf of the company.
- Keep abreast of the latest social media best practices and technologies.

Job Requirements and Qualifications:

- Must be a college sophomore or junior in advertising, digital media, marketing, communication, or a related area.
- Extensive experience and knowledge of social media required (e.g., Facebook, Twitter, LinkedIn, and Instagram).
- Experience as a social media intern in a corporate environment is preferred.
- Great communication and writing skills.
- Familiarity with web content design and publishing.
- Team player and problem-solving skills.
- Good time management skills.
- Attention to detail.

Job Candidate CV

Major & year in college

- Junior in college
- Pursuing Bachelor's degree in Communication

Total credit hours taken toward strategic communication or social media courses

Total 9 credit hours:

- Strategic Communication Campaigns (3 credits)
- Digital Strategic Communication Analytics (3 credits)
- Sports Marketing (3 credits)

Skills related to digital content creation

- Adobe Photoshop and Creative Cloud
- Final Cut Pro and other video editing programs

Previous work experience

- Production, social media intern at the Lansing Lugnuts, a minor league baseball team (Summer, 2019)

Social media accounts and number of connections and post on each account

- Twitter (@dmjerde), 1021 following, 420 followers
- Instagram (@dmjerde11), 120 posts, 30 following, 20 followers
- LinkedIn (www.linkedin.com/in/dominic), 50+ connections

Explain what makes you a good candidate for the position

- I have intricate knowledge of the key social media platforms.
- I have excellent design skills as well as written and verbal communication skills.
- Working with the Lansing Lugnuts helped me to build my confidence and gain valuable experience in social media management. I can contribute to the development and implementation of strategic social media campaigns. I am also capable of working under deadline pressure, and I am self-motivated, making me a good candidate for the position.

Tell us your unique life story that makes you to stand out among other candidates

- I am a sports enthusiast and I consume sports-related content from every possible channel. I shared the final project from my Sports Marketing class on my social media and directly sent the project and my portfolio to the Lansing Lugnuts social media account. Then, my dream came true! They were impressed by my work and proactive attitude and offered me an internship position. In my first few days working with the Lugnuts, I was already using new online tools and software that I had never used before. I quickly learned how to organize tasks and collaborate with team members. It was the most rewarding experience in my life.

Upload your most creative and visually interesting Instagram post

2. Experimental Manipulations

ADM Condition	HDM Condition
Hello! We are developing an online job recommendation platform called JobConnect. JobConnect offers personalized job recommendations for job seekers based on their experience, skills, personality, and even lifestyle. Our platform relies on an **Artificial Intelligent (AI) agent** to evaluate each job applicant's fit for the posted jobs, so job seekers can save their time searching and finding best-fitting jobs. **How does it work?** Job seekers can create their own profiles by simply answering some questions about their interests and skills. Based on the provided information, our agent matches the applicant with job ads and provides a fit score. The fit score is calculated by **our AI algorithm.** JobConnect is free to job seekers! We will show you a brief demonstration of JobConnect and ask you to evaluate the service. On the next page, you will see a job posting and you will be asked to review an applicant who is interested in the position. Then we will share a fit score calculated by **our AI agent.**	Hello! We are developing an online job recommendation platform called JobConnect. JobConnect offers personalized job recommendations for job seekers based on their experience, skills, personality, and even lifestyle. Our platform relies on a **human job coordinator** to evaluate each job applicant's fit for the posted jobs, so job seekers can save their time searching and finding best-fitting jobs. **How does it work?** Job seekers can create their own profiles by simply answering some questions about their interests and skills. Based on the provided information, our agent matches the applicant with job ads and provides a fit score. The fit score is calculated by **one of our trained human coordinators.** JobConnect is free to job seekers! We will show you a brief demonstration of JobConnect and ask you to evaluate the service. On the next page, you will see a job posting and you will be asked to review an applicant who is interested in the position. Then we will share a fit score calculated by **one of our human coordinators.**

Favorable Outcome Condition	Unfavorable Outcome Condition
Here is the fit score calculated by our coordinator. Fit score: **90/100** Our coordinator **recommends** that the user (ID: dmjerde) apply for the position based on the fit score.	Here is the fit score calculated by our agent. Fit score: **40/100** Our agent **does not recommend** that the user (ID: dmjerde) apply for the position based on the fit score.