

Thumbs Up or Down: Consumer Reactions to Decisions by Algorithms Versus Humans

Gizem Yalcin , Sarah Lim, Stefano Puntoni, and Stijn M.J. van Osselaer

Journal of Marketing Research

1-22

© The Author(s) 2022



Article reuse guidelines:
sagepub.com/journals-permissions
 DOI: 10.1177/00222437211070016
journals.sagepub.com/home/mrj



Abstract

Although companies increasingly are adopting algorithms for consumer-facing tasks (e.g., application evaluations), little research has compared consumers' reactions to favorable decisions (e.g., acceptances) versus unfavorable decisions (e.g., rejections) about themselves that are made by an algorithm versus a human. Ten studies reveal that, in contrast to managers' predictions, consumers react less positively when a favorable decision is made by an algorithmic (vs. a human) decision maker, whereas this difference is mitigated for an unfavorable decision. The effect is driven by distinct attribution processes: it is easier for consumers to internalize a favorable decision outcome that is rendered by a human than by an algorithm, but it is easy to externalize an unfavorable decision outcome regardless of the decision maker type. The authors conclude by advising managers on how to limit the likelihood of less positive reactions toward algorithmic (vs. human) acceptances.

Keywords

algorithms, decision making, decision outcome favorability, attribution theory

Online supplement: <https://doi.org/10.1177/00222437211070016>

A growing number of companies are using algorithms to make business decisions that directly affect potential and existing customers. For example, algorithms are now used to decide which applicants should be admitted to platforms (e.g., Raya) and who should receive loans (e.g., Upstart; for more examples, see Web Appendix A). As the prevalence of algorithms in consumer-facing decisions increases, so does the managerial importance of understanding consumers' reactions to algorithmic versus human decisions. We investigate consumers' reactions toward a company following a decision (favorable or unfavorable) made by an algorithmic versus a human decision maker. Specifically, we focus on contexts where the decision outcome is considered diagnostic of the consumer's qualifications, deservingness, or merit, such as when consumers submit an application to access a valued service or other benefits.

We demonstrate that consumers react less positively when a favorable decision (e.g., the acceptance of an application) is made by an algorithm rather than by a human. This difference, however, is attenuated for an unfavorable decision (e.g., the rejection of an application). We explain this interaction between the decision maker type and decision outcome favorability by drawing on attribution theory (Jones and Davis 1965; Kelley 1967). Consumers are motivated to internalize favorable decisions, but internal attribution is more difficult when the decisions are made by an algorithm (vs. a human),

so consumers react less positively (e.g., form less positive attitudes toward the company). By contrast, consumers are motivated to externalize unfavorable decisions, and this is similarly easy with algorithmic and human decision makers, so consumers' subsequent reaction is relatively indifferent to the decision maker type.

The current research makes three primary contributions (for a comprehensive literature review, see Table 1). First, our research addresses an underexplored question: How do consumers' attitudes (and related constructs) change as a function of a company's use of algorithmic versus human decision makers in consumer-facing tasks? Previous work has focused on consumers' choices, such as for advice, between an algorithmic and a human decision maker (Dietvorst, Simmons, and Massey 2015; Longoni, Bonezzi, and Morewedge 2019).

Gizem Yalcin is a doctoral candidate, Marketing Management, Rotterdam School of Management, Erasmus University Rotterdam, The Netherlands (email: yalcin@rsm.nl). Sarah Lim is Assistant Professor of Business Administration, Gies College of Business, University of Illinois at Urbana-Champaign, USA (email: lim110@illinois.edu). Stefano Puntoni is Professor of Marketing, Rotterdam School of Management, Erasmus University Rotterdam, The Netherlands (email: spuntoni@rsm.nl). Stijn M.J. van Osselaer is S.C. Johnson Professor of Marketing, Samuel Curtis Johnson Graduate School of Management, Cornell University, USA (email: stijn.vanosselaer@cornell.edu).

Table 1. Overview of Research on Consumers' Responses to Decisions by Algorithms.

Authors	Year	Main Comparison	Main Dependent Variable	Decision Context			Main Finding
				Decision Maker Is Already Chosen	Self-Diagnostic Decision	Decision Outcome Is Known	
Bigman and Gray	2018	Computer versus human	Perceived permissibility	✓		✓	It is less permissible for computers (vs. humans) to make moral decisions.
Bonezzi and Ostinelli	2021	AI versus human	Perceived bias	✓		✓	Algorithmic (vs. human) decision makers are less likely to be perceived as biased.
Cadario, Longoni, and Morewedge	2021	Algorithm versus human	Subjective understanding of decision making, preference	✓	✓		People exhibit an illusory understanding of human (vs. algorithmic) decision making, which makes them more reluctant to use algorithms.
Castelo, Bos, and Lehmann	2019	Algorithm versus human	Trust and preference		✓		People rely on algorithms less for subjective (vs. objective) tasks.
Diab et al.	2011	Formula versus interview	Perceived usefulness	✓	✓		Thorough discussions are viewed as more useful than a formula.
Dietvorst and Bharti	2020	Statistical model versus self	Preference				People prefer riskier decision making methods (humans instead of statistical models) in inherently uncertain decision domains.
Dietvorst, Simmons, and Massey	2015	Statistical model versus self	Preference			✓	Seeing an algorithm err decreases people's willingness to rely on it.
Dietvorst, Simmons, and Massey	2016	Statistical model versus self	Preference			✓	People are more willing to use an algorithm when it is modifiable.
Eastwood, Snook, and Luther	2012	Expert using a formula vs. personal experience	Preference and perceived accuracy		✓		Using an experience is preferred and seen as more accurate than using a formula.
Efendic, Van de Calseyde, and Evans	2020	Algorithm versus human	Perceived accuracy and trust	✓		✓	People judge slowly generated predictions from algorithms (vs. humans) as less accurate and are less willing to rely on them.
Jago	2019	Algorithm versus human	Perceived authenticity	✓			People believe that algorithms (vs. humans) are less authentic.
Kim and Duhachek	2020	Artificial agent versus human	Perceived appropriateness and compliance	✓		✓	Persuasive messages by artificial agents (vs. humans) are more appropriate and effective when the messages have low (vs. high) level construal features.
Lee	2018	Algorithm versus human	Trust and perceived fairness	✓			For tasks that require human (vs. mechanical) skills, algorithms are perceived as less fair and trustworthy.
Logg, Minson, and Moore	2019	Algorithm versus human/self	Weight on advice			✓	People adhere more to advice when it comes from algorithms (vs. humans).

(continued)

Table 1. (continued)

Authors	Year	Main Comparison	Main Dependent Variable	Decision Context			Main Finding
				Decision Maker Is Already Chosen	Self-Diagnostic Decision	Decision Outcome Is Known	
Longoni, Bonezzi, and Morewedge et al.	2019	AI versus human	Preference	✓	✓	People prefer to receive medical care from humans (vs AI).	
Longoni and Cian	2022	AI versus human	Preference			People prefer AI (vs. human) recommenders when utilitarian (vs. hedonic) attributes of a product are more important or salient.	
Newman, Fast, and Harmon	2020	Algorithm versus human	Perceived fairness	✓	✓	People perceive algorithms as less fair than humans.	
Onkal et al.	2009	Statistical model versus human	Weight on advice	✓	✓	People place greater weight on human (vs. algorithmic) advice.	
Promberger and Baron	2006	Computer versus human	Acceptance of advice and trust	✓	✓	Acceptance/trust is greater for humans than computers.	
Shaffer et al.	2013	Doctor soliciting advice from computers versus humans	Perceived ability	✓	✓	Soliciting aid from computers but not from humans decreases doctors' perceived ability, professionalism, and thoroughness.	
Srinivasan and Sarial-Abi	2021	Algorithm versus human	Brand evaluation	✓	✓	Consumers respond less negatively to a brand harm crisis when it is caused by an algorithmic (vs. a human-based) error.	
Yeomans et al.	2019	Algorithm versus human	Preference	✓	✓	People prefer to receive joke recommendations by humans (vs. algorithms) for themselves as well as for others.	
Current research	2022	Algorithm versus human	Attitudes toward the company	✓	✓	Consumers who receive a favorable decision form less positive attitudes toward the company if the decision was made by an algorithm (vs. human). This difference is mitigated for an unfavorable decision.	

Notes: "Decision maker is already chosen" = whether the decision maker type (algorithm vs. human) has been already chosen; if not, consumers (participants) are asked to choose between them. "Self-diagnostic decision" = whether the decisions are about the consumers themselves (i.e., reflect on the consumer's self). "Decision outcome is known": whether the decision outcome is known to consumers.

However, companies usually decide whether to rely on algorithms or humans for a given task; consumers are more often in the position of decision recipients. Unlike prior research, the current research focuses on consumers' reactions to algorithmic versus human decisions about themselves. This distinction is important because the two situations may elicit different psychological processes. Decision recipients face the task of interpreting a decision outcome reflective of one's worth in the eye of others. In such a context, one's reaction to the decision outcome often involves self-serving interpretations and motivated reasoning (Taylor and Brown 1988), a topic that has not been examined in prior research on algorithmic decisions. More generally, as consumers' choices often diverge from their reactions to the given options (Botti and Iyengar 2006), we argue that it is unclear whether findings about consumers' choice behavior (e.g., reluctance to rely on algorithmic advice) are generalizable to the reactions of consumers as decision recipients (e.g., negative reactions to algorithmic decisions made about the consumers themselves).

Second, we examine an important factor that influences consumers' reactions to different decision makers: the favorability of decision outcomes, which is known to affect people's attitudes and behaviors (e.g., Barry, Chaplin, and Grafeman 2006; Rhodewalt and Davison 1986). Both types of decision outcomes are common; companies may deliver approvals or acceptances as well as denials or rejections to existing or potential customers—and yet, the consequences of decision outcome favorability are underexplored in the research on algorithmic (vs. human) decision making. We find that most managers believe that consumers react more positively to decisions made by humans (vs. algorithms) regardless of the decision outcome (see the managerial intuitions study and Web Appendix B). We demonstrate, however, that favorable decision outcomes elicit divergent reactions to algorithmic versus human decision makers, whereas such difference is attenuated for unfavorable decision outcomes.

Third, in examining the process underlying the proposed effect, we elucidate how consumers interpret decisions made by algorithms versus by humans. Unlike prior work that focuses on consumers' diverging perceptions of humans and algorithms (e.g., moral authenticity [Jago 2019], trustworthiness [Lee 2018]), the current work examines consumers' differential attribution of a given decision outcome. Specifically, we demonstrate that for a favorable decision, a human (vs. an algorithmic) decision maker facilitates stronger internal attribution of the decision outcome, whereas for an unfavorable decision, consumers readily engage in external attribution regardless of the type of decision maker. The current research thus marries the psychological literature on attribution (McFarland and Ross 1982; Okten and Moskowitz 2018) with the marketing literature on algorithms (Castelo, Bos, and Lehmann 2019; Puntoni et al. 2021), offering a novel contribution to both.

In the following sections, we review the extant work on algorithmic and human decision making. We draw on attribution theory to make theoretical predictions about how consumers respond to favorable and unfavorable decisions made by algorithms versus humans.

Theoretical Background

An algorithm is “a set of steps that a computer can follow to perform a task” (Castelo, Bos, and Lehmann 2019, p. 809). A growing number of companies rely on algorithms; the market for artificial intelligence is expected to be worth over \$300 billion by 2026 (Markets and Markets 2021). The widespread adoption of algorithms has encouraged researchers to investigate how consumers perceive algorithms versus humans. Existing work has demonstrated that consumers perceive algorithmic and human decision makers to have different strengths and weaknesses. For instance, compared with humans, algorithms are perceived as more objective (Lee 2018; Sundar and Nass 2001) but also as less authentic, less intuitive, and less moral (Bigman and Gray 2018; Jago 2019; Yeomans et al. 2019).

Extant work has examined consumers' choices between algorithmic and human decision makers and has documented an aversion to algorithms (for an exception, see Logg, Minson, and Moore [2019]). For example, consumers are often reluctant to use algorithms to predict stock prices (Onkal et al. 2009), solicit medical advice (Cadario, Longoni, and Morewedge 2021; Longoni, Bonezzi, and Morewedge 2019; Promberger and Baron 2006), and predict people's performance (Dietvorst, Simmons, and Massey 2015). In addition, algorithm aversion varies with contextual factors such as the nature of the task (subjective vs. objective; Castelo, Bos, and Lehmann 2019) and the product (hedonic vs. utilitarian; Longoni and Cian 2022).

The current research is the first to examine consumers' attitudes toward a company in the context in which (1) a decision maker (algorithm vs. human) is already chosen, (2) the decision is made by the company about the consumers themselves (i.e., the decision is self-diagnostic), and (3) a decision outcome is known (see Table 1). Our research context is of managerial importance. Companies often deliver both types of decision outcomes—favorable (e.g., approval, acceptance) and unfavorable (e.g., denial, rejection)—to existing or potential customers. Our in-depth interviews with practitioners confirm the prevalence of algorithms in many consumer-facing tasks such as consumer application evaluations (see Web Appendix C, interviews #1, #5 and #11) and insurance premium decisions (#1). Such decisions are often based on personal information provided by the consumers, and decision outcomes are thus reflective of consumers' qualifications.

We posit that the self-diagnostic nature of many consumer-facing decisions motivates consumers to make different attributions for favorable and unfavorable outcomes. The type of decision maker (algorithm vs. human) affects consumers' internal and external attributions, leading to an interaction effect between the decision maker type and decision outcome favorability on consumers' attitudes toward the company.

Attribution of Favorable and Unfavorable Decisions as a Function of the Decision Maker Type

Consumers often make inferences about the causes of events, actions, and behaviors (Heider 1958; Jones and Davis 1965)

and attribute behaviors or outcomes to either internal or external causes (Kelley 1967). Attribution theory proposes that people are motivated to attribute self-relevant outcomes in a self-serving way: to maintain or enhance their self-worth, people are motivated to attribute favorable outcomes to themselves (i.e., “internal attribution”; Baumeister 1999; Zuckerman 1979) and to attribute unfavorable outcomes to external factors (i.e., “external attribution”; Kelley and Michela 1980; Miller and Ross 1975). In marketing research, attribution theory has been used to explain consumers’ perceptions of a company’s performance (Dunn and Dahl 2012; Folkes 1984; Wan and Wyer 2019), other consumers’ behavior (He and Bond 2015; O’Laughlin and Malle 2002), and one’s own behavior (Leung, Paolacci, and Puntoni 2018; Yoon and Simonson 2008). We contribute to this literature by demonstrating that the decision maker type (algorithm vs. human) affects how consumers attribute favorable versus unfavorable decision outcomes.

Consumers who receive a favorable decision are motivated to make an internal attribution (Luginbuhl, Crowe, and Kahan 1975), and we argue that they find it easier to do so when the decision is made by a human (vs. an algorithm). Consumers often define themselves on the basis of personal characteristics (e.g., abilities, attitudes) that make them feel unique (Brewer 1991; Fromkin and Snyder 1980). Human (vs. algorithmic) decision makers are perceived as more adept at considering individuals’ unique characteristics and qualifications (Longoni, Bonezzi, and Morewedge 2019). In contrast, algorithms usually rely on a set of precoded categories of characteristics and qualifications that are shared by many (note that an algorithm probably would not recognize characteristics that are unique to a single person) and reduce individuals into a number (Newman, Fast, and Harmon 2020). Thus, we predict that consumers view a favorable decision made by a human (vs. an algorithm) as more reflective of their individuality (i.e., unique self) and deservingness (e.g., “My application was accepted because of who I am”), so they would more easily make strong internal attributions for a favorable decision made by a human (vs. an algorithm). It is easier to attribute a good outcome to “me” when the decision maker relied on characteristics and achievements that are “uniquely me.” Put differently, it is more difficult to attribute a positive outcome to something about oneself if those qualities or that something is shared with many others.

In contrast, consumers who receive an unfavorable decision are motivated to make an external attribution, and we argue that they would find no difference in difficulty to do so regardless of whether the decision is made by a human or an algorithm. The decision maker is easily blamed for making a bad decision regardless of whether that decision maker is a human or an algorithm, but for different reasons. For instance, an algorithm can be easily blamed for ignoring consumers’ uniqueness (Longoni, Bonezzi, and Morewedge 2019), while a human can easily be blamed for not being objective (Lee 2018).

If the type of decision maker affects consumers’ ability to make attributional inferences for different decision outcomes, this should be expected to have repercussions for consumer attitudes. Causal reasoning—reasoning about what or who is

responsible for a given outcome—is a key factor in attitude formation and change (e.g., Forsyth 1980; Kelley 1973) and the marketing literature contains many demonstrations that attributions are an important determinant of attitudes toward companies (e.g., Dunn and Dahl 2012). In the context of automation, Leung, Paolacci, and Puntoni (2018) show that the extent to which the consumption context enables people to make internal or external attributions explains their product preferences. For example, in their Study 6, the authors demonstrate that framing an automated product in a way that makes it easier for people to internally attribute favorable consumption outcomes leads to more positive attitudes toward the product.

Summary of Key Predictions and Overview of Studies

We present ten studies that examine our theory (for a summary, see Table 2). Whereas most managers predict (in interviews and surveys) that consumers react more positively to human (vs. algorithmic) decision makers regardless of decision outcome favorability, we demonstrate a robust interaction effect on consumers’ attitudes toward the company (Studies 1a–8) and their word-of-mouth (WoM) intentions (specifically, the net promoter score measure, Study 1b). Furthermore, we examine the underlying attribution processes through both mediation (Studies 4 and 6) and moderation (Study 5). We also rule out alternative explanations including attention (Study 2), social presence (Study 7), and perceived fairness (a follow-up study in the “General Discussion” section). Finally, we offer managerial insights into strategies for improving reactions to favorable decisions made by algorithms (Study 8). We report all conditions and all measures. Some studies included an exploratory measure, for which we report analyses in the Web Appendix. For some of our studies, we screened participants before the study by using an attention check, such as an instructional manipulation check (Oppenheimer, Meyvis, and Davidenko 2009), and those who failed the attention check were not allowed to proceed to the actual study. Our reports of these studies include only those who participated in the actual study. Sample sizes were determined prior to data collection. All data and study materials are available on OSF (osf.io/3bnsz).

Managerial Intuitions

To evaluate the managerial importance of our findings, we examined how practitioners would predict customers’ reactions to favorable versus unfavorable decisions made by humans versus algorithms. We started with a series of in-depth interviews with 14 managers, and none correctly predicted our hypothesized interaction effect. Motivated by this preliminary result, we conducted a survey with a larger group of experienced professionals. We report the results of the in-depth interviews in Web Appendix C and the results of the survey next.

Method

We recruited 88 managers ($M_{\text{age}} = 35.05$ years; 24 women; $M_{\text{work experience}} = 11$ years) from an executive master of

business administration program at a major European business school.

We described a business situation involving consumer applications (see Web Appendix B) and asked the managers to predict how the type of decision maker would influence customer satisfaction in response to an acceptance and in response to a rejection (getting [accepted/rejected] by an algorithm would be better than getting [accepted/rejected] by an employee vs. getting [accepted/rejected] by an algorithm would be equally good as getting [accepted/rejected] by an employee vs. getting [accepted/rejected] by an employee would be better than getting [accepted/rejected] by an algorithm).

Results and Discussion

Managers expected that an algorithmic (vs. a human) decision maker would lead to lower satisfaction regardless of decision outcome favorability ($B = -.09$, $z = -.31$, $p = .758$). Specifically, 61% of the managers predicted that participants would be less satisfied with an acceptance from an algorithm (vs. a human; see Figure 1). Similarly, 59% of the managers predicted that consumers would be less satisfied with a rejection from an algorithm (vs. a human).

Only 5% (i.e., four managers) generated our predicted interaction effect: consumers would react more favorably to an acceptance made by a human (vs. an algorithm) and would be similarly satisfied with a rejection made by a human and by an algorithm. Interestingly, managers also predicted that the decision maker type would matter less for acceptance decisions than for rejection decisions (choice share of “algorithm = human”: $M_{\text{favorable}} = 23.9\%$ vs. $M_{\text{unfavorable}} = 10.2\%$; $B = -1.39$, $z = -2.28$, $p = .023$), the opposite of our hypothesized pattern.

How Are Consumers' Attitudes Toward the Company Affected by the Decision Maker Type and Decision Outcome Favorability?

The first set of studies tested the managers' prediction (i.e., consumers respond more positively to a human [vs. an algorithmic] decision maker regardless of the outcome) against our own (an interaction effect). In Studies 1a–b, we examined our hypothesized interaction effect on two dependent variables: consumers' attitudes toward the company and WoM intentions. We predicted that consumers would react less positively when a favorable decision was made by an algorithm (vs. a human); the differential reaction would be mitigated for an unfavorable decision.

Study 1a: Effect of the Decision Maker Type as a Function of Decision Outcome Favorability: Attitudes Toward the Company

Method. In this preregistered study (aspredicted.org/j7da3.pdf), we randomly assigned 993 Amazon Mechanical Turk (MTurk)

workers ($M_{\text{age}} = 40.06$ years; 531 women)¹ to one of four conditions in a 2 (decision maker type: algorithm vs. human) \times 2 (decision outcome favorability: favorable vs. unfavorable) between-participants design.

Participants read that they were applying for membership at Violethall Country Club (see Web Appendix D). Participants learned that their applications were either accepted (favorable decision condition) or rejected (unfavorable decision condition); we told participants that the decision was made by either a country club algorithm (algorithm condition) or a country club coordinator (human condition). We also told all participants that the decision was final and could not be appealed. After learning the outcome, participants indicated their attitudes toward the country club (“What is your general opinion about Violethall Country Club?”) on three bipolar items (1 = “dislike a great deal”/“very negative”/“not favorable at all,” and 11 = “like a great deal”/“very positive”/“very favorable”; $\alpha = .99$; adapted from Park et al. 2010).

Results. A 2 (decision maker type) \times 2 (decision outcome favorability) analysis of variance (ANOVA) revealed a significant main effect of the decision maker type ($M_{\text{algorithm}} = 5.16$, $SD = 3.10$ vs. $M_{\text{human}} = 5.38$, $SD = 3.29$; $F(1, 989) = 4.98$, $p = .026$, $\eta_p^2 = .01$) and of decision outcome favorability ($M_{\text{favorable}} = 7.49$, $SD = 2.61$ vs. $M_{\text{unfavorable}} = 3.07$, $SD = 1.96$; $F(1, 989) = 924.46$, $p < .001$, $\eta_p^2 = .48$). Consistent with our theory and inconsistent with the managers' predictions, we found a significant interaction effect ($F(1, 989) = 8.46$, $p = .004$, $\eta_p^2 = .01$; see Figure 2): attitudes toward the country club were less positive among participants whose applications were accepted by the algorithm than among participants whose applications were accepted by the club coordinator ($M_{\text{algorithm}} = 7.13$, $SD = 2.59$ vs. $M_{\text{human}} = 7.88$, $SD = 2.58$; $F(1, 989) = 13.15$, $p < .001$, $\eta_p^2 = .01$). Meanwhile, the effect of the decision maker type was significantly mitigated when participants' applications were rejected ($M_{\text{algorithm}} = 3.12$, $SD = 2.11$ vs. $M_{\text{human}} = 3.02$, $SD = 1.82$; $F(1, 989) = .23$, $p = .632$).

Study 1b: Effect of Decision Maker Type as a Function of Decision Outcome Favorability: WoM Intentions

With Study 1b, we aimed to replicate Study 1a with two key changes. First, we tested whether our effect generalizes to a nonsocial context: business loan applications. To further remove social cues, we used “approved” and “denied” instead of “accepted” and “rejected.” Second, we measured participants' WoM intentions, another managerially important dependent variable.

¹ One participant did not complete the demographic variables. Although in subsequent studies, we included only participants who completed all measures in our analysis, we included this one person in this study to be consistent with our preregistration. Our results held significant regardless of whether we included this participant in our analysis (see Web Appendix D). As stated in our preregistration form, we targeted 1,000 participants, but the actual sample size differed for reasons beyond our control (e.g., more people claimed their participation on MTurk than the actual number of participants).

Table 2. Summary of Study Results by Condition.

Study 1a: Effect of Decision Maker Type as a Function of Decision Outcome Favorability: Attitudes Toward the Company (Country Club Application; N = 993; MTurk)			
	Favorable Decision Outcome (N = 494)		Unfavorable Decision Outcome (N = 499)
2 (Decision Maker Type) × 2 (Decision Outcome Favorability)	Algorithm (N = 253)	Human (N = 241)	Human (N = 256)
DV: Attitudes toward the company	7.13 (2.59)	7.88 (2.58)	3.12 (2.11) 3.02 (1.82)
Key Finding: Attitudes toward the company were more positive among participants whose applications were accepted by a human (vs. an algorithm). This effect of the decision maker type was attenuated for participants whose applications were rejected.			
Study 1b: Effect of Decision Maker Type as a Function of Decision Outcome Favorability: WoM Intentions (Business Loan Application; N = 500; Prolific)			
	Favorable Decision Outcome (N = 249)		Unfavorable Decision Outcome (N = 251)
2 (Decision Maker Type) × 2 (Decision Outcome Favorability)	Algorithm (N = 126)	Human (N = 123)	Human (N = 127)
DV: Attitudes toward the company	8.06 (2.51)	9.40 (1.51)	3.39 (1.79)
DV: WoM intentions	6.54 (2.27)	7.77 (1.75)	2.10 (1.81) 2.52 (2.18)
Key Finding: We replicated the interaction effect on both DVs.			
Study 2: Replication with a Real Application Process (Research Participant Pool Application; N = 303; Prolific)			
	Favorable Decision Outcome (N = 152)		Unfavorable Decision Outcome (N = 151)
2 (Decision Maker Type) × 2 (Decision Outcome Favorability)	Algorithm (N = 74)	Human (N = 78)	Human (N = 76)
DV: Attitudes toward the company	4.57 (1.23)	5.47 (1.19)	2.88 (1.45) 3.01 (1.54)
Key Finding: We replicated the interaction effect in a real application experience, and we provided evidence against the alternative account based on inattention.			
Study 3a: Effect of (Not) Disclosing the Decision Maker (Country Club Application; N = 403; Prolific)			
	Favorable Decision Outcome		
3-Cell (Decision Maker Type)	Algorithm (N = 132)		Unspecified (N = 136)
DV: Attitudes toward the company	6.04 (2.70)		7.21 (2.83) 6.97 (2.69)
Study 3b: Effect of (Not) Disclosing the Decision Maker (Loan Application; N = 402; Prolific)			
	Favorable Decision Outcome		
3-Cell (Decision Maker Type)	Algorithm (N = 135)		Unspecified (N = 134)
DV: Attitudes toward the company	7.38 (2.39)		8.50 (1.94) 8.59 (1.98)
Key Findings of Studies 3a and 3b: An algorithmic decision maker led to worse attitudes than both a human decision maker and an unspecified decision maker.			

(continued)

Table 2. (continued)

		Favorable Decision Outcome (N = 287)		Unfavorable Decision Outcome (N = 284)	
		Algorithm (N = 132)	Human (N = 155)	Algorithm (N = 154)	Human (N = 130)
2 (Decision Maker Type) × 2 (Decision Outcome Favorability)					
DV: Attitudes toward the company		6.49 (2.88)	7.54 (2.79)	3.74 (2.36)	3.97 (2.23)
Mediator: internal attribution		6.82 (2.54)	8.15 (2.24)	6.30 (2.98)	6.22 (2.81)
<i>Key Finding:</i> Our core interaction effect on attitudes was mediated by the strength of internal attribution of the decision outcome.					
Study 5: Moderated Mediation by Internal Attribution of a Favorable Outcome (Networking Community Application; N = 443; Prolific)					
		Favorable Decision Outcome		Raffle (N = 221)	
		Evaluation (N = 222)			
2 (Decision Maker Type) × 2 (Decision Method)		Algorithm (N = 123)	Human (N = 99)	Algorithm (N = 102)	Human (N = 119)
DV: Attitudes toward the company		6.98 (2.34)	7.70 (2.49)	5.80 (2.74)	5.60 (2.73)
Mediator: internal attribution		6.84 (2.29)	8.25 (1.96)	4.08 (3.04)	4.25 (3.07)
<i>Key Finding:</i> The self-diagnosticsity of the decision method moderates the mediation effect of internal attribution in the setting of a favorable decision outcome.					
Study 6: External Attribution of an Unfavorable Decision Outcome (Country Club Application; N = 626; MTurk)					
		Unfavorable Decision Outcome		Human (N = 310)	
		Algorithm (N = 316)			
DV: Attitudes toward the company		4.84 (2.62)		4.78 (2.61)	
Mediator 1: perceived objectiveness		7.07 (2.28)		5.95 (2.47)	
Mediator 2: uniqueness consideration		4.41 (2.81)		5.35 (2.74)	
<i>Key Finding:</i> Participants engaged in external attribution (blaming the algorithm [human] for its lack of consideration of individual uniqueness [for their lack of objectivity]), resulting in a relative indifference to the decision maker type.					
Study 7: Effect of Human Decision Making Versus Mere Human Observation (Country Club Application; N = 597; MTurk)					
		Favorable Decision Outcome (N = 299)		Unfavorable Decision Outcome (N = 298)	
		Algorithm Only (N = 95)	Human (N = 98)	Algorithm Only (N = 103)	Human (N = 101)
			Algorithm with Human Monitoring (N = 106)		Algorithm with Human Monitoring (N = 94)
DV: Attitudes toward the company		7.09 (2.68)	7.82 (2.59)	4.04 (2.47)	3.76 (2.29)
			6.68 (2.65)		3.46 (1.90)
<i>Key Finding:</i> Participants had less positive attitudes toward the company when their applications were accepted by a human (vs. an algorithm) regardless of whether a human observed the algorithm's decision, contradicting the alternative account based on social presence.					
Study 8: Humanizing Algorithms to Mitigate Negative Consequences (Country Club Application; N = 601; Prolific)					
		Favorable Decision Outcome		Human-Like Algorithm (N = 201)	
		Algorithm (N = 199)		Human (N = 201)	
DV: Attitudes toward the company			7.07 (2.67)	7.87 (2.52)	7.64 (2.82)
<i>Key Finding:</i> Participants had similarly positive attitudes toward the company when their applications were accepted by a human and by a human-like algorithm, suggesting that anthropomorphization can mitigate the negative consequences of an algorithmic decision maker.					

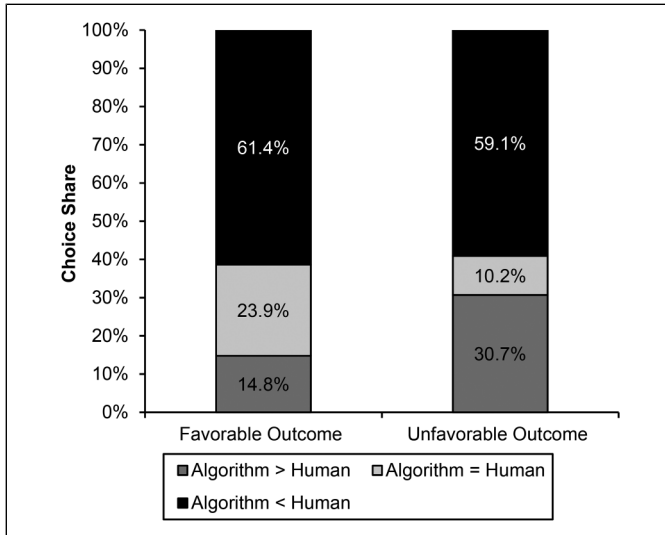


Figure 1. Managers' predictions about our interaction effect on consumers' reactions.

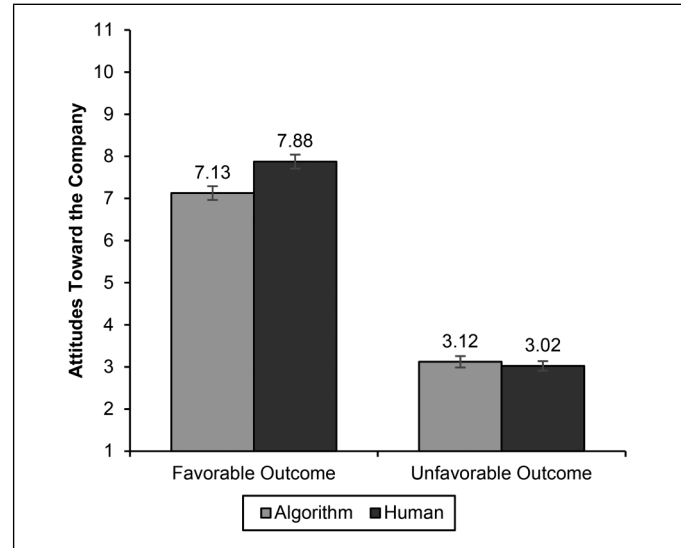


Figure 2. Study 1a results.

Method. We randomly assigned 500 Prolific workers ($M_{\text{age}} = 33.97$ years; 264 women) to one of four conditions in a 2 (decision maker type: algorithm vs. human) \times 2 (decision outcome favorability: favorable vs. unfavorable) between-participants design.

Participants read that they were applying for a business loan (see Web Appendix E). We told participants that their loan applications were either approved or denied by either a loan algorithm or a loan officer. Next, participants indicated their attitudes toward the bank ($\alpha = .99$), as in Study 1a. We also measured participants' WoM intentions using the item made famous by the net promoter score ("On a scale from 0–10, how likely are you to recommend this bank to a friend or colleague?"; 0 = "extremely unlikely," and 10 = "extremely likely").

Results. We first conducted a 2 (decision maker type) \times 2 (decision outcome favorability) ANOVA on attitudes toward the bank. We found a significant main effect of the decision maker type ($M_{\text{algorithm}} = 5.74$, $SD = 3.20$ vs. $M_{\text{human}} = 6.51$, $SD = 3.35$; $F(1, 496) = 22.17$, $p < .001$, $\eta_p^2 = .04$) and of decision outcome favorability ($M_{\text{favorable}} = 8.72$, $SD = 2.18$ vs. $M_{\text{unfavorable}} = 3.55$, $SD = 1.89$; $F(1, 496) = 853.09$, $p < .001$, $\eta_p^2 = .63$). Crucially, we replicated the significant interaction effect on consumers' attitudes ($F(1, 496) = 8.21$, $p = .004$, $\eta_p^2 = .02$): attitudes toward the bank were less positive among participants whose applications were approved by the algorithm than among participants whose applications were approved by the loan officer ($M_{\text{algorithm}} = 8.06$, $SD = 2.51$ vs. $M_{\text{human}} = 9.40$, $SD = 1.51$; $F(1, 496) = 28.56$, $p < .001$, $\eta_p^2 = .05$). Meanwhile, the effect of the decision maker type was significantly attenuated when the applications were denied ($M_{\text{algorithm}} = 3.39$, $SD = 1.79$ vs. $M_{\text{human}} = 3.71$, $SD = 1.97$; $F(1, 496) = 1.71$, $p = .192$).

Next, we conducted an analogous ANOVA on WoM intentions. We found a significant main effect of the decision maker type ($M_{\text{algorithm}} = 4.34$, $SD = 3.02$ vs. $M_{\text{human}} = 5.10$, $SD = 3.29$; $F(1, 496) = 21.06$, $p < .001$, $\eta_p^2 = .04$) and of decision outcome favorability ($M_{\text{favorable}} = 7.15$, $SD = 2.12$ vs. $M_{\text{unfavorable}} = 2.31$, $SD = 2.01$; $F(1, 496) = 722.11$, $p < .001$, $\eta_p^2 = .59$). More importantly, we found a significant interaction effect ($F(1, 496) = 5.04$, $p = .025$, $\eta_p^2 = .01$; see Figure 3): the bank was less likely to be recommended to others by participants whose applications were approved by the algorithm than participants whose applications were approved by the loan officer ($M_{\text{algorithm}} = 6.54$, $SD = 2.27$ vs. $M_{\text{human}} = 7.77$, $SD = 1.75$; $F(1, 496) = 23.25$, $p < .001$, $\eta_p^2 = .04$). Again, however, the effect of the decision maker type on WoM intentions was significantly mitigated when participants' applications were denied ($M_{\text{algorithm}} = 2.10$, $SD = 1.81$ vs. $M_{\text{human}} = 2.52$, $SD = 2.18$; $F(1, 496) = 2.76$, $p = .097$, $\eta_p^2 = .01$).

Discussion of Studies 1a–b. Studies 1a–b demonstrated that the effect of the decision maker type (algorithm vs. human) on consumers' reactions is a function of decision outcome favorability. When participants received a favorable decision outcome, the algorithm (vs. human) decision maker led to less positive reactions toward the company. However, this effect was significantly mitigated when participants received an unfavorable decision outcome.

We note the robustness of our effect thus far: it held in both social (club membership application) and nonsocial (bank loan application) contexts and with two managerially relevant measures of consumers' reactions (attitudes toward the company and WoM intentions). In addition, we demonstrated that our effect is not driven by an assumption that an algorithmic (vs. human) decision would not be the final decision. We consistently observed the key interaction effect regardless of whether we explicitly emphasized that the decision is final.

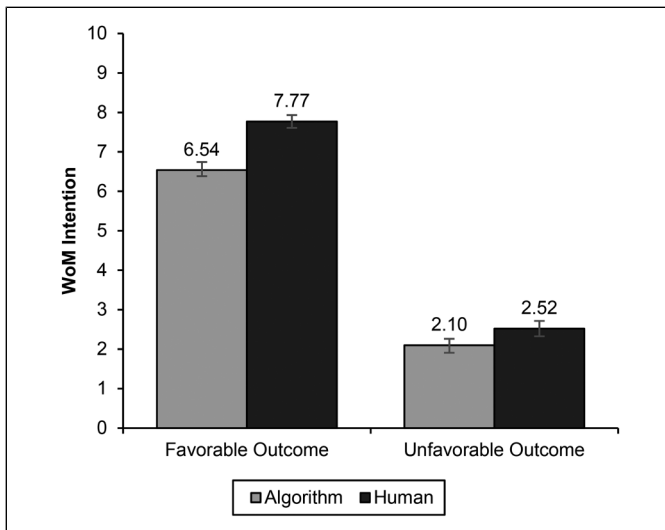


Figure 3. Study 1b results.

Note that our findings contradict the managers' intuitions, so they are managerially informative. Furthermore, it is noteworthy that our interaction effect cannot be explained by the algorithm aversion literature (e.g., Longoni, Bonezzi, and Morewedge 2019), which documents consumers' avoidance of algorithms (vs. humans) without consideration of decision outcome favorability. The interaction effect is therefore distinct from prior findings on general algorithm aversion.

Study 2: Replication with a Real Application Process

The purpose of Study 2 was twofold. First, we aimed to provide a field test of the predicted effect. Participants applied to join a research participant pool run by a research company, Johnson Customer Insight. We examined participants' attitudes toward the research company when their applications were accepted or rejected by either a human or an algorithm. Second, we aimed to rule out an alternative account: inattention to unfavorable information. People tend to avoid unfavorable information that can hurt their self-esteem (Trope and Neter 1994), so they may pay less attention to information (including the decision maker type) that is related to an unfavorable decision outcome. Accordingly, inattention may explain the apparent indifference to the decision maker type for unfavorable decision outcomes. To address this possibility, we directed participants' attention to the decision maker type in all conditions before measuring attitudes toward the company.

Method. We randomly assigned 303 Prolific workers ($M_{\text{age}} = 34.19$ years; 184 women) to one of four conditions in a 2 (decision maker type: algorithm vs. human) \times 2 (decision outcome favorability: favorable vs. unfavorable) between-participants design.

We created a Prolific researcher account under the name Johnson Customer Insight and told Prolific workers (who are essentially gig economy workers whose gig is to be a paid research participant) that the company was creating a research participant pool. Furthermore, we told participants that Johnson Customer Insight was dedicating that particular day to determining the eligibility of applicants for future surveys with generous compensation (see Web Appendix F). Participants were invited to complete an application form, which included questions about their cognitive abilities and their Prolific history; participants were told that the information reflected their diligence and attractiveness as a research participant. After submitting the application, each participant received an application number and was asked to wait while their applications were evaluated; after a few minutes, they received either an acceptance (favorable decision outcome) or a rejection (unfavorable decision outcome). Participants then rated their overall attitude toward the research company ("What is your overall evaluation of Johnson Customer Insight?") on a scale from one to ten stars.

On the next page, we informed participants of the type of decision maker: either one of the coordinators or a computer program designed by the information technology team. Participants completed another measure of attitude: "How do you feel about Johnson Customer Insight now?" (1 = "less positive," and 7 = "more positive"). Finally, we thanked and debriefed participants (including telling them that Johnson Customer Insight was a fictitious company) and paid the promised bonus to all participants.

Attitudes before receiving information about the decision maker type. As we expected, a 2 (decision maker type) \times 2 (decision outcome favorability) ANOVA on the initial rating of the research company indicated a significant main effect of decision outcome favorability ($M_{\text{favorable}} = 8.30$, $SD = 1.67$ vs. $M_{\text{unfavorable}} = 4.20$, $SD = 2.78$; $F(1, 299) = 243.30$, $p < .001$, $\eta_p^2 = .45$). Unsurprisingly, as this measure was taken before the manipulation of the decision maker type, we found neither a main effect of the decision maker type ($F(1, 299) = 1.29$, $p = .256$) nor an interaction effect between the decision maker type and decision outcome favorability ($F(1, 299) = .53$, $p = .466$), indicating successful random assignment.

Core Results. Central to our hypothesis, we tested how the decision maker type affected participants' attitudes as a function of decision outcome favorability. An ANOVA revealed a significant main effect of the decision maker type ($M_{\text{algorithm}} = 3.72$, $SD = 1.59$ vs. $M_{\text{human}} = 4.26$, $SD = 1.85$; $F(1, 299) = 11.04$, $p = .001$, $\eta_p^2 = .04$) and of decision outcome favorability ($M_{\text{favorable}} = 5.03$, $SD = 1.29$ vs. $M_{\text{unfavorable}} = 2.95$, $SD = 1.50$; $F(1, 299) = 175.69$, $p < .001$, $\eta_p^2 = .37$). Crucially, we replicated the key interaction effect ($F(1, 299) = 6.11$, $p = .014$, $\eta_p^2 = .02$; Figure 4): attitudes toward the research company were less positive among participants whose applications were accepted by the algorithm than among participants whose applications were accepted by the coordinator ($M_{\text{algorithm}} = 4.57$,

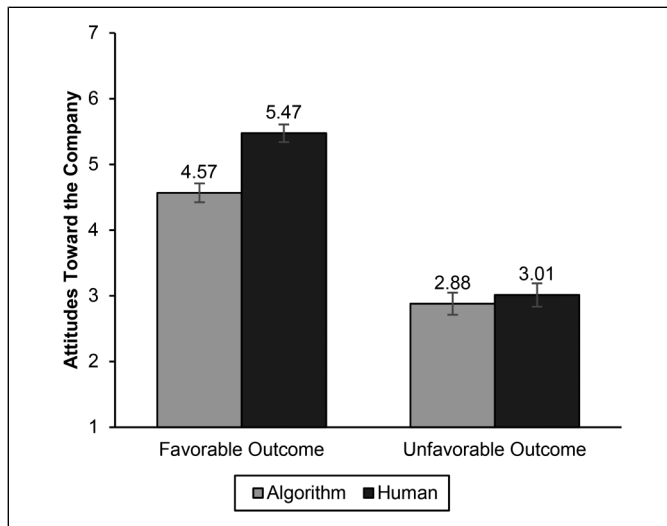


Figure 4. Study 2 results.

SD = 1.23 vs. $M_{\text{human}} = 5.47$, SD = 1.19; $F(1, 299) = 16.84$, $p < .001$, $\eta_p^2 = .05$). Meanwhile, the effect of the decision maker type on the attitudes was significantly mitigated when participants' applications were rejected ($M_{\text{algorithm}} = 2.88$, SD = 1.45 vs. $M_{\text{human}} = 3.01$, SD = 1.54; $F(1, 299) = .36$, $p = .548$). The key interaction effect remained significant after controlling for the initial rating of the research company ($F(1, 298) = 5.90$, $p = .016$, $\eta_p^2 = .02$).

Discussion. Study 2 replicated our key findings in a realistic setting where participants ostensibly were applying to a research company. Furthermore, Study 2 ruled out the alternative account based on inattention to unfavorable information by separating the decision outcome from the decision maker, thereby ensuring attention to the latter.

Studies 3a and 3b: Effect of (Not) Disclosing the Decision Maker

Studies 3a and 3b focused on favorable decision outcomes (as we did not observe a significant effect of the decision maker type for unfavorable decision outcomes in our previous studies). We aimed to clarify whether the effect of the decision maker type on reactions is driven by a positive effect of the human decision maker, a negative effect of the algorithmic decision maker, or both. The distinction is important from the perspectives of managers and business ethics because it has implications for the consequences of disclosing (vs. not disclosing) the decision maker type. Studies 3a–b included a third condition in which consumers are not informed of the decision maker, creating a baseline for assessing the effect of the decision maker type.

Methods. We randomly assigned 403 (Study 3a: $M_{\text{age}} = 32.75$ years; 251 women) and 402 (Study 3b: $M_{\text{age}} = 34.98$ years;

259 women) Prolific workers to one of three conditions (decision maker type: algorithm vs. human vs. unspecified) in a between-participants design.

In Study 3a, participants were applying for membership at Violethall Country Club (see Web Appendix G); depending on the condition, participants learned that their applications were accepted by a club algorithm (algorithm condition), accepted by a club coordinator (human condition), or simply accepted (unspecified decision maker condition). Participants completed the same attitude items ($\alpha = .98$) as in Study 1a. Study 3b was a conceptual replication of Study 3a with one difference: participants read that they were applying for a bank loan (see Web Appendix G). Similar to Study 3a, participants learned that their applications were accepted by a loan algorithm, accepted by a loan officer, or accepted by an unspecified decision maker. Participants rated their attitudes toward the bank ($\alpha = .98$).

Study 3a Results. We observed a significant effect of the decision maker type ($F(2, 400) = 6.78$, $p = .001$, $\eta_p^2 = .03$). Replicating our previous findings, attitudes toward the country club were less positive among participants whose applications were accepted by the algorithm than among participants whose applications were accepted by the club coordinator ($M_{\text{algorithm}} = 6.04$, SD = 2.70 vs. $M_{\text{human}} = 7.21$, SD = 2.83; $F(1, 400) = 12.14$, $p < .001$, $\eta_p^2 = .03$). Attitudes were significantly less positive in the algorithm condition than in the unspecified condition ($M_{\text{algorithm}} = 6.04$, SD = 2.70 vs. $M_{\text{unspecified}} = 6.97$, SD = 2.69; $F(1, 400) = 7.79$, $p = .006$, $\eta_p^2 = .02$), but attitudes were similar in the human and unspecified conditions ($M_{\text{human}} = 7.21$, SD = 2.83 vs. $M_{\text{unspecified}} = 6.97$, SD = 2.69; $F < 1$, $p = .482$; Figure 5).

Study 3b Results. We observed a significant effect of the decision maker type ($F(2, 399) = 13.79$, $p < .001$, $\eta_p^2 = .06$). Participants whose loan applications were accepted by the algorithm indicated less positive attitudes toward the bank than both participants whose loan applications were accepted by the loan officer ($M_{\text{algorithm}} = 7.38$, SD = 2.39 vs. $M_{\text{human}} = 8.50$, SD = 1.94; $F(1, 399) = 18.85$, $p < .001$, $\eta_p^2 = .05$) and participants whose loan applications were accepted by an unspecified decision maker ($M_{\text{algorithm}} = 7.38$, SD = 2.39 vs. $M_{\text{unspecified}} = 8.59$, SD = 1.98; $F(1, 399) = 22.29$, $p < .001$, $\eta_p^2 = .05$). Again, the difference between the human and unspecified conditions was not significant ($M_{\text{human}} = 8.50$, SD = 1.94 vs. $M_{\text{unspecified}} = 8.59$, SD = 1.98; $F < 1$, $p = .711$; Figure 5).

Discussion of Studies 3a and 3b. Studies 3a and 3b clarify that the effect of the decision maker type in favorable decisions occurs because the disclosure of an algorithmic decision maker hurts consumers' attitudes relative to a baseline of an undisclosed decision maker. These findings have implications for decision transparency, which we discuss in the "General Discussion" section.

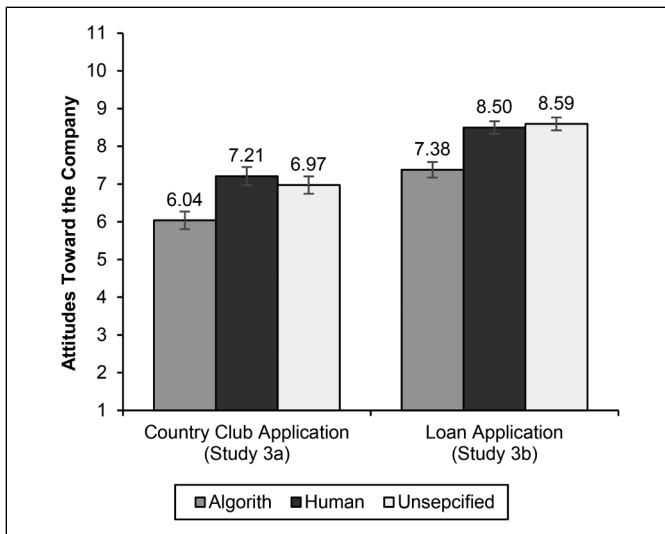


Figure 5. Results of Studies 3a and 3b.

What Psychological Mechanisms Differentiate Consumers' Reactions to Algorithmic and Human Decision Makers?

We proposed that consumers react less positively when their applications are accepted by algorithms (vs. humans) because they find it relatively more difficult to internalize an acceptance made by an algorithm (vs. by a human). By contrast, when a decision outcome is unfavorable, consumers readily externalize the decision outcome, so they react similarly toward the company regardless of the decision maker. We directly examined this attribution mechanism through mediation (Studies 4 and 6) and moderation (Study 5).

Study 4: Mediation by Internal Attribution

Study 4 examined the mediating role of attribution. We predicted that algorithmic (vs. human) decision makers would elicit distinct attributions as a function of decision outcome favorability, and the attributions would mediate our key interaction effect on attitudes toward the company.

Method. We randomly assigned 600 Prolific workers to one of four conditions in a 2 (decision maker type: algorithm vs. human) \times 2 (decision outcome favorability: favorable vs. unfavorable) between-participants design. Our final data set consisted of 571 participants ($M_{\text{age}} = 33.84$ years; 249 women) who passed our attention check.²

As in Study 3a, participants read that they were applying for membership at a country club (see Web Appendix H); participants learned that their applications were either accepted or rejected by

either the country club algorithm or the country club coordinator, and they indicated their attitudes toward the country club ($\alpha = .99$) as in Study 3a. Next, we measured internal attributions (adapted from Russell [1982]): "To what extent do you feel this decision [reflects something about yourself/can be attributed to something about yourself/is due to your personal qualities or behaviors]?" (1 = "not at all," and 11 = "very much"; $\alpha = .91$).

Results. We conducted a 2 (decision maker type) \times 2 (decision outcome favorability) ANOVA on attitudes toward the country club. We found a significant main effect of the decision maker type ($M_{\text{algorithm}} = 5.01$, $SD = 2.95$ vs. $M_{\text{human}} = 5.91$, $SD = 3.11$; $F(1, 567) = 8.62$, $p = .003$, $\eta_p^2 = .01$) and of decision outcome favorability ($M_{\text{favorable}} = 7.06$, $SD = 2.88$ vs. $M_{\text{unfavorable}} = 3.85$, $SD = 2.30$; $F(1, 567) = 211.62$, $p < .001$, $\eta_p^2 = .27$). Again, we found a marginally significant interaction between the decision maker type and decision outcome favorability ($F(1, 567) = 3.66$, $p = .056$, $\eta_p^2 = .01$; see Figure 6): attitudes toward the country club were less positive among participants whose applications were accepted by the algorithm than among participants whose applications were accepted by the coordinator ($M_{\text{algorithm}} = 6.49$, $SD = 2.88$ vs. $M_{\text{human}} = 7.54$, $SD = 2.79$; $F(1, 567) = 11.82$, $p < .001$, $\eta_p^2 = .02$). Meanwhile, this difference was significantly mitigated among participants whose applications were rejected ($M_{\text{algorithm}} = 3.74$, $SD = 2.36$ vs. $M_{\text{human}} = 3.97$, $SD = 2.23$; $F < 1$, $p = .471$).

We conducted an analogous ANOVA on internal attributions. We found a significant effect of the decision maker type ($M_{\text{algorithm}} = 6.54$, $SD = 2.79$ vs. $M_{\text{human}} = 7.27$, $SD = 2.69$; $F(1, 567) = 8.01$, $p = .005$, $\eta_p^2 = .01$) and of decision outcome favorability ($M_{\text{favorable}} = 7.54$, $SD = 2.47$ vs. $M_{\text{unfavorable}} = 6.27$, $SD = 2.90$; $F(1, 567) = 30.15$, $p < .001$, $\eta_p^2 = .05$). Importantly, we found a significant interaction effect ($F(1, 567) = 10.11$, $p = .002$, $\eta_p^2 = .02$; see Figure 6): the internal attribution was weaker when the acceptance decision was made by the algorithm than when it was made by the club coordinator ($M_{\text{algorithm}} = 6.82$, $SD = 2.54$ vs. $M_{\text{human}} = 8.15$, $SD = 2.24$; $F(1, 567) = 18.17$, $p < .001$, $\eta_p^2 = .03$). The effect of the decision maker type was significantly mitigated for the internal attribution of the rejection decision ($M_{\text{algorithm}} = 6.30$, $SD = 2.98$ vs. $M_{\text{human}} = 6.22$, $SD = 2.81$; $F < 1$, $p = .806$).

Finally, we ran a moderated mediation analysis (PROCESS Model 8, 10,000 bootstrapped samples; Hayes 2013) with attitudes toward the country club as the dependent variable, decision maker type ($-1 = \text{algorithm}$, $1 = \text{human}$) as the independent variable, decision outcome favorability ($-1 = \text{unfavorable}$, $1 = \text{favorable}$) as the moderator, and internal attribution as the mediator (see Figure 7).³ As we predicted, we found a significant moderated mediation effect ($B = .16$, 95% confidence interval [CI] = [.0536, .2780]). For a favorable decision outcome, the indirect effect of the decision maker type through internal attribution was significant ($B = .15$, 95% CI = [.0720, .2392]), suggesting

² We added an instructional manipulation check due to the concern about poor data quality during the COVID-19 crisis. We found the same results regardless of whether we filtered out those who failed the attention check (for details, see Web Appendices H and I).

³ The reported mediation model is in line with the criteria for meaningful mediation proposed by Pieters (2017).

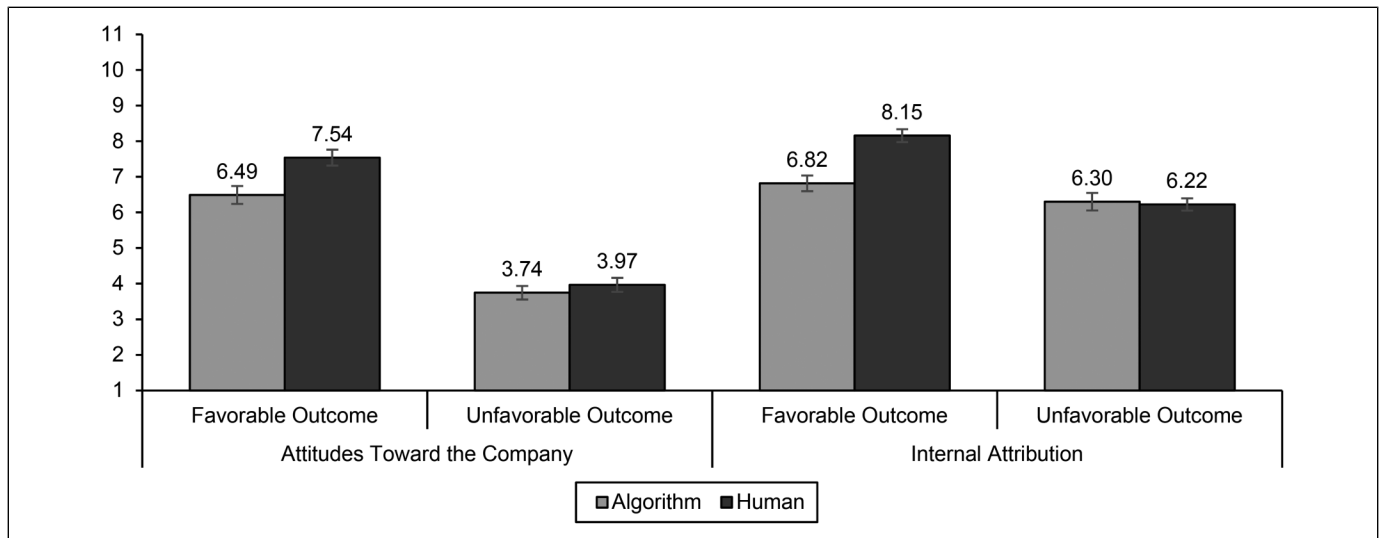


Figure 6. Study 4 results.

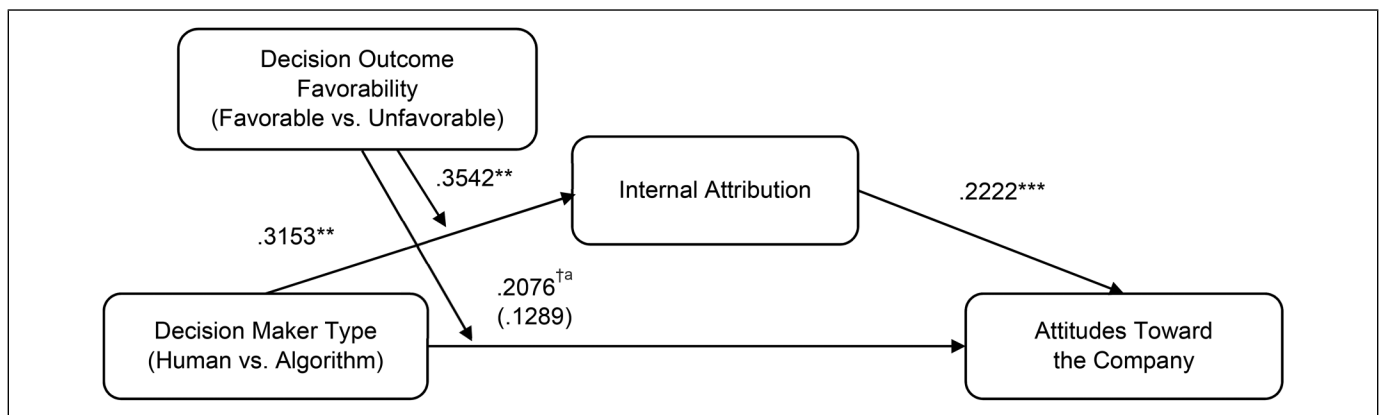


Figure 7. Study 4 results (moderated mediation model).

Notes: † $p < .06$; * $p < .05$; ** $p < .01$; *** $p < .001$. ^aEstimated coefficient for the decision outcome favorability \times decision maker type interaction in the model without (with) the mediator.

that the less positive reaction to the country club after receiving a decision from an algorithm (vs. a human) was driven by the weaker internal attribution of the favorable decision. For an unfavorable decision outcome, however, the corresponding indirect effect was not significant ($B = -.01$, 95% CI = $[-.0864, .0694]$).

In summary, Study 4 directly examined the proposed mechanism and found evidence that decision outcome favorability affects the internal attribution process of algorithmic versus human decisions, thereby leading to divergent reactions to the decisions made by the different decision makers.

Study 5: Moderated Mediation by Internal Attribution of a Favorable Outcome

We proposed that consumers react more positively when a favorable decision is made by a human (vs. an algorithm) because a human decision maker facilitates the internal attribution of the decision outcome more. If this is the case, this effect

should be mitigated when the decision outcome is not diagnostic of consumers' personal characteristics (e.g., the decision was made at random), in which case there is little justification for internal attribution regardless of the decision maker type. Study 5 tested this prediction by manipulating self-diagnostics; the decision was based on either an evaluation of the consumer's application or a raffle. Furthermore, Study 5 increased the generalizability of our effect by replicating it in another managerially relevant context: networking platforms.

Method. We randomly assigned 501 Prolific workers to one of four conditions in a 2 (decision maker type: algorithm vs. human) \times 2 (decision method: evaluation vs. raffle) between-participants design. Our final data set consisted of 443 participants ($M_{age} = 39.33$ years; 222 women) who passed our attention check.

Participants read that they were applying to join a business networking community, NetWorkLink (see Web Appendix I). Participants learned that their applications were accepted by

either the club algorithm or the club coordinator, and the decision method involved either an evaluation of the applications or a raffle (i.e., random selection). Finally, we measured participants' attitudes toward the networking club ($\alpha = .98$) and internal attributions ($\alpha = .95$) by using the same items as in Study 4.

Results. A 2 (decision maker type) \times 2 (decision method) ANOVA on attitudes revealed no significant main effect of the decision maker type ($M_{\text{algorithm}} = 6.44$, $SD = 2.59$ vs. $M_{\text{human}} = 6.55$, $SD = 2.82$; $F(1, 439) = 1.12$, $p = .290$), but a significant effect of the decision method ($M_{\text{evaluation}} = 7.30$, $SD = 2.43$ vs. $M_{\text{raffle}} = 5.69$, $SD = 2.73$; $F(1, 439) = 44.54$, $p < .001$, $\eta_p^2 = .09$). Importantly, we found a marginally significant interaction effect ($F(1, 439) = 3.44$, $p = .064$, $\eta_p^2 = .01$; see Figure 8). When the acceptance decision was based on an evaluation of the applications (i.e., when the decision was self-diagnostic), we replicated our previous findings: attitudes toward the networking club were less positive among participants whose applications were accepted by the algorithm than among participants whose applications were accepted by the club coordinator ($M_{\text{algorithm}} = 6.98$, $SD = 2.34$ vs. $M_{\text{human}} = 7.70$, $SD = 2.49$; $F(1, 439) = 4.25$, $p = .040$, $\eta_p^2 = .01$). However, when the acceptance decision was based on a raffle (i.e., when the decision was not self-diagnostic), the decision maker type did not significantly affect participants' attitudes ($M_{\text{algorithm}} = 5.80$, $SD = 2.74$ vs. $M_{\text{human}} = 5.60$, $SD = 2.73$; $F < 1$, $p = .574$).

An analogous ANOVA on internal attribution revealed a significant main effect of the decision maker type ($M_{\text{algorithm}} = 5.59$, $SD = 2.99$ vs. $M_{\text{human}} = 6.07$, $SD = 3.29$; $F(1, 439) = 9.93$, $p = .002$, $\eta_p^2 = .02$) and of the decision method ($M_{\text{evaluation}} = 7.47$, $SD = 2.26$ vs. $M_{\text{raffle}} = 4.17$, $SD = 3.05$; $F(1, 439) = 179.62$, $p < .001$, $\eta_p^2 = .29$). Crucially, we again found a significant interaction effect ($F(1, 439) = 6.01$, $p = .015$, $\eta_p^2 = .01$; Figure 8): when the decision was based on an evaluation of the applications, the internal attribution of the acceptance was weaker among participants whose applications were accepted by the algorithm than among participants whose applications were accepted by the club coordinator ($M_{\text{algorithm}} = 6.84$, $SD = 2.29$ vs. $M_{\text{human}} = 8.25$, $SD = 1.96$; $F(1, 439) = 15.69$, $p < .001$, $\eta_p^2 = .03$). However, when the acceptance decision was based on a raffle, the decision maker type did not significantly affect the internal attribution made by participants ($M_{\text{algorithm}} = 4.08$, $SD = 3.04$ vs. $M_{\text{human}} = 4.25$, $SD = 3.07$; $F < 1$, $p = .621$).

To test whether our key effect is mediated by the internal attribution of the favorable decision outcome, we conducted a moderated mediation analysis (PROCESS Model 8, 95% CI, 10,000 bootstrapped samples; Hayes 2013) with attitudes toward the networking club as the dependent variable, decision maker type ($-1 = \text{algorithm}$, $1 = \text{human}$) as the independent variable, decision method ($-1 = \text{raffle}$, $1 = \text{evaluation}$) as the moderator, and internal attribution as the mediator. In line with our theory, we found a significant moderated mediation effect ($B = .26$, 95% CI = [.0516, .4765]): when the decision was based on an evaluation of the applications and thus self-diagnostic (such that participants were motivated or able to internally attribute the favorable outcome), the indirect effect through internal attribution was significant ($B = .29$, 95% CI = [.1668, .4344]),

suggesting that the more positive attitude toward the networking club after receiving a decision from the human (vs. algorithm) was driven by the stronger internal attribution of the favorable decision. When the decision was based on a raffle and thus was not self-diagnostic, however, the indirect effect was not significant ($B = .04$, 95% CI = [−.1327, .2088]).

In summary, Study 5 corroborates our attribution mechanism by demonstrating moderation by the self-diagnosticsity of the decision. Together, the results of Studies 4 and 5 provide converging evidence that supports our attribution mechanism.

Study 6: External Attribution of an Unfavorable Decision Outcome

We proposed that the decision maker type has an attenuated effect on consumers' reactions following an unfavorable decision outcome because consumers can readily engage in external attribution of an unfavorable decision outcome regardless of the decision maker. Consumers perceive both algorithmic and human decision makers to have weaknesses: humans are less objective (Lee 2018), and algorithms neglect the uniqueness of each individual (e.g., Longoni, Bonezzi, and Morewedge 2019). Accordingly, when consumers receive an unfavorable decision outcome, they can blame a human decision maker for a lack of objectivity and blame an algorithmic decision maker for neglecting their individual uniqueness. We argued that these countervailing effects cancel each other out, resulting in consumers' relative indifference to the type of decision maker. In Study 6, we tested this proposition by measuring consumers' perceptions of the decision maker's objectivity and consideration of individual uniqueness.

Method. In this preregistered study (aspredicted.org/ah2sc.pdf), we randomly assigned 626 MTurk workers ($M_{\text{age}} = 35.51$ years; 332 women) to one of two conditions (decision maker type: algorithm vs. human) in a between-participants design.⁴

Participants read that they were applying for membership at a country club and their applications were rejected by either the club algorithm or the club coordinator (see Web Appendix J). Participants then assessed the decision maker's objectivity and consideration of the applicant's uniqueness (the order of the measures was randomized). Specifically, participants answered three items about the decision maker's objectivity: "To what extent do you think [this algorithm/club coordinator] [made an unbiased assessment of your application/made an unemotional assessment of your application/assessed your application rationally]?" (1 = "not at all," and 11 = "very much"; $\alpha = .71$). Participants also answered three items about the decision maker's consideration of their application's uniqueness: "To what extent do you think this [algorithm/club coordinator] [recognized the uniqueness of your application/considered the unique aspects of your application / tailored the decision to your unique case]?" (adapted

⁴ As stated in our preregistration form, we targeted 600 participants, but the actual sample size differed for reasons beyond our control (e.g., participants not claiming their participation on MTurk).

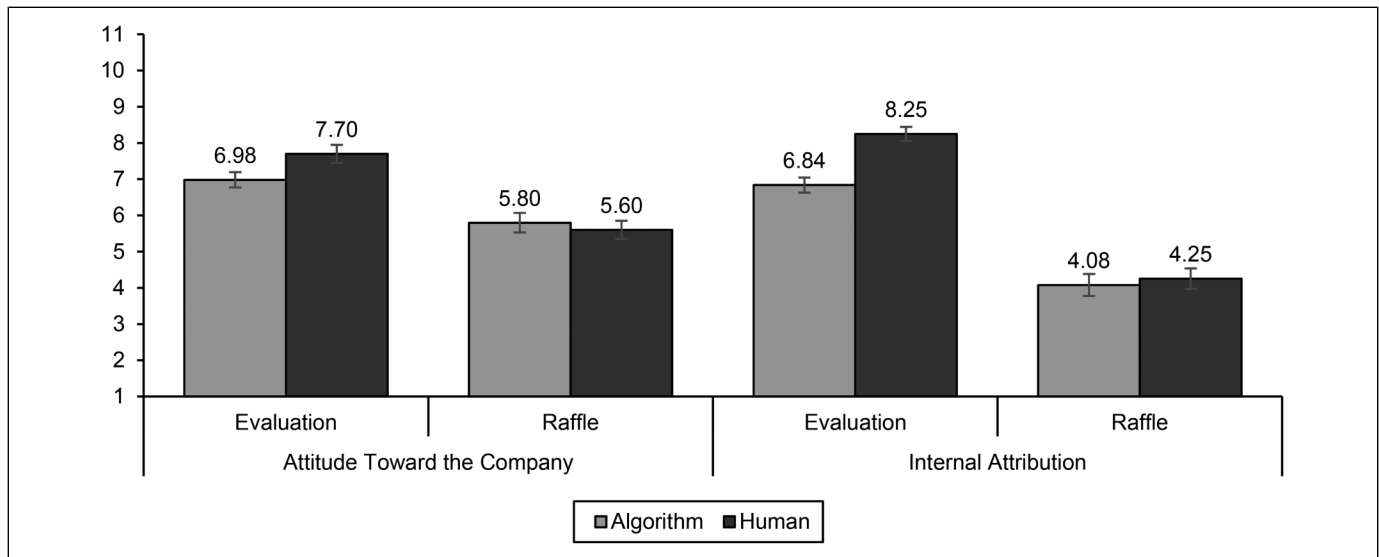


Figure 8. Study 5 results.

from Longoni, Bonezzi, and Morewedge [2019]; 1 = “not at all,” and 11 = “very much”; $\alpha = .93$). Lastly, participants completed the same attitude items as in Study 5 ($\alpha = .97$).

Results. In line with our previous findings, and as preregistered, there was no significant effect of the decision maker type on attitudes toward the country club ($M_{\text{algorithm}} = 4.84$, $SD = 2.62$ vs. $M_{\text{human}} = 4.78$, $SD = 2.61$; $F(1, 624) < 1$, $p = .782$; see Figure 9). Crucially, we found a significant effect of the decision maker type on participants’ perceptions of the decision maker’s objectivity and consideration of uniqueness: the club coordinator (vs. algorithm) was perceived as less objective ($M_{\text{human}} = 5.95$, $SD = 2.47$ vs. $M_{\text{algorithm}} = 7.07$, $SD = 2.28$; $F(1, 624) = 34.76$, $p < .001$, $\eta_p^2 = .05$), whereas the algorithm (vs. club coordinator) was perceived as less sensitive to the applicant’s uniqueness ($M_{\text{algorithm}} = 4.41$, $SD = 2.81$ vs. $M_{\text{human}} = 5.35$, $SD = 2.74$; $F(1, 624) = 17.67$, $p < .001$, $\eta_p^2 = .03$).

Finally, we conducted a mediation analysis (PROCESS Model 4; 95% CI, 10,000 bootstrapped samples; Hayes 2013) with attitudes toward the country club as the dependent variable, the decision maker type ($-1 = \text{algorithm}$, $1 = \text{human}$) as the independent variable, and perceived objectivity and uniqueness consideration as the two mediators. In line with the preregistered prediction, the indirect effects of the decision maker type via the two mediators were significant in opposite directions (perceived objectivity: $B = -.19$, 95% CI = $[-.2905, -.1144]$; uniqueness consideration: $B = .17$, 95% CI = $[.0847, .2547]$), explaining the relative indifference to the decision maker type for unfavorable decision outcomes. The direct effect was not significant ($B = .00$, 95% CI = $[-.1803, .1790]$; see Figure 10).

In summary, Study 6 corroborates our theory that consumers make external attributions about unfavorable decision outcomes for both human and algorithmic decision makers, facilitated by the perceived weakness of the decision maker—human decision makers have poor objectivity, while algorithmic decision makers do not consider each applicant’s unique characteristics.

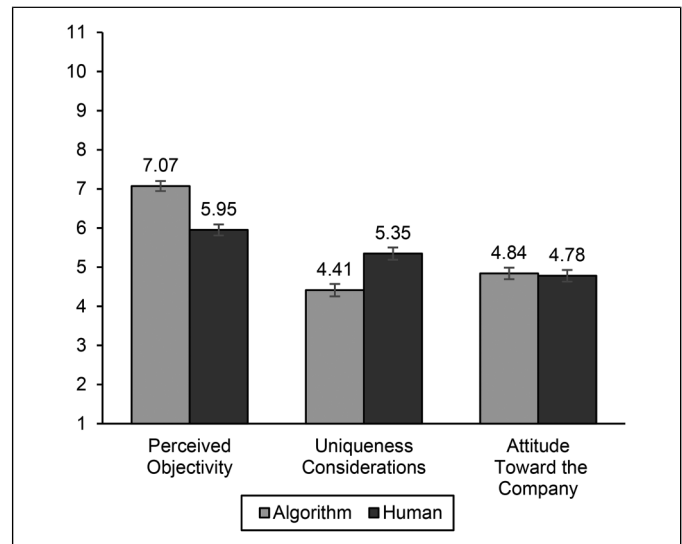


Figure 9. Study 6 results.

In addition, these results contradict an alternative explanation based on psychological numbness following a rejection, which could plausibly lead to an indifference to the type of decision maker for unfavorable decision outcomes. However, the psychological numbness account predicts psychological deactivation (including disengagement from attributional processes), which does not explain the parallel mediation processes that we found in Study 6.

Study 7: Effect of Human Decision Making Versus Mere Human Observation

One could argue that participants in our previous studies reacted more positively to acceptance by humans due to social presence (Argo, Dahl, and Manchanda 2005; McFerran and Argo 2014); when an algorithm makes an acceptance decision, no social

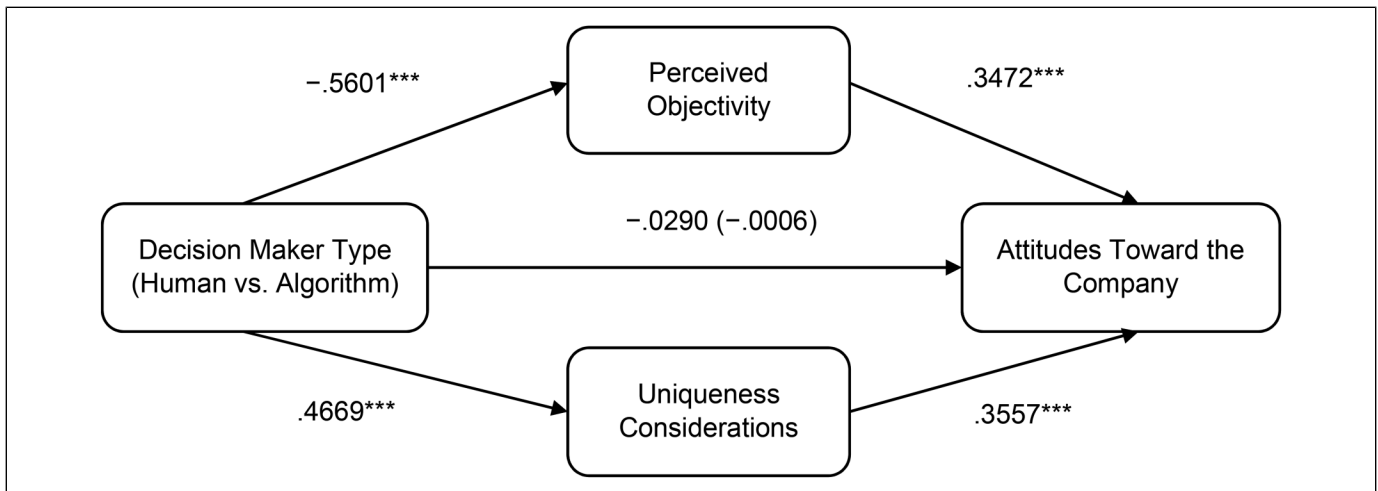


Figure 10. Study 6 results (mediation model).

*** $p < .001$.

agent is aware of the outcome. By contrast, the social presence of the human decision maker might lead participants to feel more positive about the outcome and thus react more positively toward the company.

Although it cannot explain several findings in the previous studies (e.g., the moderation in Study 5), we conducted Study 7 to directly test the alternative account of social presence by adding a new condition in which a human monitored (but did not interfere with) the algorithm's decisions. If social presence accounts for our effect, consumers should react similarly when a human makes the decision versus merely observes the favorable outcome. If our effect is due to distinct attributions under human versus algorithmic decision makers, however, then reactions should be similar when an algorithm makes the decision with versus without a human monitoring the decision process.

Method. We randomly assigned 597 MTurk workers ($M_{\text{age}} = 35.42$ years; 318 women) to one of six conditions in a 3 (decision maker type: algorithm only vs. human vs. algorithm with human monitoring) \times 2 (decision outcome favorability: favorable vs. unfavorable) between-participants design. The procedure of this study was similar to that of Study 1a with the addition of the third decision maker condition, in which the club coordinator ran and monitored the algorithm's evaluation of applications (see Web Appendix K). Participants completed the same attitude scale as in Study 6 ($\alpha = .98$).

Results. We found a significant main effect of the decision maker type ($M_{\text{algorithm only}} = 5.50$, $SD = 2.98$ vs. $M_{\text{human}} = 5.76$, $SD = 3.17$ vs. $M_{\text{algorithm w/ human monitoring}} = 5.17$, $SD = 2.83$; $F(2, 591) = 4.52$, $p = .011$, $\eta_p^2 = .02$) and of decision outcome favorability ($M_{\text{favorable}} = 7.18$, $SD = 2.67$ vs. $M_{\text{unfavorable}} = 3.76$, $SD = 2.24$; $F(1, 591) = 295.05$, $p < .001$, $\eta_p^2 = .33$). The interaction effect was marginally significant ($F(2, 591) = 2.45$, $p = .087$, $\eta_p^2 = .01$; see Figure 11).⁵

In the favorable decision outcome condition, the simple effect of the decision maker type was significant ($F(2, 591) = 5.67$, $p = .004$, $\eta_p^2 = .02$). Replicating our previous studies, attitudes toward the country club were less positive among participants whose applications were accepted by the algorithm than among participants whose applications were accepted by the club coordinator ($M_{\text{algorithm only}} = 7.09$, $SD = 2.68$ vs. $M_{\text{human}} = 7.82$, $SD = 2.59$; $F(1, 591) = 4.36$, $p = .037$, $\eta_p^2 = .01$). Moreover, we found a significant difference in attitudes between the human condition and algorithm-with-human-monitoring conditions; attitudes toward the country club were less positive in the latter condition ($M_{\text{human}} = 7.82$, $SD = 2.59$ vs. $M_{\text{algorithm w/ human monitoring}} = 6.68$, $SD = 2.65$; $F(1, 591) = 11.13$, $p < .001$, $\eta_p^2 = .02$). There was no significant difference in attitudes between the algorithm-only and algorithm-with-human-monitoring conditions ($M_{\text{algorithm only}} = 7.09$, $SD = 2.68$ vs. $M_{\text{algorithm w/ human monitoring}} = 6.68$, $SD = 2.65$; $F(1, 591) = 1.40$, $p = .238$). In the unfavorable decision outcome condition, attitudes toward the country club were not influenced by the decision maker type ($F(2, 591) = 1.37$, $p = .255$).

Discussion. Consistent with our attribution account and inconsistent with the social presence account, we found that consumers react more positively when an acceptance decision is made by a human than by an algorithm, regardless of whether a human monitors the algorithm's decisions. At first glance, our findings may seem contradictory to those of Study 9 in Longoni, Bonezzi, and Morewedge (2019), in which individuals were more likely to use a medical algorithm if it was complemented by a human dermatologist (i.e., a dermatologist reviewed the algorithm's diagnosis and made a final decision). The studies,

⁵ To test our core hypothesis more directly, we conducted a planned contrast analysis in which we aggregated the algorithm-only and algorithm-with-human-monitoring conditions, which we hypothesized to show the same results. In this case, we observed a significant interaction effect ($F(1, 591) = 4.98$, $p < .05$, $\eta_p^2 = .01$).

however, have a key difference: a human was actively engaged in the decision-making process in Longoni, Bonezzi, and Morewedge's study, whereas a human merely observed the algorithm and could not alter its decisions in our study.

What Can Managers Do to Mitigate the Negative Effects of Algorithms?

We consistently observed that consumers react less positively when a favorable decision is made by an algorithm (vs. a human). Study 8 examined a potential solution: anthropomorphizing the algorithm. Extant work suggests that humanizing a nonhuman agent (e.g., referring to an object with a personal name) leads people to attribute human-like abilities to it (Crolic et al. 2022; Epley 2018). We proposed that humanizing an algorithm should more closely align consumers' perceptions of a human decision maker and an algorithmic decision maker, enabling the human-like algorithm to lead to more positive reactions than the non-human-like algorithm.

Study 8: Humanizing Algorithms to Mitigate Negative Consequences: Attitudes Toward the Company

Method. We randomly assigned 601 Prolific workers ($M_{\text{age}} = 33.52$ years; 316 women) to one of three conditions (decision maker type: algorithm vs. human vs. human-like algorithm) in a between-participants design.

The procedure of Study 8 was similar to that of Study 1a. Participants were told they were applying for membership at a country club (see Web Appendix L); depending on the condition, the decision maker was described as a country club algorithm (depicted as a robot), a country club coordinator named Sam (depicted as a woman), or a country club algorithm named Sam (depicted as a cartoonized version of the picture of the woman from the human condition). All participants were informed that their applications were accepted. We asked participants to indicate their attitudes toward the country club using the same items as in Study 7 ($\alpha = .98$).

Pretest. We conducted a separate pretest to examine whether a human-like algorithm seems more human than a non-human-like algorithm. We presented 100 Prolific workers ($M_{\text{age}} = 30.23$ years; 41 women) with the information from the algorithm and human-like algorithm conditions in the main study. We then asked participants, "To what extent do you think that [the country club algorithm/Sam] has some human-like qualities?" and "To what extent do you think [the country club algorithm/Sam] seems like a person?" (1 = "not at all," and 7 = "very much"; $r = .80$; adapted from Kim and McGill [2018]). Results confirmed that our manipulation was successful: participants perceived the human-like algorithm to be more human than the algorithm was ($M_{\text{human-like algorithm}} = 3.92$, $SD = 1.56$ vs. $M_{\text{algorithm}} = 2.65$, $SD = 1.29$; $F(1, 598) = 19.86$, $p < .001$, $\eta_p^2 = .17$).

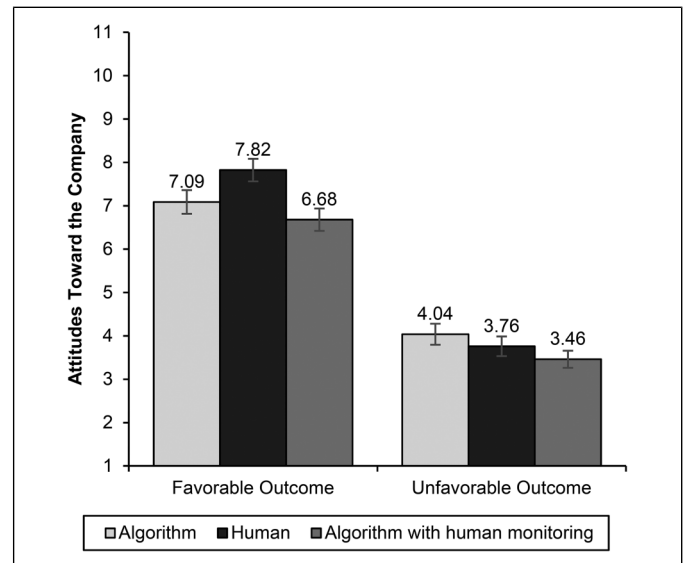


Figure 11. Study 7 results.

Results. In our main study, we conducted a one-way ANOVA on participants' attitudes toward the country club. Replicating our previous findings, the decision maker type had a significant effect ($F(2, 598) = 4.69$, $p = .009$, $\eta_p^2 = .02$). Attitudes toward the club were less positive among participants whose applications were accepted by the algorithm than among participants whose applications were accepted by the club coordinator ($M_{\text{algorithm}} = 7.07$, $SD = 2.67$ vs. $M_{\text{human}} = 7.87$, $SD = 2.52$; $F(1, 598) = 8.88$, $p = .003$, $\eta_p^2 = .01$). Importantly, humanizing the algorithm led to significantly more positive attitudes toward the country club ($M_{\text{algorithm}} = 7.07$, $SD = 2.67$ vs. $M_{\text{human-like algorithm}} = 7.64$, $SD = 2.82$; $F(1, 598) = 4.46$, $p = .035$, $\eta_p^2 = .01$) such that attitudes were similar whether the decision maker was the human-like algorithm or the club coordinator ($M_{\text{human-like algorithm}} = 7.64$, $SD = 2.82$ vs. $M_{\text{human}} = 7.87$, $SD = 2.52$; $F < 1$, $p = .384$).

Discussion. Building on our prior studies' finding that consumers react less positively when a favorable decision is made by an algorithm (vs. a human), Study 8 tested a potential solution: anthropomorphizing the algorithm. Attitudes toward the company were more positive when the favorable decision was made by a human-like (vs. a non-human-like) algorithm.

General Discussion

The current research reveals that consumers react differently to a company that uses algorithmic (vs. human) decision makers as a function of decision outcome favorability: Consumers react less positively toward a company when they receive a favorable decision made by an algorithm than by a human; however, this difference is significantly mitigated when the decision outcome is unfavorable. The effect is driven by different attributions: consumers find it relatively more difficult to internalize a favorable decision made by an algorithm (vs. a human), while it is similarly easy to externalize an unfavorable

decision made by either type of decision maker. Finally, we demonstrate that humanizing the algorithm can mitigate the relatively less positive reaction to an algorithmic (vs. a human) decision maker in the setting of favorable decision outcomes.

Alternative Accounts

Several alternative accounts merit discussion. We review these accounts and discuss how our findings and study design rule them out. In addition, we have direct evidence, in the form of both mediation and moderation, that supports attribution processes (Studies 4–6).

First, one might argue that consumers care about a favorable outcome being witnessed by (rather than made by) another human and that it is this mere human presence that leads to more positive reactions to human decision makers. Against such a social presence account, however, participants in Study 7 reacted more positively only when a human (vs. an algorithm) made the favorable decision on them, but not when a human merely observed the algorithm and thus knew about the favorable decision outcome.

Second, our results might be explained by social cues. For instance, being accepted by a human might create a sense of social belonging, while being evaluated by an algorithm might engender feelings of disrespect. However, we observed the key interaction effect even in contexts in which social relationships are less salient (i.e., business loan application, market research participant panel). Moreover, if algorithmic (vs. human) evaluation creates feelings of disrespect, we should have found a main effect of the decision maker type, but not necessarily the interaction effect between the decision maker type and decision outcome favorability.

Third, consumers might pay less attention to unfavorable information about the self because they inherently avoid information that can hurt their self-esteem (Trope and Neter 1994). In this regard, consumers might be inattentive to any unfavorable information about the self, including the type of decision maker that was involved in the unfavorable decision. However, we replicated our interaction effect even when we explicitly directed participants' attention to the decision maker type (Study 2), ruling out the inattention account. In addition, the inattention account does not explain the two opposing mediation processes for unfavorable decisions in Study 6.

Fourth, one could argue that psychological numbness explains the relative indifference to the decision maker type for unfavorable decision outcomes. An experience of social exclusion (e.g., ostracism) can impair people's emotional sensitivity and cognitive function (Williams 2007), and even social rejections by nonhuman agents (e.g., robots) can lead to negative psychological consequences (Nash et al. 2018). If psychological numbness explains our effect, however, then it should be limited to contexts in which social relationships are salient—but our effect is significant in nonsocial contexts as well. Moreover, the psychological numbness account would predict that consumers who receive unfavorable decision outcomes

should be less likely to engage in any cognitive processes including attributions, but in Study 6, participants' reactions to unfavorable decisions were due to external attribution processes.

Fifth, one could argue that the perceived fairness of algorithmic versus human decision makers explains our results. Consumers are known to perceive decisions made by algorithms (vs. humans) as less fair (Lee 2018). Differential perceptions of decision fairness should produce a main effect of the decision maker type, but not necessarily the interaction effect that we observed consistently. Nonetheless, we conducted a follow-up study (see Web Appendix N) that measured the perceived fairness of the decision. We found a main effect of the decision maker type: participants perceived the human to be fairer than the algorithm ($M_{\text{human}} = 4.12$, $SD = 1.55$ vs. $M_{\text{algorithm}} = 3.28$, $SD = 1.57$; $F(1, 317) = 23.63$, $p < .001$, $\eta_p^2 = .07$). However, this effect was not moderated by decision outcome favorability ($F < 1$, $p = .578$), ruling out perceived fairness as a viable explanation for our effect.

Finally, one could be concerned about scale insensitivity as an explanation for the interaction between the decision maker type and decision outcome favorability. Specifically, one could argue that there could be differences in consumers' reactions to unfavorable decision outcomes by different decision maker types, but that our measures are not sensitive enough to capture these differences (e.g., because such reactions are in general quite negative). Study 2 rules out this concern. In Study 2, we first elicited a response to the outcome (favorable or unfavorable) and then provided information about the decision maker to probe how the information of the decision maker type changes participants' attitudes toward the company. In this study, we still observed that participants reacted to rejections by humans and algorithms similarly.

Theoretical Implications

The current research makes several theoretical contributions. Extending prior research on how consumers decide between algorithms and humans (Dietvorst, Simmons, and Massey 2015; Longoni, Bonezzi, and Morewedge 2019), we shed light on how consumers' reactions to a self-diagnostic decision (i.e., decisions about the consumers themselves) are affected by the decision maker type (human vs. algorithm). Second, our work identifies a theoretically and managerially relevant moderator (decision outcome favorability) that has been underexplored in the existing literature on algorithmic decision making. Finally, we extend the existing work on consumers' perceptions of the different decision makers (e.g., Lee 2018) by examining how algorithmic (vs. human) decisions prompt different attributions as a function of decision outcome favorability. In doing so, our research marries the social psychology literature on attribution processes with the marketing literature on algorithmic decision making.

Our article opens several avenues for future research. First, future research could examine consumers' perceptions of decisions that are made through human–algorithm collaboration. Consumers may react differently depending on the nature of the collaboration (e.g., who conducts the first round of screening vs. makes a final decision). Second, future research could

examine whether our interactive effect is influenced by the nature of decision criteria. Companies use a variety of criteria to accept or reject consumers (e.g., high/low performance, passing/failing a threshold). In our studies, we did not specify why an application was accepted or rejected. We encourage researchers to investigate whether specific reasoning affects the interaction between the decision maker type and decision outcome favorability. Third, even though big data has improved the quality of decisions made by both humans and algorithms, there are still concerns about the representativeness of data used by firms (Bolukbasi et al. 2016). Given that minority groups are often underrepresented in data sets (Sheikh 2006), the effect of the decision maker type on consumers' reactions may differ for consumers from a minority versus majority group. Future research can incorporate consumer demographics to understand such differences. Fourth, our work focuses on consumers' attitudes toward the company, but more research is needed to understand how algorithmic decisionmaking impacts consumers' psychological security. For instance, future research can investigate how the decision maker type affects consumers' perceived threat and anxiety (Mende et al. 2019). Fifth, future research could investigate whether consumers' reactions change depending on whether the decision outcome is communicated by a person or through a nonhuman medium (e.g., email). Although we manipulated only the decision maker type and held all other communication about the decision outcome constant, future research could examine the effect of how decisions are communicated to consumers (e.g., Campbell 2007; price tag vs. store owner). Sixth, although the current research primarily focuses on consumers' attitudes toward the company, which is a managerially important consumer indicator, future research can extend our findings to other behavioral measures.

Lastly, it is interesting to consider under which conditions algorithmic (vs. human) decisions might be more likely to facilitate internal attributions. Although we observed a consistent pattern across different consumer contexts, responses, and procedures, it is possible that in some situations algorithmic acceptance might offer an especially salient cue of diagnosticity and facilitate internal attributions to a larger extent. In general, more research is needed to understand how our effects can be moderated by the nature of the evaluation context. For instance, if a decision process is based on a simple objective criterion (e.g., if one's grade point average is above the 80th percentile), a favorable decision might facilitate internal attributions regardless of whether the decision maker is an algorithm or a human, mitigating the effect of the decision maker type.

Managerial Implications

The current work has several managerial implications. First, our results offer insights—perhaps surprising to many managers—into how the adoption of algorithms for consumer-facing decisions may affect consumers' reactions toward the company. We found that some managers hesitate to automate consumer-facing decisions because they are concerned about exacerbating

consumers' negative reactions to unfavorable decision outcomes (see in-depth interviews #2, #3, and #12 in Web Appendix C; Dietvorst, Simmons, and Massey 2015; Luo et al. 2019). Our results, however, demonstrate that an algorithmic (vs. a human) decision maker hurts consumers' reactions for favorable outcomes, not for unfavorable ones.

Second, in our interviews with managers, some managers expected that consumers would respond more positively when human and algorithmic decision makers collaborate, and some mentioned that their companies are already using this strategy (see in-depth interviews #2 and #9 in Web Appendix C). Our results indicate that consumers may not necessarily respond more positively to companies if humans are merely observing the algorithms without active involvement in decision making (Study 7). By showing this, we offer managerial insights on how companies can design their evaluation processes. In addition, we demonstrate that the effect of the decision maker type is mitigated when the favorable decision outcome is not self-diagnostic (i.e., when the decision was based on a raffle; Study 5). Managers can leverage these findings to improve consumers' reactions to companies that use algorithms for consumer-facing decisions.

Third, we explored a possible approach to mitigate the risk of less positive reactions following algorithmic acceptance: making the algorithm more human-like. In Study 8, the addition of simple anthropomorphic cues eliminated the effect of the decision maker type in the case of an acceptance decision. We also observed a similar pattern in field data from a financial services company. These data provide click-through rates on a link to the company's services after receiving financial feedback from human-like algorithms (vs. non-human-like algorithms; for details, see Web Appendix M). Once consumers answered a questionnaire, the company provided feedback based on an algorithmic assessment of the consumer's financial health. Some consumers received feedback that was highly favorable (good financial health with just a check-up needed), mimicking the favorable outcome condition of Study 8. Replicating the effect with a behavioral measure, these consumers were more likely to seek information about the company's services when the favorable feedback came from a human-like (vs. non-human-like) algorithm. These preliminary findings mimic those in Study 8 and corroborate the conclusion that negative consequences of algorithmic decision making may be averted by making algorithms more human-like (using, e.g., a more conversational format, a human name, a human-like photo).

Finally, we offer insights for policy makers. When the decision maker type is not disclosed, consumers are likely to react similarly as they do to a human decision maker (Studies 3a–b), offering firms an incentive to avoid transparency, which is not in the interest of consumers. Our results align with recent movements calling practitioners to be more transparent about their use of algorithms (Davenport et al. 2020; Rai 2020) and laws in the United States and European Union that require companies to disclose whether they use algorithms in consumer-related tasks (Castelluccia and Le Métayer 2019; Smith 2020).

Associate Editor

Dhruv Grewal

Contribution Note

The first two authors contributed equally.


Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This is funded by the Erasmus Research Institute of Management

ORCID iD

Gizem Yalcin  <https://orcid.org/0000-0001-7671-4892>

References

- Argo, Jennifer J., Darren W. Dahl, and Rajesh V. Manchanda (2005), "The Influence of a Mere Social Presence in a Retail Context," *Journal of Consumer Research*, 32 (2), 207–12.
- Barry, Christopher T., William F. Chaplin, and Sarah J. Grafeman (2006), "Aggression Following Performance Feedback: The Influences of Narcissism, Feedback Valence, and Comparative Standard," *Personality and Individual Differences*, 41 (1), 177–87.
- Baumeister, Roy F. (1999), *The Self in Social Psychology*. Philadelphia: Psychology Press.
- Bigman, Yochanan E. and Kurt Gray (2018), "People Are Averse to Machines Making Moral Decisions," *Cognition*, 181, 21–34.
- Bolukbasi, Tolga, Kai-Wei Chang, James Y. Zou, Venkatesh Saligrama, and Adam T. Kalai (2016), "Man Is to Computer Programmer as Woman Is to Homemaker? Debiasing Word Embeddings," in *Advances in Neural Information Processing Systems*, arXiv preprint, <https://arxiv.org/abs/1607.06520>.
- Bonezzi, Andrea and Mario Ostinelli (2021), "Can Algorithms Legitimize Discrimination?" *Journal of Experimental Psychology: Applied*, 27 (2), 447–59.
- Botti, Simona and Sheena S. Iyengar (2006), "The Dark Side of Choice: When Choice Impairs Social Welfare," *Journal of Public Policy & Marketing*, 25 (1), 24–38.
- Brewer, Marylinn B. (1991), "The Social Self: On Being the Same and Different at the Same Time," *Personality and Social Psychology Bulletin*, 17 (5), 475–82.
- Cadario, Romain, Chiara Longoni, and Carey K. Morewedge (2021), "Understanding, Explaining, and Utilizing Medical Artificial Intelligence," *Nature Human Behavior*, 5, 1636–42.
- Campbell, Margaret C. (2007), "'Says Who?!' How the Source of Price Information and Affect Influence Perceived Price (Un)Fairness," *Journal of Marketing Research*, 44 (2), 261–71.
- Castelluccia, Claude and Daniel Le Métayer (2019), "Understanding Algorithmic Decision-Making: Opportunities and Challenges," *European Parliament*, [https://www.europarl.europa.eu/thinktank/en/document/EPRS_STU\(2019\)624261](https://www.europarl.europa.eu/thinktank/en/document/EPRS_STU(2019)624261).
- Castelo, Noah, Maarten W. Bos, and Don R. Lehmann (2019), "Task-Dependent Algorithm Aversion," *Journal of Marketing Research*, 56 (5), 809–25.
- Crolic, Cammy, Felipe Thomaz, Rhonda Hadi, and Andrew T. Stephen (2022), "Blame the Bot: Anthropomorphism and Anger in Customer-Chatbot Interactions," *Journal of Marketing*, 86 (1), 132–48.
- Davenport, Thomas, Abhijit Guha, Dhruv Grewal, and Timna Bressgott (2020), "How Artificial Intelligence Will Change the Future of Marketing," *Journal of the Academy of Marketing Science*, 48 (1), 24–42.
- Diab, Dalia L., Shuang-Yueh Pui, Maya Yankelevich, and Scott Highhouse (2011), "Lay Perceptions of Selection Decision Aids in US and Non-US Samples," *International Journal of Selection and Assessment*, 19 (2), 209–16.
- Dietvorst, Berkeley J. and Soaham Bharti (2020), "People Reject Algorithms in Uncertain Decision Domains Because They Have Diminishing Sensitivity to Forecasting Error," *Psychological Science*, 31 (10), 1302–14.
- Dietvorst, Berkeley J., Joseph P. Simmons, and Cade Massey (2015), "Algorithm Aversion: People Erroneously Avoid Algorithms After Seeing Them Err," *Journal of Experimental Psychology: General*, 144 (1), 114–26.
- Dietvorst, Berkeley J., Joseph P. Simmons, and Cade Massey (2016), "Overcoming Algorithm Aversion: People Will Use Imperfect Algorithms if They Can (Even Slightly) Modify Them," *Management Science*, 64 (3), 1155–70.
- Dunn, Lea and Darren W. Dahl (2012), "Self-Threat and Product Failure: How Internal Attributions of Blame Affect Consumer Complaining Behavior," *Journal of Marketing Research*, 49 (5), 670–81.
- Eastwood, Joseph, Brent Snook, and Kirk Luther (2012), "What People Want from Their Professionals: Attitudes Toward Decision-Making Strategies," *Behavioral Decision Making*, 25 (5), 458–68.
- Efendić, Emir, Philippe P.F.M. Van de Calseyde, and Anthony M. Evans (2020), "Slow Response Times Undermine Trust in Algorithmic (But Not Human) Predictions," *Organizational Behavior and Human Decision Processes*, 157, 103–14.
- Epley, Nicholas (2018), "A Mind Like Mine: The Exceptionally Ordinary Underpinnings of Anthropomorphism," *Journal of the Association for Consumer Research*, 3 (4), 591–98.
- Folkes, Valerie S. (1984), "Consumer Reactions to Product Failure: An Attributional Approach," *Journal of Consumer Research*, 10 (4), 398–409.
- Forsyth, Donelson R. (1980), "The Functions of Attributions," *Social Psychology Quarterly*, 43 (2), 184–89.
- Fromkin, Howard L. and Charles R. Snyder (1980), "The Search for Uniqueness and Valuation of Scarcity," in *Social Exchange: Advances in Theory and Research*, Kenneth J. Gergen, Martin S. Greenberg, and Richard H. Willis, eds. Boston: Springer, 57–75.
- Hayes, Andrew F. (2013), *Introduction to Mediation, Moderation, and Conditional Process Analysis: A Regression-Based Approach*. New York: The Guilford Press.
- He, Stephen X. and Samuel D. Bond (2015), "Why Is the Crowd Divided? Attribution for Dispersion in Online Word of Mouth," *Journal of Consumer Research*, 41 (6), 1509–27.

- Heider, Fritz (1958), "The Naive Analysis of Action," in *The Psychology of Interpersonal Relations*. Hoboken, NJ: John Wiley & Sons, 79–124.
- Jago, Arthur (2019), "Algorithms and Authenticity," *Academy of Management Discoveries*, 5 (1), 38–56.
- Jones, Edward E. and Keith E. Davis (1965), "From Acts to Dispositions the Attribution Process in Person Perception," in *Advances in Experimental Social Psychology*, Vol. 2, Leonard Berkowitz, ed. Cambridge, MA: Academic Press, 219–66.
- Kelley, Harold H. (1967), "Attribution Theory in Social Psychology," *Nebraska Symposium on Motivation*, 15, 192–238.
- Kelley, Harold H. (1973), "The Processes of Causal Attribution," *American Psychologist*, 28 (2), 1078–2.
- Kelley, Harold H. and John L. Michela (1980), "Attribution Theory and Research," *Annual Review of Psychology*, 31 (1), 457–501.
- Kim, Tae Woo and Adam Duhachek (2020), "Artificial Intelligence and Persuasion: A Construal-Level Account," *Psychological Science*, 31 (4), 363–80.
- Kim, Hye-Young and Ann L. McGill (2018), "Minions for the Rich? Financial Status Changes How Consumers See Products with Anthropomorphic Features," *Journal of Consumer Research*, 45 (2), 429–50.
- Lee, Min Kyung (2018), "Understanding Perception of Algorithmic Decisions: Fairness, Trust, and Emotion in Response to Algorithmic Management," *Big Data & Society*, 5 (1), 1–16.
- Leung, Eugina, Gabriele Paolacci, and Stefano Puntoni (2018), "Man Versus Machine: Resisting Automation in Identity-Based Consumer Behavior," *Journal of Marketing Research*, 55 (6), 818–31.
- Logg, Jennifer M., Julia A. Minson, and Don A. Moore (2019), "Algorithm Appreciation: People Prefer Algorithmic to Human Judgment," *Organizational Behavior and Human Decision Processes*, 151, 90–103.
- Longoni, Chiara, Andrea Bonezzi, and Carey Morewedge (2019), "Resistance to Medical Artificial Intelligence," *Journal of Consumer Research*, 46 (4), 629–50.
- Longoni, Chiara and Luca Cian (2022), "Artificial Intelligence in Utilitarian vs. Hedonic Contexts: The 'Word-of-Machine' Effect," *Journal of Marketing*, 86 (1), 91–108.
- Luginbuhl, James E., Donald H. Crowe, and James P. Kahan (1975), "Causal Attributions for Success and Failure," *Journal of Personality and Social Psychology*, 31 (1), 86–93.
- Luo, Xueming, Siliang Tong, Zheng Fang, and Zhe Qu (2019), "Frontiers: Machines vs. Humans: The Impact of Artificial Intelligence Chatbot Disclosure on Customer Purchases," *Marketing Science*, 38 (6), 937–47.
- Markets and Markets (2021), "Artificial Intelligence (AI) Market Research Report," press release, <https://www.marketsandmarkets.com/PressReleases/artificial-intelligence.asp%20.asp>.
- McFarland, Cathy and Michael Ross (1982), "Impact of Causal Attributions on Affective Reactions to Success and Failure," *Journal of Personality and Social Psychology*, 43 (5), 937–46.
- McFerran, Brendan and Jennifer J. Argo (2014), "The Entourage Effect," *Journal of Consumer Research*, 40 (5), 871–84.
- Mende, Martin, Maura L. Scott, Jenny van Doorn, Dhruw Grewal, and Ilana Shanks (2019), "Service Robots Rising: How Humanoid Robots Influence Service Experiences and Elicit Compensatory Consumer Responses," *Journal of Marketing Research*, 56 (4), 535–56.
- Miller, Dale T. and Michael Ross (1975), "Self-Serving Biases in the Attribution of Causality: Fact or Fiction?" *Psychological Bulletin*, 82 (2), 213–25.
- Nash, Kyle, Johanna M. Lea, Thomas Davies, and Kumar Yogeeswaran (2018), "The Bionic Blues: Robot Rejection Lowers Self-Esteem," *Computers in Human Behavior*, 78, 59–63.
- Newman, David T., Nathanael J. Fast, and Derek J. Harmon (2020), "When Eliminating Bias Isn't Fair: Algorithmic Reductionism and Procedural Justice in Human Resource Decisions," *Organizational Behavior and Human Decision Processes*, 160, 149–67.
- Okten, Irmak Olcaysoy and Gordon B. Moskowitz (2018), "Goal Versus Trait Explanations: Causal Attributions Beyond the Trait-Situation Dichotomy," *Journal of Personality and Social Psychology*, 114 (2), 211–29.
- O'Laughlin, Matthew J. and Bertram F. Malle (2002), "How People Explain Actions Performed by Groups and Individuals," *Journal of Personality and Social Psychology*, 82 (1), 33–48.
- Onkal, Dilek, Paul Goodwin, Mary Thomson, Sinan Gonul, and Andrew Pollock (2009), "The Relative Influence of Advice from Human Experts and Statistical Methods on Forecast Adjustments," *Journal of Behavioral Decision Making*, 22 (4), 390–409.
- Oppenheimer, Daniel M., Tom Meyvis, and Nicolas Davidenko (2009), "Instructional Manipulation Checks: Detecting Satisficing to Increase Statistical Power," *Journal of Experimental Social Psychology*, 45 (4), 867–72.
- Park, C. Whan, Deborah J. MacInnis, Joseph Priester, Andreas B. Eisingerich, and Dawn Iacobucci (2010), "Brand Attachment and Brand Attitude Strength: Conceptual and Empirical Differentiation of Two Critical Brand Equity Drivers," *Journal of Marketing*, 74 (6), 1–17.
- Pieters, Rik (2017), "Meaningful Mediation Analysis: Plausible Causal Inference and Informative Communication," *Journal of Consumer Research*, 44 (3), 692–716.
- Promberger, Marianne and Jonathan Baron (2006), "Do Patients Trust Computers?" *Journal of Behavioral Decision Making*, 19 (5), 455–68.
- Puntoni, Stefano, Rebecca Walker Rezek, Markus Giesler, and Simona Botti (2021), "Consumers and Artificial Intelligence: An Experiential Perspective," *Journal of Marketing*, 85 (1), 131–51.
- Rai, Arun (2020), "Explainable AI: From Black Box to Glass Box," *Journal of the Academy of Marketing Science*, 48 (1), 137–41.
- Rhodewalt, Frederick and James Davison Jr. (1986), "Self-Handicapping and Subsequent Performance: Role of Outcome Valence and Attributional Certainty," *Basic and Applied Social Psychology*, 7 (4), 307–22.
- Russell, Dan (1982), "The Causal Dimension Scale: A Measure of How Individuals Perceive Causes," *Journal of Personality and Social Psychology*, 42 (6), 1137–45.
- Shaffer, Victoria A., C. Adam Probst, Edgar C. Merkle, Hal R. Arkes, and Mitchell A. Medow (2013), "Why Do Patients Derogate Physicians Who Use a Computer-Based Diagnostic Support System?" *Medical Decision Making*, 33 (1), 108–18.
- Sheikh, Aziz (2006), "Why Are Ethnic Minorities Under-Represented in US Research Studies?" *PLoS Medicine*, 3 (2), 166–67.
- Smith, Andrew (2020), "Using Artificial Intelligence and Algorithms," Federal Trade Commission (April 8), <https://www.ftc.gov/news-events/blogs/business-blog/2020/04/using-artificial-intelligence-algorithms>.

- Srinivasan, Raji and Gülen Sarial-Abi (2021), "When Algorithms Fail: Consumers' Responses to Brand Harm Crises Caused by Algorithm Errors," *Journal of Marketing*, 85 (5), 74–91.
- Sundar, S. Shyam and Clifford Nass (2001), "Conceptualizing Sources in Online News," *Journal of Communication*, 51 (1), 52–72.
- Taylor, Shelley E. and Jonathon D. Brown (1988), "Illusion and Well-Being: A Social Psychological Perspective on Mental Health," *Psychological Bulletin*, 103 (2), 193–210.
- Trope, Yaacov and Efrat Neter (1994), "Reconciling Competing Motives in Self-Evaluation: The Role of Self-Control in Feedback Seeking," *Journal of Personality and Social Psychology*, 66 (4), 646–57.
- Wan, Lisa C. and Robert S. Wyer Jr. (2019), "The Influence of Incidental Similarity on Observers' Causal Attributions and Reactions to a Service Failure," *Journal of Consumer Research*, 45 (6), 1350–68.
- Williams, Kipling D. (2007), "Ostracism," *Annual Review of Psychology*, 58, 425–52.
- Yeomans, Michael, Anuj Shah, Sendhil Mullainathan, and Jon Kleinberg (2019), "Making Sense of Recommendations," *Journal of Behavioral Decision Making*, 32 (4), 403–14.
- Yoon, Song-Oh and Itamar Simonson (2008), "Choice Set Configuration as a Determinant of Preference Attribution and Strength," *Journal of Consumer Research*, 35 (2), 324–36.
- Zuckerman, Miron (1979), "Attribution of Success and Failure Revisited, or: The Motivational Bias Is Alive and Well in Attribution Theory," *Journal of Personality*, 47 (2), 245–87.