scientific reports

OPEN

What algorithmic evaluation fails to deliver: respectful treatment and individualized consideration

Jinseok S. Chun[1], David De Cremer[2], Eun-Ji Oh[1] & Youngsang Kim[1]

As firms increasingly depend on artificial intelligence to evaluate people across various contexts (e.g., job interviews, performance reviews), research has explored the specific impact of algorithmic evaluations in the workplace. In particular, the extant body of work focuses on the possibility that employees may perceive biases from algorithmic evaluations. We show that although perceptions of biases are indeed a notable outcome of AI-driven assessments (vs. those performed by humans), a crucial risk inherent in algorithmic evaluations is that individuals perceive them as lacking respect and dignity. Specifically, we find that the effect of algorithmic (vs. human) evaluations on perceptions of disrespectful treatment (a) remains significant while controlling for perceived biases (but not vice versa), (b) is significant even when the effect on perceived biases is not, and (c) is larger in size than the effect on perceived biases. The effect of algorithmic evaluations on disrespectful treatment is explained by perceptions that individuals' detailed characteristics are not properly considered during the evaluation process conducted by AI.

Keywords Artificial intelligence, AI, Algorithmic evaluations, Respect, Individualized consideration, Biases

As the utilization of artificial intelligence (AI) is extended to the domains that have been traditionally reserved for human efforts, firms are continuously expanding their investments on AI, particularly with the focus of using it to aid their managerial decisions and evaluations[1,2]. One of the reasons for this trend is that firms consume large amounts of digitized data about their employees. If sense can be made of those data by means of algorithm-driven systems, organizations can gain a competitive edge by making more accurate evaluation and thus promoting its overall efficiency[3]. Firms are in fact using algorithms to evaluate which employees to promote[4]and which people to recruit[5]. Given that algorithms can sometimes make superior judgments than those made by human experts[6], it is no surprise that firms are increasingly relying on AI to optimize organizational evaluations[7].

On the other hand, research has also identified risks tied to the incorporation of algorithms into the workplace. One significant concern is that, as AI is used for managerial decision making, the ones affected by those decisions (e.g., employees) may question whether they are being evaluated in a fair, unbiased manner[8,9]. Relevant examples abound, such as Amazon's algorithmic recruitment efforts, where candidates were ranked in a gender-biased manner, and the case in the United Kingdom, where algorithms intended to reduce grading biases instead ended up penalizing students from poorer schools or neighborhoods[10]. Research has found that employees can actually perceive biases involved in decisions and evaluations generated by AI[11,12].

Although this line of research has shown that algorithmic evaluations (i.e., evaluations conducted by AI) are viewed as problematic because they seem biased, we focus on a related but distinct type of risk rooted in AI-driven evaluations: people perceive algorithmic evaluations as lacking respect[13,14]. We formulate and test the argument that individuals evaluated by AI, as opposed to human evaluators, perceive the treatment as disrespectful. More importantly, we present evidence suggesting that perceptions of disrespectful treatment may constitute a more essential risk embedded in algorithmic evaluations than perceived biases.

Perceptions of respectful treatment hold significance because not only they are differentiated from perceptions of unbiasedness but they present distinct challenges to important workplace outcomes[13,14]. In fact, research has found that perceptions of disrespect can determine employees' social behaviors toward coworkers over and above the experience of biases[15]. Moreover, those behaviors shaped by the perceived lack of respectful treatment are also the ones that determine the overall tone of interpersonal dynamics in the workplace, which eventually contribute to the organizational bottom lines[16].

Existing evidence indicates that perceptions of respectful treatment may indeed suffer under algorithm-driven evaluations and decisions. For example, Binn and colleagues[17]provided qualitative evidence on perceptions of

[1]Sungkyunkwan University, Seoul, South Korea. [2]Northeastern University, Boston, USA. email: jschun2000@gmail.com; ykim03@skku.edu

indignity as a potential consequence of algorithmic decisions. More recently, Bankins and colleagues[18]offered quantitative data showing how algorithmic (vs. human) decisions resulted in lower perceptions of respectful treatment. These findings can be explained by at least two interrelated patterns. First, people tend to draw a line between humans and non-human entities, including computers and algorithms[19,20]. Believing that human intricacies take humans to correctly appraise, people are likely to judge that AI would fail in understanding their true characteristics[21,22]. Second, people may worry that the details of their individual profiles are suppressed during algorithmic evaluations, which are often believed to focus on developing models that explain overall patterns rather than discerning the specific features of each case[23]. As a result, they may fear that important aspects of their profiles are overlooked in these evaluations. Based on these reactions, people may conclude that their individual characteristics are not properly considered in algorithmic evaluations. Because individualized consideration is a key indicator of respect in organizations[24–26], algorithmic evaluations may result in lower levels of respectful treatment compared to human evaluations.

Juxtaposing the perspectives on how algorithmic evaluations may lead to perceived disrespect and bias, we posit that perceptions of disrespect represent a more fundamental reaction to algorithmic (vs. human) evaluations than perceptions of bias. Specifically, when people receive evaluations conducted by AI and experience a lack of individualized consideration, they may see this as a stronger indication of disrespect and indignity than of bias and discrimination. Along with realizing that their individual characteristics might not be adequately considered, people undergoing algorithmic evaluations may also recognize that the same processes are applied to others. Believing that algorithms are used to detect patterns across large datasets and apply these patterns universally, they may suspect that the same algorithms have been used to evaluate everyone, including themselves. Given this belief, individuals may not necessarily develop strong perceptions of bias and discrimination, as they do not feel singled out during the evaluation process. Although perceptions of disrespectful treatment driven by algorithmic evaluations may have a downstream consequence of inducing perceptions of bias, this effect may diminish once disrespect is accounted for. In contrast, perceptions of disrespect may persist even when biases are addressed. People can perceive disrespect even when they encounter evaluation processes that seem unbiased. Even if algorithms advance to better address potential biases compared to human evaluators[27,28], they may still be seen as non-human evaluators assessing numerous targets simultaneously. As a result, individuals evaluated by AI are likely to feel a lack of individualized consideration, leading to perceptions of disrespectful treatment. We empirically test these predictions in the context of work organizations.

Plan of study

We present four studies that examine the effect of algorithmic evaluations (vs. human evaluations) on people's perceptions of respectful treatment and unbiasedness. In Studies 1 and 2, we survey people's actual experiences of being evaluated by AI versus human evaluators during job interviews. In subsequent studies, we ask participants to imagine being evaluated by an AI evaluator versus a human evaluator. In Study 3, we describe performance reviews conducted by AI or a human manager, and participants imagine that the recipient is either themselves or a third-party employee. In Study 4, we incorporate perceptions of individualized consideration as the underlying process that explains the effect of algorithmic evaluations on perceptions of respectful treatment.

As described below, although we draw on Colquitt's[14] scales that have been found to capture participants' perceptions of respectful treatment and unbiasedness, our supplemental analyses using the most representative item for respectful treatment and unbiasedness (e.g., "The evaluation treats me with respect" and "The evaluation is free of bias," respectively) produced results that are identical in terms of statistical significance to the results reported below. We use a 1 "not at all" to 5 "very much" scale for all our dependent variables in all our studies. Raw data, computer codes, and study materials for the present research can be found on the following webpage: https://osf.io/568cn.

Methods

Study 1

In Study 1, we examined the effect of algorithmic evaluations on perceptions of respectful treatment and unbiasedness in the context of job interviews. We surveyed people's experience of being evaluated by AI versus human interviewers. Considering that job interviews with human interviewers often involve in-person meetings where interpersonal exchanges (e.g., handshakes) could significantly influence job candidates' experiences[29], we focused on human interviews conducted via Zoom (the online communication application). To collect data on participants' recent experience, we only recruited those with relevant job interviews during the past 12 months.

In addition, we considered that algorithmic versus human evaluations may be different from each other in other features including the duration of the interview, whether the interview was for a permanent or a temporary job, the outcome of the interview, the stage of the interviews in which participants experienced algorithmic (vs. human) evaluations, and the projected salaries of the job. We thus measured these differences and used them as control variables.

Sample. We recruited American participants with an experience of having a job interview where they were evaluated by (a) either an AI interviewer (i.e., algorithmic evaluations) (b) or a human interviewer (i.e., human evaluations) via Zoom. The recruitment for these two groups of participants occurred on the same date and was done on Prolific, an online platform supporting academic research[3031]. We ensured that no individual participated in both surveys. We received responses from 98 to 99 participants in the algorithmic and human evaluation conditions, respectively ($n_{\text{algorithmic}} = 98$, $n_{\text{human}} = 99$).

Procedures. Participants first provided information about the control variables discussed above (e.g., duration of the interactions). Participants then reported perceptions of respectful treatment and unbiasedness. We drew on Colquitt's[14] three items to measure respectful treatment (e.g., "The interview treated me with respect") and three items to measure unbiasedness (e.g., "The interview was free of bias).

Variables	M	SD	α	1	2	3	4	5	6	7	
1. Algorithmic (1) vs. human (0)	0.50	0.50	–	–							
2. Permanent (1) vs. temporary (0)	0.54	0.50	–	−0.14	–						
3. Job offer (1) vs. no job offer (0)	0.50	0.50	–	−0.29	−0.21	–					
4. Duration of interactions	2.18	0.80	–	−0.27	0.28	0.11	–				
5. Stage: Middle (1) vs. initial (0)	0.35	0.48	–	−0.08	0.05	−0.09	0.20	–			
6. Stage: Last (1) vs. initial (0)	0.12	0.32	–	−0.08	−0.01	0.20	0.21	−0.26	–		
7. Salary	2.93	2.13	–	0.16	0.40	−0.26	0.29	0.18	−0.05	–	
8. Respectful treatment	4.20	0.92	0.92	−0.44	0.03	0.38	0.12	−0.07	0.08	−0.18	–
9. Unbiasedness	4.14	0.81	0.85	−0.31	0.02	0.29	−0.02	−0.06	0.07	−0.15	0.74

Table 1. Descriptive statistics, reliability estimates, and intercorrelations (study 1). $n = 197$. $|r|s > 0.14$ are significant at 0.05 level.

	Respectful treatment		Unbiasedness	
Variables	Model 1	Model 2	Model 3	Model 4
Algorithmic evaluations	−0.80 (0.12)*	−0.62 (0.13)*	−0.51 (0.11)*	−0.43 (0.12)*
Permanent		0.13 (0.13)		0.13 (0.12)
Job offer		0.50 (0.13)*		0.37 (0.12)*
Duration of interactions		0.03 (0.08)		−0.12 (0.08)
Stage: Middle (vs. initial)		−0.14 (0.13)		−0.05 (0.12)
Stage: Final (vs. initial)		−0.09 (0.19)		0.05 (0.18)
Salary		−0.03 (0.03)		−0.01 (0.03)
R^2	0.19*	0.27*	0.10*	0.16*

Table 2. OLS regression analyses on perceptions of respectful treatment and unbiasedness (study 1). $n = 199$. Values in parentheses are standard errors. $*p < 0.05$.

Results. The descriptive statistics and intercorrelations can be found in Table 1. Participants reported that algorithmic and human evaluations were different from each other in terms of the duration of the interactions: participants reported that interactions with AI interviewers were shorter than those with human interviewers, $p < 0.001$. In addition, more participants reported that they received a job offer after the interview involving algorithmic evaluations than human evaluations, $p < 0.001$. There was no significant difference between algorithmic and human evaluations in the stage of job interviews they were used, $p = 0.184$. Finally, algorithmic evaluations were used for higher-paying (not lower-paying) jobs than human evaluations, $p = 0.027$. As described below, we accounted for all these features in our analyses.

As can be found in Table 2, the OLS regression showed that participants perceived lower levels of respectful treatment from algorithmic evaluations ($M = 3.91$, $SD = 0.89$) than human evaluations ($M = 4.61$, $SD = 0.57$), Cohen's $d = -0.85$, t = -6.56, $p < 0.001$ (M1). They also perceived lower levels of unbiasedness from algorithmic evaluations ($M = 3.88$, $SD = 0.88$) than human evaluations ($M = 4.39$, $SD = 0.64$), Cohen's $d = -0.63$, t = -4.61, $p < 0.001$ (M3). In addition, the effects of algorithmic versus human evaluations on respectful treatment and unbiasedness remained statistically significant when control variables were included in the models (M2 and M4).

More importantly, we examined whether the negative effect of algorithmic (vs. human) evaluations on respectful treatment remained significant controlling for unbiasedness and vice versa. The effect on respectful treatment remained significant controlling for unbiasedness, t = -4.44, $p < 0.001$. In contrast, the effect on unbiasedness became nonsignificant controlling for respectful treatment, $t = -0.23$, $p = 0.819$.

These findings suggest that job candidates perceived lower levels of respectful treatment and unbiasedness from algorithmic evaluations than human evaluations, and these differences existed over and above other important features including whether they received a job offer after the interview. Moreover, we found that the negative effect of algorithmic (vs. human) evaluations on perceptions of respectful treatment remained significant when perceptions of unbiasedness were accounted for. In contrast, the effect on perceived unbiasedness became nonsignificant when taking into account perceptions of respectful treatment.

Study 2

In Study 1, we compared the effect of algorithmic versus human evaluations while controlling for other important features of the interviews. However, we could not eliminate the possibilities that there may have been some other differences between the two types of evaluations that could have influenced job candidates' experiences. For example, although we controlled whether participants received a job offer after the interview, participants might have accepted the offer less frequently after an interview with algorithmic (vs. human) evaluations, which

Variables	M	SD	α	1	2	3	4	5	6
1. Algorithmic (1) vs. human (0)	0.48	0.50	–	–					
2. Permanent (1) vs. temporary (0)	0.55	0.50	–	0.02	–				
3. Job offer (1) vs. no job offer (0)	0.43	0.50	–	− 0.04	0.06	–			
4. Working (1) vs. not working (0)	0.22	0.42	–	0.06	0.18	0.59	–		
5. Duration of interactions	2.33	0.90	–	− 0.14	0.35	0.15	0.11	–	
6. Respectful treatment	3.96	0.92	0.88	− 0.31	− 0.02	0.31	0.22	0.06	–
7. Unbiasedness	3.83	0.87	0.84	− 0.16	0.06	0.35	0.30	0.03	0.57

Table 3. Descriptive statistics, reliability estimates, and intercorrelations (study 2). $n = 195$. $|r|$s > 0.14 are significant at 0.05 level.

	Respectful treatment		Unbiasedness	
Variables	Model 1	Model 2	Model 3	Model 4
Algorithmic evaluations	− 0.57 (0.13)*	− 0.57 (0.12)*	− 0.28 (0.12)*	− 0.30 (0.12)*
Permanent		− 0.08 (0.13)		0.08 (0.13)
Job offer		0.46 (0.15)*		0.45 (0.15)*
Working		0.22 (0.18)		0.33 (0.18)
Duration of interactions		− 0.02 (0.07)		− 0.07 (0.07)
R^2	0.10*	0.20*	0.03*	0.17*

Table 4. OLS regression analyses on perceptions of respectful treatment and unbiasedness (study 2). $n = 195$. Values in parentheses are standard errors. *$p < 0.05$.

could have had further downstream consequences on participants' beliefs regarding the evaluation processes. To address this concern, in Study 2 we recruited people who experienced *both* algorithmic and human evaluations *in a given job interview*. We then randomly assigned participants to the algorithmic evaluation or the human evaluation condition by asking participants to recall the corresponding elements that they experienced during the interview.

Sample. We received responses from 197 American participants who had a job interview involving both AI and human evaluators on Prolific. Two of them later described that they did not have an AI evaluator, thus being excluded from the sample (which did not affect the statistical significance of the findings).

Procedures. Participants first provided information regarding the control variables used in Study 1 (except for the stage of the job interview, which did not significantly differ in Study 1). In addition, participants also reported how long ago the interview had taken place and whether they were still working for the company.

We then manipulated the salience of algorithmic versus human evaluations. In the [algorithmic evaluation vs. human evaluation] condition, we asked participants to describe the evaluation processes with the [AI vs. human] evaluators in terms of the questions the evaluators asked, the answers they offered, and the feelings they experienced during the evaluation processes. Participants then reported perceptions of respectful treatment and unbiasedness, using the same items as Study 1.

Results. The descriptive statistics and intercorrelations can be found in Table 3. Participants in the algorithmic vs. human evaluation conditions provided similar responses regarding all control variables except for the duration of interactions that demonstrated a marginal difference between the two conditions, $p = 0.059$ (algorithmic evaluations being shorter than human evaluations.

As can be found in Table 4, participants perceived lower levels of respectful treatment from algorithmic evaluations ($M = 3.80$, $SD = 0.86$) than human evaluations ($M = 4.30$, $SD = 0.69$), Cohen's $d = -0.61$, $t = -4.45$, $p < 0.001$ (M1). A similar difference was found for perceived unbiasedness of algorithmic evaluations ($M = 3.68$, $SD = 0.95$) and human evaluations ($M = 3.96$, $SD = 0.76$), Cohen's $d = -0.32$, $t = -2.28$, $p = 0.024$ (M3). We found that the effect of algorithmic versus human evaluations existed over and above control variables (M2 and M4).

Moreover, we found the negative effect of algorithmic (vs. human) evaluations on perceptions of respectful treatment controlling for perceptions of unbiasedness ($t = -3.79$, $p < 0.001$). In contrast, when respectful treatment was controlled for, the effect of algorithmic versus human evaluations on unbiasedness became nonsignificant ($t = -0.17$, $p = 0.864$). These results replicated and extended the findings from Study 1 by demonstrating the effect of algorithmic evaluations on perceptions of respectful treatment using random assignment.

Study 3

In Study 3, we investigated another important context of workplace evaluations: performance reviews. We tested the effect of a performance report generated by AI versus human evaluators. We also sought to examine whether the reaction to algorithmic versus human evaluations was an egocentric phenomenon that only applies to a situation where the recipient was the self, or it was a general process that emerged even when people observed another person receiving the evaluations[32]. To achieve this goal, we used vignettes and asked participants to

Variables	M	SD	α	1	2	3
1. Algorithmic (1) vs. human (0)	0.50	0.50	–	–		
2. Self (1) vs. third-party (0)	0.51	0.50	--	0.00	–	
3. Respectful treatment	3.06	1.15	0.96	− 0.41	0.14	–
4. Unbiasedness	3.35	0.85	0.68	− 0.15	0.04	0.65

Table 5. Descriptive statistics, reliability estimates, and intercorrelations (study 3). $n = 403$. $|r|$s > 0.10 are significant at 0.05 level.

	Respectful treatment		Unbiasedness	
Recipient	Algorithmic evaluations	Human evaluations	Algorithmic evaluations	Human evaluations
Self	2.73 (1.03)	3.71 (1.03)	3.25 (0.83)	3.53 (0.83)
Third-party	2.44 (1.07)	3.36 (1.01)	3.20 (0.86)	3.43 (0.87)
Overall	2.59 (1.06)	3.54 (1.03)	3.23 (0.84)	3.48 (0.85)

Table 6. Perceptions of respectful treatment and unbiasedness (study 3). $n = 403$. Values in parentheses are standard deviations.

imagine situations where either they or a third-party was on the receiving end of algorithmic versus human evaluations.

Sample. We received responses from 403 American participants on Prolific ($n_{\text{algorithmic|self}} = 102$, $n_{\text{human|self}} = 101$, $n_{\text{algorithmic|third-party}} = 99$, $n_{\text{human|third-party}} = 101$).

Procedures. This study had a 2 (evaluation: algorithmic vs. human) × 2 (recipient: self vs. third party) factorial design, and participants were randomly assigned to one of the four vignettes. In the self condition, the vignette asked participants to imagine that they were an employee of a large investment bank in the United States. It described that they just received a performance report for the previous quarter as part of the formal evaluation procedures. In the algorithmic versus human evaluation conditions, participants imagined that the performance report was created by an algorithm versus a human manager supervising their work. The description in the third-party condition was identical to the self condition, except that the participants in the third-party condition were asked to imagine that an employee named Jason was working at the bank and received the performance report (created by an algorithm vs. a human manager).

Based on these descriptions, participants reported their perceptions of respectful treatment and unbiasedness. Similar to Studies 1 and 2, we adapted Colquitt's three items[14] to measure perceptions of respectful treatment and unbiasedness.

Results. The descriptive statistics and intercorrelations can be found in Table 5, and the mean perceptions of respectful treatment and unbiasedness from each experimental condition can be found in Table 6. We first examined perceptions of respectful treatment and unbiasedness using 2 × 2 ANOVAs. Participants perceived lower levels of respectful treatment from algorithmic evaluations ($M = 2.59$, $SD = 1.06$) than human evaluations ($M = 3.53$, $SD = 1.03$), Cohen's $d = -0.82$, $F(1, 399) = 83.82$, $p < 0.001$. Algorithmic evaluations also led to lower perceptions of unbiasedness ($M = 3.23$, $SD = 0.84$) than human evaluations ($M = 3.48$, $SD = 0.85$), Cohen's $d = -0.29$, $F(1, 399) = 9.12$, $p = 0.003$.

While there was also a significant main effect of the recipient on perceptions of respectful treatment (perceiving less respectful treatment when the recipient was a third party, $p = 0.003$; the recipient did not have a significant effect on perceptions of unbiasedness, $p = 0.493$), there were no significant interactions between evaluations and the recipient predicting perceptions of respectful treatment and unbiasedness, $ps > 0.849$. This pattern indicated that the effects of algorithmic versus human evaluations generalized regardless of whether the recipient was the self or a third party.

The negative effect of algorithmic (vs. human) evaluations on perceptions of respectful treatment remained significant controlling for unbiasedness, $t = -9.25$, $p < 0.001$. In contrast, controlling for respectful treatment, the effect of algorithmic (vs. human) evaluations on unbiasedness became *positive* and significant, $t = 3.45$, $p < 0.001$.

In this study, we found that algorithmic evaluations evoked lower perceptions of respectful treatment than human evaluations in performance reviews of work organizations. This effect was observed regardless of whether the recipient was the self or a third party, indicating that the effect of algorithmic evaluations did not emerge as an egocentric phenomenon[32]. In addition, in contrast to the negative effect of algorithmic (vs. human) evaluations on respectful treatment that remained significant when controlling for unbiasedness, the negative effect of algorithmic (vs. human) evaluations on unbiasedness disappeared when controlling for respectful treatment—it even went the opposite direction. These patterns hint at the possibility that when respectful treatment is accounted for, algorithmic evaluations' effect on perceived unbiasedness may become less robust.

Study 4

In this study, we incorporated perceptions of receiving individualized consideration and examined its effect as the underlying process explaining the effect of algorithmic evaluations on perceptions of respectful treatment. In addition, we attempted to address the possibility that participants in the human evaluation condition might infer that they were evaluated by a manager with whom they shared a trusting and supportive relationship. To do so, we created vignettes in which participants in the human evaluation condition imagined that their evaluation came from not their direct supervisor but instead a senior manager in charge of performance management.

Sample. We received responses from 198 American participants on Connect, an online platform that supports data collection for academic research ($n_{\text{algorithmic}} = 102$, $n_{\text{human}} = 96$).

Procedures. Participants were given vignettes describing performance reviews. All vignettes asked participants to imagine working at a medium-sized insurance company in the United States. The vignettes described that every employee's performance profile was evaluated within the same company-wide pool. Participants were randomly assigned to either the algorithmic or human evaluation condition. Those in the algorithmic versus human evaluation conditions were informed that the evaluation was done by AI versus a senior VP of Performance Management.

Based on the descriptions, participants reported their perceptions of respectful treatment and unbiasedness using the items adapted from Colquitt[14]. Participants also answered three items that captured their perceptions of receiving individualized consideration during the evaluation process (e.g., "The evaluation appropriately considers my individual characteristics")[26].

Results. The descriptive statistics and intercorrelations can be found in Table 7. Participants perceived lower levels of respectful treatment from algorithmic evaluations ($M = 3.04$, $SD = 1.31$) than human evaluations ($M = 3.49$, $SD = 1.05$), Cohen's $d = -0.37$, t = -2.68, $p = 0.008$. In contrast, there was no significant difference in perceptions of unbiasedness between algorithmic evaluations ($M = 3.27$, $SD = 1.30$) and human evaluations ($M = 3.51$, $SD = 0.96$), Cohen's $d = -0.21$, t = -1.47, $p = 0.142$.

The negative effect of algorithmic (vs. human) evaluations on perceptions of respectful treatment remained significant when controlling for perceived unbiasedness, t = -2.49, $p = 0.014$. The effect on perceived unbiasedness was nonsignificant when controlling for perceptions of respectful treatment, $t = 1.15$, $p = 0.253$.

Participants also reported lower perceptions of receiving individualized consideration from algorithmic evaluations ($M = 2.74$, $SD = 1.21$) than human evaluations ($M = 3.23$, $SD = 1.20$), Cohen's $d = -0.40$, t = -2.87, $p = 0.005$. To test whether perceptions of receiving individualized consideration explained the effect of algorithmic (vs. human) evaluations on perceptions of respectful treatment, we estimated its indirect effect using a quasi-Bayesian approximation with 5,000 Monte-Carlo draws[33]. The indirect effect was statistically significant, point estimate $= -0.32$, 95% CI = [-0.561, -0.087].

Algorithmic evaluations led to lower perceptions of respectful treatment than human evaluations, even when the human evaluations were conducted by a senior manager in charge of performance management instead of employees' direct supervisor. In contrast, the difference between algorithmic and human evaluations in perceived biases was nonsignificant. Finally, perceptions of receiving individualized consideration during the evaluation process provided insights regarding the underlying process that explained the effect of algorithmic evaluations on perceptions of respectful treatment.

Discussion

In the present research, we reveal the centrality of perceived disrespect in understanding people's reactions to algorithmic evaluations in the workplace. Our empirical evidence showed that perceptions of receiving respectful treatment were a key element that was damaged by receiving algorithmic evaluations in work organizations[13–15]. While perceptions of unbiased evaluations were also harmed when people were evaluated by AI versus human evaluators, reactions to algorithmic evaluations were demonstrated in a more significant and consistent manner in the domain of respect and dignity. Specifically, the detrimental effect of algorithmic (vs. human) evaluations on perceptions of respectful treatment was significant (a) when controlling for perceived unbiasedness (but not vice versa) and (b) even when the effect on perceived unbiasedness was nonsignificant.

As an additional step to assess the overall effects of algorithmic evaluations on perceptions of respectful treatment and unbiasedness, we ran a meta-analysis on our findings from Studies 1 to 4 (using a fixed-effect model given the internal consistency of our study designs and samples). The results demonstrated that the effect of algorithmic evaluations on respectful treatment (-0.75) was larger than unbiasedness (-0.35), and their 95% confidence intervals did not overlap with each other ([-0.88, -0.62] vs. [-0.48, -0.23]). Again, this evidence suggests that the perceptions of respect were more prone to be harmed by algorithmic evaluations than perceptions of unbiased evaluations. The results from the meta-analysis are depicted in Fig. 1.

Before discussing the implications of our findings, it should be noted that we do *not* argue that the effect of algorithmic evaluations on perceptions of biases and discrimination is unimportant. In Studies 1 to 3, we did find that algorithmic evaluations produced lower perceptions of unbiasedness than human evaluations. Given the risks driven by perceived biases in organizations, the effect of algorithmic evaluations on the inferences of biases should be properly recognized[27,34]. That said, our findings consistently showed that perceptions of

Variables	M	SD	α	1	2	3
1. Algorithmic (1) vs. human (0)	0.52	0.50	–	–		
2. Respectful treatment	3.26	1.21	0.96	−0.19	–	
3. Unbiasedness	3.38	1.15	0.92	−0.10	0.81	–
4. Individualized consideration	2.98	1.23	0.92	−0.20	0.72	0.73

Table 7. Descriptive statistics, reliability estimates, and intercorrelations (study 4). $n = 198$. $|r|$s > 0.14 are significant at 0.05 level.

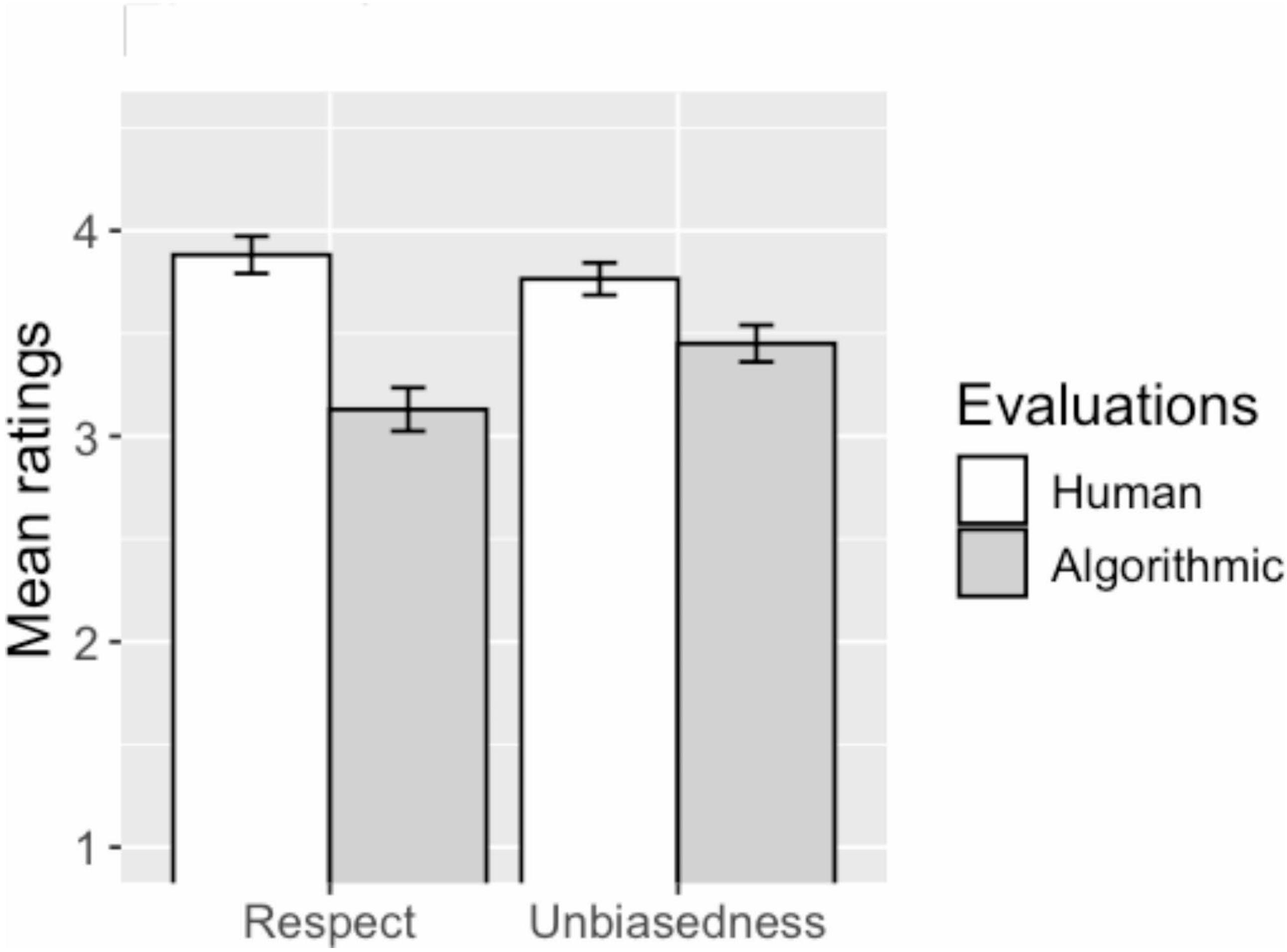

Fig. 1. Perceptions of respectful treatment and unbiasedness as afunction of algorithmic vs. human evaluations

respectful treatment were impaired more significantly by algorithmic evaluations, revealing the essential role that respect and dignity play in understanding the way AI-driven evaluations influence people's experiences in the workplace.

Implications

Our findings in Study 4 suggest that, even in situations where firms can minimize people's worries about potential biases included in algorithmic evaluations, the issue of disrespect may persist. As discussed earlier, firms have strived to improve their algorithms in terms of reducing biases[11,12], and such efforts have led to meaningful achievements[28]. We observe that such efforts may not necessarily eliminate people's concerns regarding evaluations conducted by AI.

Algorithmic evaluations' impact on the experience of respectful treatment can be further discussed in relation to quantification[17]. Researchers and practitioners have made remarkable progress in enhancing the accuracy of their prediction models that rely on quantified data[35,36]. However, during this very process of quantification, those who are assessed by algorithms may believe that their individual characteristics are obscured and/or disregarded[12]. While being reduced to data points, people who receive the evaluations may believe that the intricacies of their characteristics have been somehow stripped off, concluding that their organizations have not attempted to respect their individual identities. If the development of AI focuses solely on improving the accuracy of predictions and minimizing biases, it may continue to overlook the issues of dignity and identity rooted in human experience.

In terms of practical recommendations, our findings indicate that firms may need to develop plans on how to secure perceptions of respectful treatment from AI-driven evaluations. In our Study 2, we found that people perceived more respectful treatment when recalling the details of a given job interview in terms of human evaluations versus algorithmic evaluations. As discussed earlier, when people notice that they are evaluated by algorithms in the workplace, they may feel being treated as just another bit of quantified data. Organizations can reduce such unwanted consequences by emphasizing the presence of a human element. For example, they can provide opportunities for recipients to interact with human evaluators in the initial and/or final stages of the evaluation[37,38].

Our findings further emphasize the importance of attending to the *recipients*' reactions to algorithmic evaluations[39,40]. Although AI is seeing rapid progress in its ability to reduce biases and increase accuracy, that may only represent what *organizations* are doing to evaluate their targets in a fair and effective way[35,36]. It should be valuable for them to actively communicate how those implementations are used to better understand the recipients so that those who receive the evaluations can recognize the firms' efforts to respect their individual profiles[26,41,42]. For example, firms can consider offering evidence of individualization and discussing it to secure the recipients' beliefs that their individual identities are recognized. By doing so, they may be able to protect the sense of respect and dignity, which is something that is so core to what people need as humans.

Data availability

Raw data, computer codes, and study materials for the present research can be found on the following webpage: https://osf.io/568cn.

Received: 30 November 2023; Accepted: 14 October 2024
Published online: 29 October 2024

References

1. Bean, R. Why is it so hard to become a data-driven company? *Harvard Business Rev.* (2021).
2. De Cremer, D., Narayanan, D., Nagpal, M., McGuire, J. & Schweitzer, S. AI Fairness in Action: A human-computer perspective on AI fairness in organizations and society. *Int. J. Hum. Comput. Interact.* **40**, 1–3 (2024).
3. Taskesen, B., Blanchet, J., Kuhn, D. & Nguyen, V. A. A statistical test for probabilistic fairness. in *Proceedings of the ACM Conference on Fairness, Accountability, and Transparency* 648–665 (Association for Computing Machinery, New York, NY, USA, 2021). doi: (2021). https://doi.org/10.1145/3442188.3445927
4. Gee, K. In Unilever's radical hiring experiment, resumes are out, algorithms are in. *Wall Str. J.* (2017).
5. Hoffman, M., Kahn, L. B. & Li, D. Discretion in hiring. *Quart. J. Econom.* **133**, 765–800 (2018).
6. Colson, E. What AI-driven decision making looks like. *Harvard Business Rev.* (2019).
7. Reeves, M. Algorithms can make your organization self-tuning. *Harvard Business Rev.* (2015).
8. De Cremer, D. *Leadership by Algorithm: Who Leads and Who Follows in the Ai Era* (Harriman House Publishing, 2020).
9. Lavanchy, M., Reichert, P., Narayanan, J. & Savani, K. Applicants' fairness perceptions of algorithm-driven hiring procedures. *J. Bus. Ethics.* **188**, 125–150 (2023).
10. Asher-Schapiro, A. Exam grading algorithms amid coronavirus: What's the row about? *Reuters* (2020).
11. Lee, M. K., Kim, J. T. & Lizarondo, L. A human-centered approach to algorithmic services: Considerations for fair and motivating smart community service management that allocates donations to non-profit organizations. in *Proceedings of the CHI Conference on Human Factors in Computing Systems* 3365–3376 (Association for Computing Machinery, New York, NY, USA, 2017). doi: (2017). https://doi.org/10.1145/3025453.3025884
12. Newman, D. T., Fast, N. J. & Harmon, D. J. When eliminating bias isn't fair: Algorithmic reductionism and procedural justice in human resource decisions. *Organ. Behav. Hum. Decis. Process.* **160**, 149–167 (2020).
13. Bies, R. J. & Shapiro, D. L. Interactional fairness judgments: The influence of causal accounts. *Soc. Justice Res.* **1**, 199–218 (1987).
14. Colquitt, J. A. On the dimensionality of organizational justice: A construct validation of a measure. *J. Appl. Psych.* **86**, 386–400 (2001).
15. Aquino, K., Lewis, M. U. & Bradfield, M. Justice constructs, negative affectivity, and employee deviance: A proposed model and empirical test. *J. Organ. Behav.* **20**, 1073–1091 (1999).
16. Ambrose, M. L., Schminke, M. & Mayer, D. M. Trickle-down effects of supervisor perceptions of interactional justice: A moderated mediation approach. *J. Appl. Psych.* **98**, 678–689 (2013).
17. Binns, R. et al. 'It's reducing a human being to a percentage': Perceptions of justice in algorithmic decisions. in *Proceedings of the CHI Conference on Human Factors in Computing Systems* 1–14 (Association for Computing Machinery, New York, NY, USA, 2018). doi: (2018). https://doi.org/10.1145/3173574.3173951
18. Bankins, S., Formosa, P., Griep, Y. & Richards, D. A. I. Decision making with dignity? Contrasting workers' justice perceptions of human and AI decision making in a human Resource Management Context. *Inf. Syst. Front.* **24**, 857–875 (2022).
19. VanRullen, R. & Thorpe, S. J. Is it a bird? Is it a plane? Ultra-rapid visual categorisation of natural and artifactual objects. *Perception.* **30**, 655–668 (2001).
20. Waytz, A., Gray, K., Epley, N. & Wegner, D. M. Causes and consequences of mind perception. *Trends Cogn. Sci.* **14**, 383–388 (2010).
21. Kim, S., Razi, A., Stringhini, G., Wisniewski, P. J. & De Choudhury, M. A. Human-centered systematic literature review of cyberbullying detection algorithms. in *Proc. ACM Hum.-Comput. Interact.* 5, 325:1-325:34 (2021).
22. Haslam, N. Dehumanization: An integrative review. *Pers. Soc. Psychol. Rev.* **10**, 252–264 (2006).
23. Longoni, C., Bonezzi, A. & Morewedge, C. K. Resistance to medical artificial intelligence. *J. Consum. Res.* **46**, 629–650 (2019).
24. Gorman, B. & Seguin, C. Reporting the international system: Attention to foreign leaders in the US news media, 1950–2008. *Soc. Forces.* **94**, 775–799 (2015).
25. Sluss, D. M. & Thompson, B. S. Socializing the newcomer: The mediating role of leader–member exchange. *Organ. Behav. Hum. Decis. Process.* **119**, 114–125 (2012).
26. Chun, J. S., Brockner, J. & De Cremer, D. How temporal and social comparisons in performance evaluation affect fairness perceptions. *Organ. Behav. Hum. Decis. Processes.* **145**, 1–15 (2018).
27. Dietvorst, B. J., Simmons, J. P. & Massey, C. Algorithm aversion: People erroneously avoid algorithms after seeing them err. *J. Exp. Psychol. Gen.* **144**, 114–126 (2015).
28. Boegershausen, J., Castelo, N., Hildebrand, C. & Henkel, A. P. Creating customer service bots that people don't hate. *Harvard Business Rev.* (2023).
29. Schroeder, J., Risen, J. L., Gino, F. & Norton, M. I. Handshaking promotes deal-making by signaling cooperative intent. *J. Personal. Soc. Psychol.* **116**, 743–768 (2019).
30. Palan, S. & Schitter, C. Prolific.ac – a subject pool for online experiments. *J. Behav. Exp. Finance.* **17**, 22–27 (2018).
31. Peer, E., Brandimarte, L., Samat, S. & Acquisti, A. Beyond the Turk: Alternative platforms for crowdsourcing behavioral research. *J. Exp. Soc. Psych.* **70**, 153–163 (2017).
32. Epley, N., Keysar, B., Van Boven, L. & Gilovich, T. Perspective taking as egocentric anchoring and adjustment. *J. Personal. Soc. Psychol.* **87**, 327–339 (2004).
33. Tingley, D. et al. (2019). (Version 4.4.7).
34. Brockner, J., Chen, Y. R., Mannix, E., Leung, K. & Skarlicki, D. Culture and procedural fairness: When the effects of what you do depend on how you do it. *Admin Sci. Quart.* **45**, 138–159 (2000).
35. Albarghouthi, A., D'Antoni, L., Drews, S. & Nori, A. V. FairSquare: Probabilistic verification of program fairness. *Proc. ACM Program. Lang.* 1, 80:1–80:30 (2017).

36. Taskesen, B., Blanchet, J., Kuhn, D. & Nguyen, V. A. A statistical test for probabilistic fairness. (2020).
37. Olkkonen, M. E. & Lipponen, J. Relationships between organizational justice, identification with organization and work unit, and group-related outcomes. *Organ. Behav. Hum. Decis. Process.* **100**, 202–215 (2006).
38. Tyler, T. R. & Lind, E. A. A relational model of authority in groups. in *Advances in Experimental Social Psychology* (ed Zanna, M. P.) vol. 25 115–191 (Academic, (1992).
39. Lind, E. A. et al. In the eye of the beholder: tort litigants' evaluations of their experiences in the civil justice system. *Law Soc. Rev.* **24**, 953 (1990).
40. Narayanan, D., Nagpal, M., McGuire, J. & Schweitzer, S. De Cremer, D. Fairness Perceptions of Artificial Intelligence: A review and path Forward. *Int. J. Human–Computer Interact.* **40**, 4–23 (2024).
41. Polzer, J. T., Milton, L. P. & Swarm, W. B. Capitalizing on diversity: Interpersonal congruence in small work groups. *Admin Sci. Quart.* **47**, 296–324 (2002).
42. Swann, W. B., Kwan, V. S. Y., Polzer, J. T. & Milton, L. P. Fostering group identification and creativity in diverse groups: The role of individuation and self-verification. *Pers. Soc. Psych Bull.* **29**, 1396–1406 (2003).

Author contributions
J.S.C. designed the studies, collected the data, and performed the analyses. All authors discussed the results. J.S.C. wrote the main manuscript. All authors revised the manuscript and approved the final version.

Declarations

Competing interests
The authors declare no competing interests.

Additional information
Correspondence and requests for materials should be addressed to J.S.C. or Y.K.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.