Received: 22 December 2021 | Revised: 18 June 2023 | Accepted: 26 September 2023

DOI: 10.1111/isj.12482

RESEARCH ARTICLE

WILEY

Perceived algorithmic fairness: An empirical study of transparency and anthropomorphism in algorithmic recruiting

Jessica Ochmann[1] | Leonard Michels[1] | Verena Tiefenbeck[1] | Christian Maier[2] | Sven Laumer[1]

[1]School of Business, Economics and Society, Institute of Information Systems, Friedrich-Alexander-Universität Erlangen-Nürnberg, Nuremberg, Germany

[2]Information Systems, Health and Society in the Digital Age, University of Bamberg, Bamberg, Germany

Correspondence
Sven Laumer, School of Business, Economics and Society, Institute of Information Systems, Friedrich-Alexander-Universität Erlangen-Nürnberg, Nuremberg, Germany.
Email: sven.laumer@fau.de

Abstract

Despite constant efforts of organisations to ensure a fair and transparent personnel selection process, hiring is still characterised by systematic inequality. The potential of algorithms to produce fair and objective decision outcomes has attracted the attention of academic scholars and practitioners as a conceivable alternative to human decision-making. However, applicants do not necessarily consider an objective algorithm as fairer than a human decision maker. This study examines the conditions under which applicants perceive algorithms as fair and establishes a theoretical foundation of algorithmic fairness perceptions. We further propose and investigate transparency and anthropomorphism interventions as strategies to actively shape these fairness perceptions. In an online application scenario with eight experimental groups ($N = 801$), we analyse determinants for algorithmic fairness perceptions and the impact of the proposed interventions. Embedded in a stimulus-organism-response framework and drawing from organisational justice theory, our study reveals four justice dimensions (procedural, distributive, interpersonal, informational justice) that determine algorithmic fairness perceptions. The results further show that transparency and anthropomorphism interventions mainly affect dimensions of interpersonal

and informational justice, highlighting the importance of algorithmic fairness perceptions as critical determinants for individual choices.

KEYWORDS
anthropomorphism, organisational justice, perceived algorithmic fairness, personnel selection, stimulus-organism-response, transparency

1 | INTRODUCTION

Organisations and their human resources (HR) departments aim to establish fair and transparent personnel selection processes (Gilliland, 1993). To that end, they employ various methods and strategies, such as offering diversity and inclusion training, sensitising employees to fair personnel selection processes, and applying standardised application criteria. Despite these efforts, applicants often have the impression that fairness continues to be wishful thinking in the personnel selection process (Thielsch et al., 2012). For example, some groups of applicants are significantly disadvantaged due to their demographic characteristics (Avery et al., 2008). A recent study from the UK illustrates that ethnic minorities have to submit 60% more applications to be considered in the personnel selection process than the majority population when they disclose their ethnicity (Di Stasio & Heath, 2020). Similar effects occur when considering how gender influences the success of equally qualified applicants (Bertrand & Mullainathan, 2004; Moss-Racusin et al., 2012). Recent statistics indicate that the share of women at the board level of the largest listed companies in the EU amounts to 30.6% (Eurostat, 2022). The inconvenient truth is that such systematic run through multiple societies and industries (Annabi & Lebovitz, 2018; Cheryan & Markus, 2020; Dworkin et al., 2020; Quillian et al., 2017). The impression of this consistent persistence of discriminatory tendencies in personnel selection is also confirmed by 79% of HR managers (Souza, 2020).

One of the most promising paths to increase fairness in personnel selection processes is using sophisticated algorithms[1] (Newlands, 2021; van den Broek et al., 2021). If algorithms are carefully designed and based on unbiased data, there is a potential to produce fair decisions (De-Arteaga et al., 2022; Dolata et al., 2022; Robert et al., 2020). Nonetheless, there are indications that applicants would not necessarily consider an objective algorithm as fair (Lee, 2018).

The discrepancy between objective fairness criteria and perceived fairness is mainly rooted in the multidimensional nature of fairness. Fairness is often context-specific and formed as a subjective, personal evaluation (Colquitt, 2001). For algorithmic decisions, fairness perceptions can be detached from an algorithm's actual performance (Köchling & Wehner, 2020), and even if an algorithm were to apply objectively fair criteria, applicants might have contrary perceptions. Therefore, it is an opportune and critical time to understand which factors make applicants perceive algorithms to be fair and to build knowledge on their mental models of perceived algorithmic fairness (Lee, 2018).

To better leverage the advantages of potentially fair algorithms, it is also of value to identify potential interventions that influence applicants' algorithmic fairness perceptions. Prior research has shown that organisations can shape applicants' perceptions of fair selection procedures (Cropanzano et al., 2007) by providing causal explanations of the personnel selection process (Ployhart & Ryan, 1997) or ensuring interpersonal treatment (e.g., being empathetic and understanding) (Cropanzano et al., 2007). As individuals often consider algorithms to be reductionistic and opaque (Newman et al., 2020), we need to find ways to transfer these conventions and rules to algorithmic recruiting.

[1]In line with prior literature, we refer to algorithms as encoded procedures that encompass statistical models based on evidence-based rules and specific mathematical calculations to perform an action without human involvement.

Recent related literature has identified two main strategies for shaping perceptions of algorithms: (1) establishing transparency by providing information about an algorithm's underlying processes (Felzmann et al., 2019) and (2) using anthropomorphism to increase an individual's familiarity with the algorithmic decision maker (Schanke et al., 2021). So far, research has predominantly investigated the implications of *transparency* and *anthropomorphism* on adopting algorithms in decision-making. These research streams are especially interested in increasing the acceptance of algorithms and overcoming applicants' algorithm aversion (Burton et al., 2020) and little is known about how the proposed mechanisms are appropriate measures to shape algorithmic fairness perceptions (Binns et al., 2018; Schanke et al., 2021). We will complement those insights by evaluating *how* the proposed mechanisms influence applicants' algorithmic fairness perceptions (Binns et al., 2018; Schanke et al., 2021). This is important as fair treatment of applicants in the application process fosters their satisfaction with the organisation and word-of-mouth recommendations (McCarthy et al., 2017). Both outcomes help companies build a positive organisational reputation (Cropanzano et al., 2007).

To summarise, the objective of our research is twofold. First, *we aim to obtain a nuanced picture of perceived algorithmic fairness*. We gain an understanding of the emergence of perceived algorithmic fairness by relying primarily on current research activities related to the application of organisational justice theory. This theory describes fairness as a combined evaluation of distributive, procedural, interpersonal, and informational justice (Colquitt, 2001). However, these studies tend to focus on a particular fairness dimension, leading to an incomplete picture of perceived algorithmic fairness (Robert et al., 2020). Therefore, we take a holistic approach that considers all justice dimensions in one study and relies on extensive research on organisational justice. This more differentiated view of perceived algorithmic fairness enables us to pursue our second objective, which is *to analyse how transparency and anthropomorphism shape perceptions of algorithmic fairness*.

We conducted a randomised survey-based online experiment with 801 participants to achieve our two research objectives. Participants on the online sampling platform Prolific were invited to apply for well-compensated participation in future market studies of the company 'Skillistic Insights'. We created an application scenario to investigate algorithmic fairness perceptions in a real-world setting. Drawing on organisational justice theory (Colquitt, 2001) and embedding our hypotheses in a stimulus-organism-response network (Jacoby, 2002), we examined the underlying dimensions of perceived algorithmic fairness. Additionally, we investigated the effects of transparency and anthropomorphism as two interventions to increase algorithmic fairness perceptions in the personnel selection process. Lastly, we assessed the implications of algorithmic fairness perceptions for organisational outcomes (e.g., satisfaction and recommendation likelihood).

The theoretical account of our paper provides three key contributions to IS research. First, we systematically investigate how and when justice dimensions originally formulated in organisational justice theory can be applied to measure algorithmic fairness perceptions. Second, we show that transparency and anthropomorphism are effective interventions to influence algorithmic fairness perceptions. Third, we identify the implications of algorithmic fairness perceptions by investigating related individual responses and outline the implications for organisations.

The paper proceeds as follows: First, we present related work on algorithmic decision-making and fairness and discuss the development of perceived algorithmic fairness. We further embed our research into the stimulus-organism-response model. Departing from these insights, we derive our hypotheses and describe our empirical method used to investigate these questions and present the results. Finally, we discuss our findings and conclude with implications for academia and practice.

2 | THEORETICAL BACKGROUND

2.1 | Organisational justice theory and individual responses

In the history of humankind, fairness—often interchangeably used with justice—has been thought of as a key factor enabling collaboration among humans (Almås et al., 2010). Central to the definition of fairness is the unbiased and equal distribution of goods and rights according to the performance and needs of individuals (McAuliffe et al., 2017).

While researchers throughout the ages have been captivated by the evolutionary puzzle of the distribution of limited resources (McAuliffe et al., 2017), fairness has become increasingly visible in organisational science over the last decades (Colquitt, 2001; Cropanzano et al., 2007). Unlike early philosophical approaches that have been intrigued to define logical rules for an objectively just resource allocation (Miller, 1996), managerial scientists have been mainly interested in fairness perceptions. Under the umbrella of *organisational justice theory*, these scholars have described how individuals assess organisational decision outcomes (e.g., hiring decisions) and discussed the discrepancy between objective fairness characteristics (e.g., equal hiring chances for similarly skilled candidates) and fairness perceptions (e.g., an individual's subjective belief to be better qualified for a particular job than another candidate) (Colquitt & Zipay, 2015). In other words, they are interested in capturing when individuals believe a specific decision is fair, rather than ascertaining an objective reality (Cropanzano et al., 2007).

Following this view, fairness perceptions result from evaluating four dimensions: procedural justice, distributive justice, interpersonal justice, and informational justice. While *distributive justice* refers to the evaluation of the appropriateness of a decision outcome regarding allocation rules that were applied in the decision-making process to make a certain decision, *procedural justice* concerns the evaluation of the procedures used to determine the allocation of resources and, finally, the decision outcome (Colquitt et al., 2001). In contrast, interpersonal and informational justice, often summarised as *interactional justice*, focus on the interpersonal treatment associated with a decision outcome. In essence, *interpersonal justice* refers to the degree to which people that are affected by the decision outcome were treated with politeness, dignity, and respect (Bies & Shapiro, 1987), whereas *informational justice* puts emphasis on explaining the procedures used and the distribution of resources provided to people that are affected by the decision outcome (Colquitt, 2001).

Organisations seek to understand what sort of activities influences the four justice dimensions because there is a consensus among scientists that higher fairness perceptions can provide powerful benefits for organisations (Colquitt, 2001). Up to now, some studies have identified the positive relationship between organisational justice and various organisational outcomes, including job performance, organisational citizenship behaviours, trust, and commitment (Cohen-Charash & Spector, 2001; Colquitt et al., 2001). Even though a great deal of previous research into organisational justice has focused on fairness perceptions of employees, some studies have postulated the applicability of organisational justice theory to examine candidates' reactions to recruiting procedures (Gilliland, 1993; Hausknecht, 2004). Building on the early work of Smith and Robertson (1989) and Schuler (1993) who identified the characteristics of the selection method as well as the nature of the decision feedback as the key determinants for applicants' reactions to the recruiting process, this line of evidence suggested that the four justice dimensions are a valuable reflection of these two key determinants. Consequently, researchers investigating personnel selection from the organisational justice perspective argue that the four justice dimensions are an appropriate measure to provide a differentiated picture of the selection process (Gilliland, 1993; Ployhart & Ryan, 1997). Using the organisational justice lens, researchers have been able to show, for example, that candidates who feel fairly treated in the application process build a positive impression of the organisation (Bauer et al., 2001), have a higher motivation to perform in the job interview (Lukacik et al., 2022), and are more likely to refer the organisation to peers (Schinkel et al., 2016)—regardless of the actual decision outcome.

In IS research, organisational justice theory has been adopted to explain employees' intentions to abuse computers (Willison et al., 2018), the influence of fairness perceptions in strategic systems planning (Mirchandani & Lederer, 2014) and supply chain integration and firm performance (Wei et al., 2021), customer demands for privacy (Greenaway et al., 2015) and how to rebuild trust after customer information privacy violation (Bansal & Zahedi, 2015), compliance with internet use policy (Li et al., 2014), and general customer reactions to digital services (Carr, 2007). Findings in those streams agree that organisational justice theory can be used on IS-specific topics and offer valuable theoretical insights. More recently, some specific dimensions of organisational justice, mainly including procedural and distributive justice, were also applied to study the fairness of algorithmic decision makers (Dolata et al., 2022; Feuerriegel et al., 2020; Morse et al., 2021). So far, however, the impact of fairness in an algorithm-based recruiting context is still unclear, thus highlighting the need for further research activities (Binns et al., 2018).

2.2 | Perceived algorithmic fairness

In the past years, algorithms have been increasingly implemented to support human decision-making, which guided scholars to consider the meaning of organisational justice theory for algorithmic fairness (Robert et al., 2020). A number of studies have assessed the mathematical formalisation of algorithmic fairness. By developing fair algorithms, these studies have particularly taken account of procedural and distributive justice (De-Arteaga et al., 2022). To complete these efforts, more recent attention has focused on perceived algorithmic fairness (Dolata et al., 2022). This line of work is particularly interested in potential differences regarding the fairness perceptions of human and algorithmic decision makers (Lee, 2018; Newman et al., 2020) and has demonstrated that individuals form fairness perceptions detached from the performance of an algorithm (Köchling & Wehner, 2020).

However, these studies have not been able to provide robust evidence on how individuals evaluate algorithmic fairness. For example, previous research has established that algorithms' objective and rule-based nature can defeat human bias and lead to fairer recruiting decisions (Newman et al., 2020; van den Broek et al., 2021). Surprisingly, a key study comparing fairness perceptions of human and algorithmic decision makers found that individuals have higher fairness perceptions for humans compared to algorithms (Newman et al., 2020; Yalcin et al., 2022). In contrast, other studies have concluded that an algorithm's ability to generate optimised and fair decisions can lead to higher fairness perceptions for algorithmic compared to human decision makers (Bai et al., 2021; Fumagalli et al., 2022; Lee, 2018). Unlike studies revealing significant differences in the fairness perceptions of human and algorithmic decision makers, recent evidence reported no differences in the fairness perceptions of the decision maker (human vs. algorithm) (Oetting & Maier, 2018; Suen et al., 2019).

One explanation for these varying findings is the need for more consideration of fairness as a multidimensional construct. Perceptions of algorithmic fairness have been studied by many researchers focusing on one particular justice dimension. For example, a considerable amount of research strongly focused on distributive and procedural justice, whereas more attention should be paid to interactional justice (Robert et al., 2020). While these studies have provided a useful account of how to shape fairness perceptions in the particular justice dimensions, they ignore the complexity of fairness formation and that procedural, distributive, and interactional justice dimensions are indeed distinct but still interrelated (Robert et al., 2020). In addition, much uncertainty still exists regarding the applicability of organisational justice to the measurement of perceived algorithmic fairness. For example, Starke et al. (2021) criticised that organisational justice was originally developed for the fairness evaluation of human decision makers. The authors suggested that a natural progression of prior work is to validate how the factors proposed by organisational justice theory determine algorithmic fairness perceptions.

To summarise, no studies have been performed that have fully addressed the multidimensionality of the construct fairness in algorithmic decision-making (Robert et al., 2020). This is especially critical, as prior research has emphasised that all four justice dimensions interact (Cropanzano et al., 2007). Hence, to explain a high degree of variance, researchers should include multiple dimensions of justice within one study and investigate them in relation to each other (Colquitt et al., 2001). To complement these research endeavours it is also required to validate the suitability of organisational justice theory for the explanation of algorithmic fairness perceptions and assess if there are any differences between human and algorithmic decision makers.

2.3 | Transparency and anthropomorphism in algorithmic decision-making

For many years, one of the key characteristics of highly sophisticated algorithms in decision-making was their opaqueness or 'black box' nature, referring to an incomprehensibility of how their decision was reached (Adadi & Berrada, 2018; Barredo Arrieta et al., 2020; Castelvecchi, 2016). Multiple scholars have identified the resulting inability of humans to understand the processes underlying an algorithm's output as one of the key obstacles hindering

their acceptance (Ajunwa, 2020; Cadario et al., 2021; Yeomans et al., 2019; Zhu et al., 2018). Recent advances in the field of explainable AI (XAI) and machine learning, however, have led to the development of techniques that produce models and algorithms that are easier to interpret and understand by humans (Barredo Arrieta et al., 2020; Felzmann et al., 2019).

These advancements enable developers and organisations to provide individuals relying on algorithmic decision makers and individuals affected by them with additional, transparent information about their underlying reasoning. Revealing this information helps humans to understand the factors that influenced the algorithm's decision and the mechanisms of how they were integrated in the decision-making process (Burton et al., 2020; Gönül et al., 2009). Consequently, increasing an algorithm's transparency has been shown to increase its acceptance, provided this information is comprehensible (Cadario et al., 2021; Castelo et al., 2019; Glikson & Woolley, 2020). In addition, Mahmud et al. (2022) suggest that increasing an algorithmic decision maker's transparency can increase its trustworthiness.

Prior research has, thereby, also suggested that increasing the transparency of an algorithm and explaining its underlying processes can be a viable intervention to increase its fairness and, thus, potentially people's fairness perceptions of algorithmic decision makers (Barredo Arrieta et al., 2020; Doshi-Velez & Kim, 2017; Sonboli et al., 2021; Wang et al., 2019).

While transparency in algorithmic decision-making is mainly concerned with the amount and content of the information provided about an algorithmic decision maker, another important aspect is how the decision maker and their decisions are presented and communicated.

When interpreting information in communication scenarios, humans commonly focus on objective criteria such as the information value and take non-verbal and social cues in the transmission of the information into account. Even though computers are non-human agents, prior research in the domain of computer-mediated communication has shown that humans are also looking for such cues in their interaction with computers (Liew & Tan, 2021; Nass & Moon, 2000; Seeger et al., 2021) and can greatly affect people's perceptions of the underlying technology (Liew & Tan, 2021; Qiu & Benbasat, 2009; Seeger et al., 2021). Consequently, scholars put great effort into identifying potential ways to induce a perception of computers as mindful, social, and even human-like.

One particularly promising approach in this domain is the concept of anthropomorphism (Blut et al., 2021; Epley et al., 2007; Waytz et al., 2010), which refers to the attribution of human characteristics, motivations, intentions, or emotions to non-human agents (e.g., robots, chatbots, algorithms). Prior research has shown that equipping non-human agents with anthropomorphic cues (e.g., visual, auditory, or mental) leads humans to perceive them as more human-like and, consequently, ascribe human-like abilities to them (Crolic et al., 2022; Waytz et al., 2010; Yalcin et al., 2022). This can increase people's acceptance of new technologies (Benlian et al., 2019; Maedche et al., 2019; Pfeuffer et al., 2019), trust in the non-human agent (de Visser et al., 2016), and even render the perception of interactions with the agent as more meaningful and important (Seeger et al., 2021).

These aspects can also have important consequences on people's fairness perceptions. Newman et al. (2020), for example, suggested that algorithms are generally perceived as less fair than humans. Increasing the human-likeness of an algorithm might increase fairness perceptions of an algorithmic decision maker. In addition, research in marketing has shown that brand anthropomorphization can affect people's fairness and justice perceptions (Kwak et al., 2017).

3 | THEORETICAL MODEL AND HYPOTHESES DEVELOPMENT

3.1 | Stimulus-organism-response theory

We drew on stimulus-organism response (S-O-R) theory to build a nomological network for algorithmic fairness perceptions. S-O-R has three stages and claims that (environmental) stimuli influence the internal psychological states of individuals from which individuals' responses are derived (Mehrabian & Russell, 1974). According to S-O-R, stimuli

are defined as external cues that arouse or excite individuals (Belk, 1975). Organism refers to cognitive and affective processes (e.g., individuals' perception, feelings, and beliefs) that mediate between the stimuli and the reactions of an individual. Those reactions may manifest themselves in different responses including the formation of impressions (i.e., internal non-visible responses) or performance of concrete actions (i.e., external visible responses) (Jacoby, 2002). In this study, we point to two widely discussed interventions, namely transparency and anthropomorphism as environmental stimuli (Felzmann et al., 2019; Schanke et al., 2021). Both interventions can be directly influenced by the organisation and act as external cues that arouse and excite candidates in the personnel selection process. We suggest that the proposed interventions may influence the perceived algorithmic fairness of individuals (Kordzadeh & Ghasemaghaei, 2022), which lead to individuals' responses such as satisfaction or referral likelihood (Cropanzano et al., 2007). Based on these theoretical relationships, we develop our hypotheses and elaborate on the specific associations among interventions, perceived algorithmic fairness, and individual responses. The research model depicting these relationships is illustrated in Figure 1.

3.2 | Embedding perceived algorithmic fairness in organisational justice theory

Taking into account the contradictory findings on the fairness perceptions of human and algorithmic decision makers, we argue to simultaneously investigate justice dimensions to expand our knowledge of algorithmic fairness perceptions. We, thereby, propose that much of the points elaborated by organisational justice theory do not only reflect the requirements for human decision makers but also the demands formulated for algorithms. Consequently, the holistic consideration of all four justice dimensions should provide a more nuanced view of algorithmic fairness perceptions and enable us to identify promising ways to actively shape them.

For example, prior research has recognised the demand for consistent, unbiased, correctable, representative, accurate, and ethical algorithms in order to increase algorithmic fairness perceptions (Feuerriegel et al., 2020). These requirements are mirrored by the procedural justice dimension that has linked a fair allocation process to multiple criteria: consistent application of procedures across people and time, adherence to unbiased, accurate, and correctable judgements and ethical standards, and the reflection of basic concerns and values of all individuals affected by the decision outcome (Colquitt, 2001; Leventhal, 1980). Scholars have also advocated that individuals have higher fairness perceptions of algorithms if the algorithms' underlying allocation rules are in line with leading social norms including equality, equity, and individual needs (Saxena et al., 2020). The requirements for these fair allocation rules closely align with distributive justice (Kordzadeh & Ghasemaghaei, 2022). Beyond the denotation of fair decision outcomes and the processes that lead to those outcomes, individuals expect that decision makers are transparent and behave friendly and attentive (McCarthy et al., 2017). This fundamental human need for respectful interpersonal

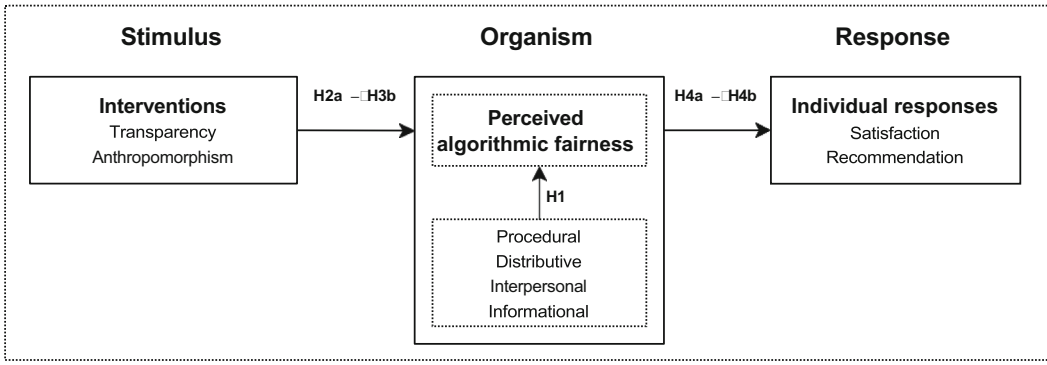

FIGURE 1 Research model and hypotheses.

treatment is emphasised by the organisational justice dimensions of interpersonal and informational justice. As it is well established from a variety of studies that individuals apply the same social rules in the interaction with technologies as in the interaction with humans (Nass & Moon, 2000), the impression of polite algorithms that treat individuals with respect and share relevant information should result in higher fairness perceptions (Acikgoz et al., 2020). Hence, we suggest applying the firm ideas of interpersonal and informational justice to shed further light on algorithmic fairness perceptions.

To summarise, organisational justice theory provides a promising foundation to build a differentiated picture of algorithmic fairness perceptions. Even though the four justice dimensions were originally developed to explain fairness perceptions of human decision makers, individuals seem to have the same requirements for algorithms. Therefore, they should be examined in their entirety in order to gain a comprehensive picture of algorithmic fairness. This is especially crucial against the backdrop of recent contradictory findings and the increasing scholarly attempt to understand fairness perceptions in algorithmic decision-making. Therefore, we hypothesize:

> **H1.** Algorithmic fairness perceptions consist of procedural, distributive, interpersonal, and informational justice perceptions.

While developing a holistic understanding of algorithmic fairness perceptions is an important research end itself, it is also a key prerequisite to determining and investigating ways to shape an individual's fairness perceptions. This is especially important in light of the results of prior research suggesting that people might unfoundedly devalue algorithms in terms of fairness.

So far, we have developed a differentiated understanding of how to shape fairness perceptions of human decision makers. Algorithms, however, differ from humans in their abstract and technical nature and require the application of adapted and tailored interventions to promote fairness perceptions. For example, we know from research on human decision makers that individuals evaluate personnel selection processes as fairer (even when they are rejected for a position) when the recruiter communicated the recruiting process transparently and was friendly (Cropanzano et al., 2007). Building on organisational justice theory, we propose that this knowledge can be adapted to algorithmic decision makers and guide the implementation of interventions to promote perceived algorithmic fairness.

3.3 | Interventions and perceived algorithmic fairness

The implementation of interventions can stimulate individuals to adopt new technology by, for example, transparently communicating an algorithm's reasoning (Cadario et al., 2021) or using anthropomorphic design elements (Seeger et al., 2021). These approaches have led to promising results in the domain of algorithm adoption and stimulated a growing body of literature concerned with the applicability of transparency and anthropomorphism interventions to influence fairness perceptions of algorithms (Newman et al., 2020; Schanke et al., 2021).

3.3.1 | Transparency

As outlined before, algorithmic fairness perceptions might be a crucial factor in determining individual responses resulting from an algorithmic decision-making process. It is, therefore, in an organisation's best interest to maximise individuals' algorithmic fairness perceptions. An essential step to achieve this objective is the development and adjustment of algorithmic models in a way that they recommend fair decisions or allocate resources in an unbiased and equitable manner (Dressel & Farid, 2018; Obermeyer et al., 2019). However, laying such technical foundations for a fair algorithmic decision does not necessarily translate to individuals' fairness perceptions of the decision's outcome. To strengthen this relationship, individuals need to understand and comprehend the processes that lead to

the outcome of an algorithm (Shin & Park, 2019). This notion is further supported by prior research that obtained empirical evidence for the importance of transparency for algorithmic fairness perceptions (Binns et al., 2018; Dodge et al., 2019; Kizilcec, 2016).

In the context of our theoretical model, additional information on an algorithm's underlying processes should influence procedural, distributive, and informational justice dimensions (Starke et al., 2021). Procedural justice denotes internal algorithmic decision-making processes such as the rules and logic incorporated into algorithmic models. Algorithmic decision processes are often perceived to be opaque because individuals cannot explain them. In contrast, individuals often rate human decision making to be more transparent as they assume other people apply the same decision heuristics as oneself. Giving individuals objective knowledge of how an algorithm comes up with a decision should balance this effect and enable individuals to believe that they have access to algorithmic decision making processes (Cadario et al., 2021). Providing information on the procedures of how the algorithm derived the decision should, therefore, result in higher levels of procedural justice. Similarly, distributive justice indicates whether individuals perceive an algorithm and its decision outcome as fair. In situations where individuals expect a standard to exist, they are more receptive to a provided reference point. Additional information about the underlying processes of an algorithm should induce individuals to form arguments that are consistent with the given reference point. Providing transparent information that helps individuals understand a certain decision outcome should therefore lead individuals to integrate that information into their evaluation and adjust their fairness judgements based on that information (Bauer & Gill, 2023). In the same vein, informational justice refers to whether individuals receive adequate justifications for a decision outcome. Prior research has shown that individuals often use algorithmic advice to confirm or disconfirm their initial assessment of a decision outcome (Jussupow et al., 2021). Explaining the reasoning for how the algorithm derived a decision outcome can thus result in the feeling that the algorithm made its decision based on a thoughtful and transparent process, which should increase individuals' ratings of informational justice.

Based on the presented results, we derive the following three hypotheses:

H2a. Transparently communicating the algorithm's underlying reasoning positively influences procedural justice.

H2b. Transparently communicating the algorithm's underlying reasoning positively influences distributive justice.

H2c. Transparently communicating the algorithm's underlying reasoning positively influences informational justice.

Next to the amount and content of information that is communicated about an algorithmic decision maker, organisations can also implement interventions that adjust the way the information or the decision maker itself is presented. One potential intervention in this domain builds on the concept of anthropomorphism.

3.3.2 | Anthropomorphism

Individuals often criticise the aloofness or 'strangeness' of algorithmic compared to human decision makers (Binns et al., 2018). Algorithms are, in nature, abstract and, as a result, often impersonal and inaccessible for individuals (Dietvorst et al., 2015). The general impression that algorithms are reductionistic and unable to take individual characteristics into account contributes to this effect (Newman et al., 2020). For example, previous studies found that applicants respond negatively to impersonal selection procedures (Lee, 2018) and criticise the poor interactional treatment in asynchronous digital interviews (Langer et al., 2017). Conducting face-to-face interviews instead of virtual interviews, on the other hand, led to higher expressed fairness perceptions (Chapman et al., 2003).

As the quality of interactional treatment is decisive for the perception of fairness (Colquitt & Zipay, 2015), research has introduced the concept of anthropomorphism, which describes the attribution of human qualities to non-human agents, as a promising lever to induce a sense of personal interaction (Pfeuffer et al., 2019; Seeger et al., 2021). Combining algorithmic decision-making with anthropomorphisation will help to increase the human-likeness of algorithms, which is likely to be reflected in perceptions of a higher quality interactional treatment. In addition, it helps to overcome an algorithm's reductionistic attribution, which might positively influence the corresponding fairness perceptions.

To map factors associated with interactional treatment to fairness perceptions, organisational justice theory proposes the dimensions of interpersonal and informational justice (Cohen-Charash & Spector, 2001).

More precisely, fair interactional treatment requires the decision maker to set up a respectful relationship (i.e., interpersonal justice) and give the individual the feeling of being truthful (i.e., informational justice). While friendly and empathetic communication can conceptualise a respectful relationship, candid and timely communication supports the notion of truthfulness (Cropanzano et al., 2007). Since individuals have similar expectations for interacting with algorithms as they do for interacting with humans (Nass & Moon, 2000), both requirements can be fulfilled by using anthropomorphisation. Research has shown that the use of anthropomorphic design elements can activate the knowledge individuals have about human decision-making (Epley et al., 2007). As a result, individuals tend to evaluate a human-like algorithm similarly to the way they evaluate the decision-making of other humans and also attribute the ability to behave intentionally to it (Puzakova et al., 2013). Furthermore, according to the person-positivity bias (Moon & Conlon, 2002), human decision-making should be more favourably evaluated than algorithmic decision-making. This is also underlined by the finding that individuals broadly anticipate positive actions from other humans and judge these actions to be guided by positive intentions (Ybarra, 2002). Thus, assigning human attributes to algorithmic decision makers might lead to the perception of algorithms as acting not solely reductionistic but also empathising with the individual. On the other hand, mimicking the interaction with a human decision maker supports the individual's perceptions of whether the algorithmic decision maker is truthful and justifies the decision outcome adequately. Accordingly, we propose the following two hypotheses:

H3a. Anthropomorphisation of the algorithmic decision maker positively influences interpersonal justice.

H3b. Anthropomorphisation of the algorithmic decision maker positively influences informational justice.

3.3.3 | Perceived algorithmic fairness and individual responses

Central to organisational justice theory is the investigation of individual responses as a consequence of fairness perceptions (McCarthy et al., 2017). Several lines of evidence have suggested that it pays for organisations to engage with the concept of fairness, for example, to prevent negative word of mouth (Cropanzano et al., 2007).

These assumptions led scholars to associate higher ratings of 'what is fair' with positive individual responses toward the organisation. Conversely, low fairness perceptions generate unfavourable individual responses (Gilliland, 1993). Thereby, individual responses can be either internal and non-visible or external and visible (Jacoby, 2002). So far, however, the fairness literature does not provide clear guidance regarding possible individual responses to algorithmic fairness perceptions.

For the context of algorithmic decision-making, we argue that an important internal response is the individual's satisfaction with the algorithmic decision maker (Colquitt, 2001). Traditionally, satisfaction with the decision maker has played a fundamental role in fairness research. Previous research has established that the decision maker is the initial point of contact and characterises the quality of the relationship between an organisation and an individual (Masterson

et al., 2000). Indeed, applicants often see the hiring process as a predictor of employee treatment and organisational values (Walker et al., 2013). Therefore, applicants who perceive a company to be fair are more likely to be satisfied with the decision maker, while applicants who believe the company is behaving unfairly will be dissatisfied with the decision maker (Ababneh et al., 2014). This evidence, along with the notion that individuals have the same requirements for algorithmic and human decision makers (Nass & Moon, 2000), led us to propose the following hypothesis:

> **H4a.** Perceived algorithmic fairness positively influences satisfaction with an algorithmic decision maker.

Next to satisfaction as an internal individual response, prior research on organisational justice suggested recommendation intentions as a critical external individual response toward fairness perceptions (Ababneh et al., 2014). High fairness perceptions can encourage applicants to recommend the organisation to others, whereas low fairness perceptions generate unfavourable individual responses such as negative worth-of-mouth (Ababneh et al., 2014; Gilliland, 1993). One explanation for this observation is that applicants, who perceive that organisations implemented a fair recruiting process, are more likely to reciprocate by expressing positive responses, while those who perceive fairness rules to be violated are more likely to express negative responses (Hausknecht, 2004; Schinkel et al., 2016). Therefore, we argue that the extent to which applicants perceive the algorithmic hiring process to be fair determines the likelihood that applicants respond reciprocally by recommending the organisation to others. Thus, we hypothesize:

> **H4b.** Perceived algorithmic fairness positively influences the likelihood to recommend an organisation to friends and colleagues.

To test the derived hypotheses, we conducted a randomised online experiment with 801 participants.

4 | METHODOLOGY

In the present study, we employed a quantitative experimental approach and empirically tested our hypotheses. The following sections outline the methodology.

4.1 | Experimental design

We focused on the personnel selection process. Participants applied for admission to a fictitious 'special participant pool for future market studies'. Throughout the experiment, participants were made to believe that they were actually applying for the participant pool, which gave us the opportunity to investigate the effects of our experimental manipulations in a realistic setting.

The experiment consisted of eight experimental groups to which participants were randomly allocated at the beginning of the experiment. The experimental factors were *transparency of the algorithmic decision* (transparent vs. non-transparent) and *appearance of the algorithmic decision maker* (anthropomorphic vs. non-anthropomorphic). With prior research suggesting that a decision's outcome can have significant effects on fairness perceptions (Colquitt et al., 2001; Newman et al., 2020; Ployhart & Ryan, 1997; Yalcin et al., 2022), we also manipulated the *outcome of the personnel selection decision* (positive vs. negative). We further added two groups in which the origin of the hiring decision was human and either positive or negative as an additional baseline comparison to show whether the dimensions of organisational justice theory that were initially used to evaluate the fairness perceptions of human decision makers could also be applied to algorithmic decision makers. In these groups, we informed our participants

that an HR employee was responsible for assessing their application. Table 1 displays the experimental groups with the respective treatments.

To manipulate the *outcome of the hiring decision*, participants were either informed that their application for the special participant pool was accepted (positive) or rejected (negative) during the experiment (Yalcin et al., 2022). The *transparency of the algorithmic decision* was manipulated by adding a short and simple explanation (Cadario et al., 2021) informing participants that the algorithm would evaluate their application by reviewing their demographic criteria and their mindset and using prior data of successful applicants to assess their fit for the special participant pool. Lastly, we manipulated the design of the algorithmic decision maker (see Figure 2) by implementing human-identity design (assigning it the avatar of a human person and giving it the name 'Emily'), non-verbal design (simulating a personal communication with animations of a chat-like conversation), and verbal design (using greetings, typos, and emojis in the communication to the application process; Benlian et al., 2020; Han et al., 2019; Seeger et al., 2021; Yalcin et al., 2022).

4.2 | Experimental task

For our experimental task, participants on the online sampling platform Prolific were invited to apply for being admitted to the special participant pool of the company Skillistic Insights. Prolific participants are essentially gig economy workers who are paid to participate in surveys and often earn their primary income from doing that (Keith et al., 2019; Peer et al., 2021). Our participants were paid for their application (on average 8 £/h) and informed that Skillistic Insights generally provided generous compensation as well as bonus payments. The application procedure was described as involving the participant's mindset and their Prolific experience. In addition, participants were informed that they would

TABLE 1 Overview of experimental groups.

Decision outcome	Decision maker			
	Human	Algorithm	Transparent algorithm	Anthropomorphic algorithm
Positive	H^+	A^+	A^+_T	A^+_A
Negative	H^-	A^-	A^-_T	A^-_A

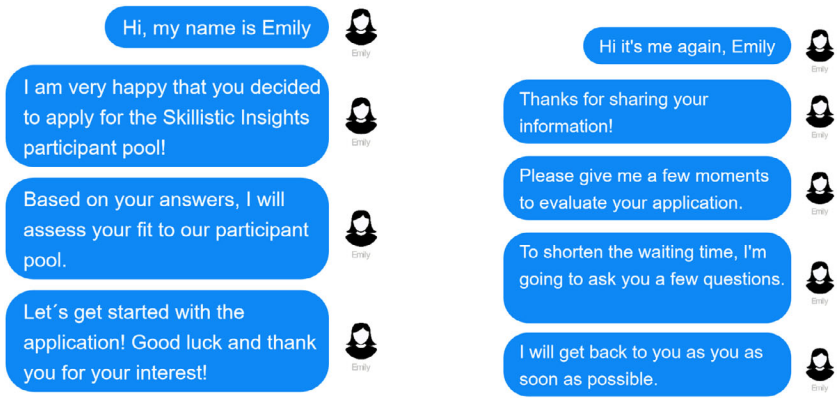

FIGURE 2 Figure shows examples of the chat-like conversations in the anthropomorphic design condition. Instead of the female user icon, a real woman's face was displayed, and the chat also contained some emojis. Unfortunately, the visuals cannot be display here due to copyright reasons. However, the authors can provide them upon request.

receive the outcome of their application after a short waiting period. Once participants accepted that they wanted to apply, they were 'redirected' to the application portal (i.e., the next survey page that was designed accordingly).

On the first page of the application portal, participants received additional information according to their experimental group. To make sure that participants understood that their application was reviewed by an algorithm (or human), they had to choose the correct answer to a control question that directly followed the information (i.e., 'Based on the information you have just read, how is the evaluation of your application organised?'). The subsequent pages included the self-efficacy scale by Chen et al. (2001) as a filler task, a question regarding the approximate total number of studies the participants had completed so far, and an open-ended question asking them to elaborate on why the respective decision maker should consider them for the participant pool (50–250 characters).

Afterward, participants were informed about the decision maker's decision and answered multiple additional questions and scales. At the end of the survey, participants were fully debriefed regarding the non-existence of Skillistic Insights, the non-evaluation of their input data and the randomness of the decision, and the study's purpose. They further received additional compensation of 0.10 £ to thank them for their input.

4.3 | Measures and scales

4.3.1 | Algorithmic fairness perceptions

To measure participants' (algorithmic) fairness perceptions, we relied on the organisational justice scales by Colquitt et al. (2001) and adapted them to our specific use case (see Table A1 in Appendix A).

Procedural justice

We measured the participants' procedural justice perceptions regarding the decision maker using an adapted version of the procedural justice scale by Colquitt et al. (2001). Participants stated the extent to which they agreed with the respective statements using a 7-point Likert scale (1 = Strongly disagree, 7 = Strongly agree). The scale showed sufficient internal consistency (Cronbach's $\alpha = 0.89$).

Distributive justice

We measured the participants' distributive justice perceptions regarding the decision maker using an adapted version of the distributive justice scale by Colquitt et al. (2001). Participants stated the extent to which they agreed with the respective statements using a 7-point Likert scale (1 = Strongly disagree, 7 = Strongly agree). The scale showed very high internal consistency (Cronbach's $\alpha = 0.97$).

Interpersonal justice

The participants' interpersonal justice perceptions regarding the decision maker were measured using an adapted version of the interpersonal justice scale by Colquitt et al. (2001). Participants stated the extent to which they agreed with the statements using a 7-point Likert scale (1 = Strongly disagree, 7 = Strongly agree). The scale showed sufficient internal consistency (Cronbach's $\alpha = 0.90$).

Informational justice

To measure the participants' informational justice perceptions regarding the decision maker, we adapted the informational justice scale by Colquitt et al. (2001). Participants stated the extent to which they agreed with the statements using a 7-point Likert scale (1 = Strongly disagree, 7 = Strongly agree). The scale showed sufficient internal consistency (Cronbach's $\alpha = 0.86$).

4.3.2 | Individual responses

To assess the participants' attitude toward the company, we used the item 'What is your overall evaluation of Skillistic Insights?', adapted from Yalcin et al. (2022). Participants stated the extent to which they agreed with the statement using a 7-point Likert scale (1 = 'Strongly disagree', 7 = 'Strongly agree').

The participants' satisfaction with the decision maker was measured using the satisfaction scale by Lin et al. (2021). Participants stated the extent to which they agreed with the statements using a 7-point Likert scale (1 = 'Strongly disagree', 7 = 'Strongly agree'). The scale showed very high internal consistency (Cronbach's $\alpha = 0.98$).

We further used an 11-point Kunin scale (Kunin, 1955) to assess how likely participants were to recommend the company 'Skillistic Insights' to their friends or other Prolific participants. The scale ranged from 0 = 'Not likely at all' to 10 = 'Extremely likely'. While this item also resembled typical Net Promoter Score (NPS) questions, we did not categorise the participants accordingly but rather used this question as many consumers are familiar with it due to its widespread application (Baehre et al., 2022).

4.3.3 | Manipulation checks

For our manipulation check, we omitted the two experimental groups in the human decision maker groups, as they did not receive either the transparency or the anthropomorphic design manipulation.

To measure whether our transparency manipulation led to participants perceiving the underlying reasoning of the algorithmic decision as more transparent, we used an adaptation of the data use transparency scale by Martin et al. (2017). Participants stated the extent to which they agreed with the statements using a 7-point Likert scale (1 = 'Strongly disagree', 7 = 'Strongly agree'). The scale showed high internal consistency (Cronbach's $\alpha = 0.93$). Participants in the transparency conditions ($M = 4.55$, $SD = 1.46$) and the non-transparency conditions ($M = 3.93$, $SD = 1.55$) differed significantly regarding their perceived transparency of the underlying reasoning of the algorithmic decision ($t(430.04) = -4.80$, $p < 0.001$). This result indicates that our transparency manipulation was successful.

We measured the success of our anthropomorphic design manipulation using two items proposed by Yalcin et al. (2022). Participants stated the extent to which they agreed with the statements using a 7-point Likert scale (1 = 'Strongly disagree', 7 = 'Strongly agree'). The scale showed high internal consistency (Cronbach's $\alpha = 0.91$). Participants in the anthropomorphic design conditions ($M = 3.80$, $SD = 1.44$) and the non-anthropomorphic design conditions ($M = 2.65$, $SD = 1.45$) differed significantly regarding their perception of the decision maker as anthropomorphic ($t(392.43) = -9.22$, $p < 0.001$). This result indicates that our anthropomorphic design manipulation was successful.

4.4 | Participants

We used the following prescreening criteria to collect a sample that closely resembles people that could be confronted with algorithms in hiring processes without being too restrictive (see Table 2 for a demographic overview): location (UK), age (18–60), fluent languages (English), employment status (full-time, part-time, unemployed (and job seeking), and highest education completed (at least high school diploma/A-levels). In total, 801 participants completed our experiment with a median completion time of approximately 7 min.

TABLE 2 Demographic summary of participants.

	Variable	H^+	H^-	A^+	A^-	A^+_T	A^-_T	A^+_A	A^-_A
n		100	96	101	102	102	102	99	99
Gender	Female	54 (54.0%)	47 (49.0%)	52 (51.5%)	47 (46.1%)	46 (45.1%)	56 (54.9%)	49 (49.5%)	49 (49.5%)
	Male	46 (46.0%)	49 (51.0%)	49 (48.5%)	55 (53.9%)	56 (54.9%)	46 (45.1%)	50 (50.5%)	50 (50.5%)
Age (yrs)	*M* (SD)	35.7 (10.9)	38.1 (11.0)	36.5 (10.8)	36.6 (9.7)	36.2 (10.4)	34.2 (10.4)	35.5 (9.8)	36.1 (10.3)
Education	Doctorate degree	7	4	2	3	1	1	4	1
	Graduate degree	22	23	17	18	12	17	15	18
	Undergraduate degree	33	40	47	50	44	52	48	46
	Technical college	8	13	14	14	14	17	13	15
	High school diploma	29	16	21	16	31	25	19	19
	NA	1	0	0	1	0	0	0	0
Employment	Full-Time	67	63	73	69	80	64	77	68
	Part-Time	21	25	23	23	15	27	14	19
	Unemployed (job seeking)	11	8	4	10	7	11	8	2
	NA	1	0	1	0	0	0	0	0

5 | RESULTS

5.1 | Randomization checks

We calculated multiple randomization checks to ensure that our procedure led to an equal distribution of key demographics across our experimental groups. Neither for gender distribution ($\chi^2(7) = 3.38$, $p = 0.848$), age ($F(7, 793) = 1.14$, $p = 0.338$), highest level of education ($\chi^2(35) = 42.66$, $p = 0.175$), nor current status of employment ($\chi^2(21) = 28.02$, $p = 0.140$) did the analyses reveal any significant differences between the groups.

5.2 | Hypotheses testing

To test our hypotheses, we fitted regression models to the data and considered p values below the threshold of $\alpha = 0.05$ as significant. In Table 3, we summarised the hypothesis tests and their corresponding results.

5.2.1 | Perceived algorithmic fairness

For our first hypothesis, we investigated whether the adopted algorithmic fairness perceptions based on organisational justice theory (Colquitt, 2001) were similarly indicative of the participant's attitude toward the company for the experimental groups with an algorithmic decision maker as for the groups with a human decision maker.

We included the attitude toward the company to test our hypothesis H1 as prior research has considerably demonstrated the consequences of fairness perceptions on individuals' attitudes and behaviours (Judge & Colquitt, 2004). Fairness perceptions generally affect the formation and change of attitudes, thus influencing an individual's attitude toward a particular object (Ajzen, 1982). For example, a number of studies have found fairness to be a reliable predictor of attitudes toward institutions and agencies, a particular decision outcome, and contextual outcomes (Cohen-Charash & Spector, 2001). As organisational justice theory proposed no single fairness variable, we were required to find a proxy to test our hypothesis. Including attitude, which is a central variable in fairness literature, thus enabled us to test if the four justice dimensions are suitable to broaden our understanding of algorithmic fairness perceptions and if there are any differences between human and algorithmic decision makers.

We fitted a multiple linear regression model predicting the participants' attitudes with the following independent variables: procedural justice, distributive justice, interpersonal justice, and informational justice (all reflecting the participants' scores on the respective scales) as well as algorithmic decision maker (dummy variable: 1 if the decision maker was an algorithm, 0 otherwise), and the respective interactions of the dummy variable and the scale values.

TABLE 3 Overview of hypothesis tests and corresponding results.

Hypothesis	Regression models	Results
H1	1	Supported
H_{2a}	2	Not supported
H_{2b}	3	Not supported
H_{2c}	5	Supported
H_{3a}	4	Supported
H_{3b}	5	Supported
H_{4a}	6	Supported
H_{4b}	7	Supported

The model was highly significant in explaining participants' attitude toward the company ($F(9, 791) = 267.8$, $p < 0.001$, $R^2 = 0.75$) and revealed multiple significant main effects (see Table 4). In line with the findings of prior research, all organisational justice dimensions predicted the participants' attitude toward the company, indicated by their significant main effects. There was no significant main effect of algorithmic decision maker, suggesting that participants equal attitudes irrespective of the nature of the decision maker. Importantly, none of the interaction effects of the justice dimensions with the algorithmic decision maker were significant. This indicates that these dimensions were equally predictive for the algorithmic decision maker as they were for the human decision maker, supporting our H1.

5.2.2 | Technological interventions and perceived algorithmic fairness

For our hypotheses H_{2a}–H_{2c} and H_{3a}–H_{3b}, we calculated four multiple regression models to predict each of the algorithmic fairness perception dimensions. The respective models predicted the respective dimension with the following independent variables: algorithmic decision maker (dummy variable: 1 if the decision maker was an algorithm, 0 otherwise), transparency (dummy variable: 1 if the algorithm's decision was transparent), anthropomorphism (dummy variable: 1 if the algorithmic decision maker was designed in an anthropomorphic way, 0 otherwise), positive outcome (dummy variable: 1 if the decision outcome was positive, 0 otherwise), and the respective interaction terms of a positive outcome and the other predictors.

For the participants' procedural justice, the model (see Table 5) was significant ($F(7, 793) = 50.35$, $p < 0.001$, $R^2 = 0.30$). It showed a significant positive main effect of positive outcome, indicating that participants who received a positive decision outcome considered the procedure as fairer than participants who received a negative decision outcome. The regression further showed a marginally significant negative main effect of anthropomorphism, providing suggestive evidence that using anthropomorphic design features for algorithmic decision makers reduces procedural fairness. No other effects reached statistical significance, which is not in line with our H_{2a}.

TABLE 4 Linear regression model 1—Participants' attitudes toward the company.

	Dependent variable Attitude toward the company
Procedural justice	0.22** (0.09)
Distributive justice	0.45*** (0.04)
Interpersonal justice	0.27*** (0.07)
Informational justice	0.19*** (0.07)
Algorithmic decision maker	−0.41 (0.33)
Procedural justice × Algorithmic decision maker	0.11 (0.10)
Distributive justice × Algorithmic decision maker	0.04 (0.05)
Interpersonal justice × Algorithmic decision maker	−0.03 (0.08)
Informational justice × Algorithmic decision maker	−0.03 (0.08)
Constant	−0.61** (0.28)
Observations	801
R^2	0.75
Adjusted R^2	0.75
Residual Std. Error	0.89 (df = 791)
F statistic	267.82*** (df = 9; 791)

*$p < 0.1$;
**$p < 0.05$;
***$p < 0.01$.

TABLE 5 Linear regression model 2—Participants' procedural justice.

	Dependent variable Procedural justice
Algorithmic decision maker	−0.01 (0.14)
Transparency	0.21 (0.14)
Anthropomorphism	−0.25* (0.14)
Positive outcome	1.27*** (0.14)
Algorithmic decision maker × Positive outcome	0.10 (0.20)
Transparency × Positive outcome	−0.23 (0.20)
Anthropomorphism × Positive outcome	0.17 (0.20)
Constant	3.37*** (0.10)
Observations	801
R^2	0.31
Adjusted R^2	0.30
Residual Std. Error	1.01 (df = 793)
F statistic	50.35*** (df = 7; 793)

*$p < 0.1$;
**$p < 0.05$;
***$p < 0.01$.

TABLE 6 Linear regression model 3—Participants' distributive justice.

	Dependent variable Distributive justice
Algorithmic decision maker	0.26 (0.16)
Transparency	0.04 (0.16)
Anthropomorphism	−0.07 (0.16)
Positive outcome	3.27*** (0.17)
Algorithmic decision maker × Positive outcome	−0.18 (0.23)
Transparency × Positive outcome	−0.18 (0.23)
Anthropomorphism × Positive outcome	−0.01 (0.23)
Constant	2.28*** (0.12)
Observations	801
R^2	0.64
Adjusted R^2	0.64
Residual Std. Error	1.16 (df = 793)
F statistic	203.58*** (df = 7; 793)

*$p < 0.1$;
**$p < 0.05$;
***$p < 0.01$.

The model for the participants' distributive justice (see Table 6) was also significant in its prediction ($F(7, 793) = 203.58$, $p < 0.001$, $R^2 = 0.64$). It showed a significant positive main effect of positive outcome, indicating that participants who received a positive decision outcome had higher distributive justice than participants who received a negative decision outcome. No other effects reached statistical significance, which is not in line with our H_{2b}.

TABLE 7 Linear regression model 4—Participants' interpersonal justice.

	Dependent variable Interpersonal justice
Algorithmic decision maker	0.02 (0.15)
Transparency	0.24 (0.15)
Anthropomorphism	0.84*** (0.15)
Positive outcome	1.10*** (0.15)
Algorithmic decision maker × Positive outcome	0.18 (0.22)
Transparency × Positive outcome	−0.28 (0.21)
Anthropomorphism × Positive outcome	−0.47** (0.21)
Constant	4.53*** (0.11)
Observations	801
R^2	0.24
Adjusted R^2	0.23
Residual Std. Error	1.08 (df = 793)
F statistic	34.95*** (df = 7; 793)

*$p < 0.1$;
**$p < 0.05$;
***$p < 0.01$.

The participants' interpersonal justice were also significantly predicted by the independent variables ($F(7, 793) = 34.95$, $p < 0.001$, $R^2 = 0.23$). The model (see Table 7) showed a significant positive main effect of anthropomorphism, indicating that participants for which the algorithmic decision maker was designed in an anthropomorphic way had higher interpersonal justice than participants who interacted with a non-anthropomorphic algorithmic decision maker. This result is in line with our H_{3a}. In addition, the model showed a significant positive main effect of positive outcome, indicating that participants who received a positive decision outcome had higher interpersonal justice than participants who received a negative decision outcome. Lastly, the model showed a significant negative interaction effect of anthropomorphism and positive outcome. This suggests that for positive decision outcomes, anthropomorphism has a weaker effect on interpersonal justice than for negative decision outcomes.

For the participants' informational justice, the model (see Table 8) was significant ($F(7, 793) = 23.03$, $p < 0.001$, $R^2 = 0.16$). It showed a significant positive main effect of transparency, indicating that participants who received a transparent explanation for the algorithm's decision had higher informational justice than participants who did not receive this explanation. This result is in line with our H_{2c}. The model also showed a significant positive main effect of anthropomorphism, indicating that participants for which the algorithmic decision maker was designed in an anthropomorphic way had higher informational justice than participants who interacted with a non-anthropomorphic algorithmic decision maker. This result is in line with our H_{3b}. The regression further showed a significant positive main effect of positive outcome, indicating that participants who received a positive decision outcome had higher informational justice than participants who received a negative decision outcome.

5.2.3 | Perceived algorithmic fairness and organisational outcomes

To assess whether the participants' perceived algorithmic fairness perceptions affected their satisfaction with the decision maker, we ran a linear regression model. To fit the model, we included the algorithmic fairness perception dimensions as predictors of the satisfaction with the decision maker and only took the groups that were informed that an algorithm assessed their application into account.

TABLE 8 Linear regression model 5—Participants' informational justice.

	Dependent variable Informational justice
Algorithmic decision maker	−0.07 (0.17)
Transparency	0.48*** (0.17)
Anthropomorphism	0.36** (0.17)
Positive outcome	0.90*** (0.17)
Algorithmic decision maker × Positive outcome	0.07 (0.24)
Transparency × Positive outcome	0.05 (0.24)
Anthropomorphism × Positive outcome	0.14 (0.24)
Constant	3.68*** (0.12)
Observations	801
R^2	0.17
Adjusted R^2	0.16
Residual Std. Error	1.22 (df = 793)
F statistic	23.03*** (df = 7; 793)

*$p < 0.1$;
**$p < 0.05$;
***$p < 0.01$.

TABLE 9 Linear regression model 6—Participants' satisfaction with the decision maker.

	Dependent variable Satisfaction
Procedural justice	0.33*** (0.05)
Distributive justice	0.52*** (0.03)
Interpersonal justice	0.17*** (0.04)
Informational justice	0.19*** (0.04)
Constant	−1.17*** (0.18)
Observations	605
R^2	0.76
Adjusted R^2	0.76
Residual Std. Error	0.90 (df = 600)
F statistic	467.24*** (df = 4; 600)

*$p < 0.1$;
**$p < 0.05$;
***$p < 0.01$.

The model (see Table 9) was significant in predicting the participants' satisfaction with the decision maker ($F(4, 600) = 467.24$, $p < 0.001$). It showed significant positive main effects for all predictors. This means that higher algorithmic fairness perceptions, indicated by the four dimensions, positively influenced the satisfaction with the algorithmic decision maker. This result is in line with our H_{4a}.

For our H_{4b}, we calculated an additional regression model, predicting participants' subjective likelihood ratings for recommending the company that assessed their application. We used the same predictors as for the participants' satisfaction with the decision maker and, again, only took the groups that were informed that an algorithm assessed their application into account.

TABLE 10 Linear regression model 7—Participants' recommendation likelihoods.

	Dependent variable Recommendation likelihood
Procedural justice	0.66*** (0.10)
Distributive justice	0.91*** (0.06)
Interpersonal justice	0.30*** (0.08)
Informational justice	0.16* (0.09)
Constant	−2.41*** (0.39)
Observations	605
R^2	0.65
Adjusted R^2	0.65
Residual Std. Error	1.97 (df = 600)
F statistic	282.64*** (df = 4; 600)

*$p < 0.1$;
**$p < 0.05$;
***$p < 0.01$.

The model was significant in predicting the participants' ratings ($F(4, 600) = 282.6$, $p < 0.001$, $R^2 = 0.65$) and showed multiple significant main effects (see Table 10). Procedural justice, distributive justice, and interpersonal justice had a significant positive main effect on participants' recommendation likelihood ratings. Informational justice also showed a positive main effect that was, however, only marginally significant. This result supports our H_{4b}.

6 | DISCUSSION

With AI entering the HR departments, (van den Broek et al., 2021) the study follows two objectives. We first, aim to identify which factors determine the perceptions of algorithmic fairness. With S-O-R, we investigated whether the four justice dimensions, procedural, distributive, interpersonal, and informational justice, are applicable to capture algorithmic fairness perceptions. Second, we explored the implementation of transparency and anthropomorphism interventions to shape algorithmic fairness perceptions and uncovered the implications of algorithmic fairness perceptions.

Based on data from an online experiment with 801 participants, we demonstrated that the four justice dimensions are appropriate for algorithmic decision making, as procedural, distributive, interpersonal, and informational justice mirror algorithmic fairness requirements. We further identified transparency as a suitable mechanism to influence informational justice. Contrary to our hypotheses, this study did not find empirical support for transparency as an antecedent of procedural and distributive justice. However, the findings indicated a positive influence of anthropomorphism on informational and interpersonal justice. Another finding was the impact of algorithmic fairness perceptions on individual responses, such as satisfaction with the decision maker and the likelihood of referring the organisation to others. The following sections describe our contributions and implications in more detail.

6.1 | Theoretical implications

The findings contribute to research on organisational justice theory and algorithmic fairness perceptions.

First, fairness perceptions are highly subjective (Cropanzano et al., 2007). This is visible in the ambiguous findings on perceived algorithmic fairness from earlier literature (Newman et al., 2020; Suen et al., 2019; Yalcin et al., 2022). Given the consequences of algorithmic fairness perceptions, we demonstrate that perceptions of

fairness should be considered equally important for individual responses as an algorithm's ability to produce fair decision outcomes. This is in line with existing literature (Dolata et al., 2022; Feuerriegel et al., 2020) theorising the importance of fairness perceptions and extends our current understanding by capturing the multidimensional nature of algorithmic fairness perceptions and proposing factors to determine them.

Second, we test the applicability of organisational justice theory in the context of algorithmic decision-making (Starke et al., 2021). Organisational justice theory was developed to describe fairness perceptions of human decision makers and it remains to be clarified if individuals have the same requirements for the fairness evaluation of algorithmic compared to human decision makers. We complement prior research (Lee, 2018; Robert et al., 2020) by systematically demonstrating that the four dimensions originally proposed by organisational justice theory can be used for the fairness assessment of algorithmic decision makers. We contribute by highlighting the importance of equally considering all four dimensions to gain a holistic understanding and a differentiated picture of algorithmic fairness perceptions. This contribution guides research with the further investigation of algorithmic fairness perceptions. Thereby, it addresses the second shortcoming of prior research: While studies have been conducted that investigated selected justice dimensions to broaden our understanding of algorithmic fairness perceptions, they have not yet treated the multidimensionality of fairness in much detail. However, it is well established from several studies that fairness dimensions are distinct yet complement each other (Colquitt et al., 2001; Cropanzano et al., 2007). This indicates that when one justice dimension is violated, potential adverse effects can be compensated by adequately addressing the further justice dimensions. For example, organisations can compensate applicants' perceptions of an unfair selection process by ensuring candid and respectful communication with the applicant (Cropanzano et al., 2007). Indeed, prior research has argued that perceptions of interpersonal and informational justice are powerful predictors of procedural justice (Colquitt et al., 2001). Our study considers these possible interdependencies adding to a holistic understanding and a differentiated picture of algorithmic fairness perceptions.

Third, our study also proposes concrete mechanisms to form algorithmic fairness perceptions. Individuals must perceive them as fair to benefit from algorithms that produce fair selection decisions. Prior research has demonstrated the importance of transparent and empathetic decision makers for individual fairness perceptions in personnel selection procedures (Cropanzano et al., 2007). We unveiled that additional information on an algorithm's underlying processes and the use of anthropomorphic design elements shape perceptions of interpersonal and informational justice. Following this approach, we complement prior research on interventions aiming to induce algorithmic acceptance and enrich these research endeavours by considering the interventions' implications for algorithmic fairness perceptions (Felzmann et al., 2019; Schanke et al., 2021). The higher algorithmic fairness perceptions, in turn, result in satisfaction with the algorithmic decision maker and the likelihood of recommending the organisation to others.

Forth, by drawing on S-O-R, we follow the call of IS research to contextualise established theories, thereby enriching their explanatory and predicative power (Hong et al., 2014). We used S-O-R as a guiding theory to show that it can be applied to the context of perceived algorithmic fairness. Thereby we demonstrate that the contextualization of the theory helps researchers and practitioners to manage and solve practical problems such as proactively increasing fairness perceptions of algorithms.

6.2 | Practical implications

The findings also provide implications for managers and practitioners. Based on our findings, we derive the following recommendations.

First, applicants seem to be generally open to algorithmic selection procedures as they do not differentiate their evaluation of human vs. algorithmic decision makers. While advanced algorithmic methods can help overcome the current systematic discrimination of certain demographic groups, it is important to ensure that these methods are used in a fair and unbiased manner. Therefore, we encourage organisations to use advanced algorithmic methods to foster applicants' equality only if they have been thoroughly tested, reveal superior performance, and lead evidently to objectively fair outcomes. In this realm, the combination of human and algorithmic strengths can be also

considered to achieve the fairest results. Moreover, organisations should be aware of the multidimensional nature of fairness perceptions. Focusing on different fairness aspects in the personnel selection process enables organisations to increase the probability that applicants perceive the AI in HR processes to be fair. For example, offering objective algorithmic allocation processes, following social norms, and respectful interpersonal treatment is expected to lead to higher algorithmic fairness perceptions. However, organisations could compensate applicants' perceptions of an unfair selection process by ensuring candid and respectful communication with the applicant through the interdependencies of these requirements (Cropanzano et al., 2007).

Second, applicants have highly similar demands for the interaction with algorithms as they have with humans. Due to the abstract nature of algorithms, however, fulfilling these demands might require different actions. One such action could be the implementation of transparency and anthropomorphism interventions. Both mechanisms help organisations shape informational and interpersonal justice and overcome the impression of unempathetic and impersonal algorithms.

Third, algorithmic fairness perceptions influence applicants' satisfaction with a decision maker and the recommendation purpose, both essential aspects for a favourable reputation of organisations. The proposed interventions may be of assistance to ensure positive responses and help organisations overcome the potential concerns of applicants.

6.3 | Limitation and future research

Even though our research provides valuable insights into algorithmic fairness perceptions and introduces interventions to shape them actively, our work includes some limitations that offer opportunities for future research.

First, we empirically tested our hypotheses using a scenario-based vignette in an online survey. The employed experiment created a realistic personnel selection scenario with gig workers that persuaded applicants to apply for a unique participant pool. Testing the hypotheses in an application setting that is not driven by the specifics of the gig working economy such as conducting a field study within a company can provide additional insights into the effectiveness of the proposed interventions and the individual responses to algorithmic fairness perceptions.

Second, while our study design allowed us to influence informational and interpersonal justice, our interventions were not successful in positively affecting procedural and distributive justice. To expand our knowledge of algorithmic fairness perceptions, we encourage further research to systematically investigate why interactional justice dimensions can be influenced more easily. One possible explanation would be that interpersonal treatment is easier to manipulate as it is based on human communication rules that address the subjective requirements of empathetic and transparent language. In contrast, procedural and distributive justice might be less easy to grasp and need, therefore, more detailed explanations of the concrete selection procedures used and allocation rules applied. Correspondingly, how we manipulated anthropomorphism and transparency might also influence how the respondents perceived algorithmic fairness. Future research could address the question of the impact that different levels of human likeness and transparency might have on the perceived fairness of algorithms. It is conceivable, for example, that overuse of anthropomorphic design elements might backfire due to the uncanny valley effect (Mori et al., 2012) or that additional information is overwhelming for non-experts (You et al., 2022).

Third, although the outcome favourability was included as a control variable in our regression models, the effect was comparably strong across the different models. This finding aligns with several other studies (Yalcin et al., 2022). However, it is all the more compelling that our interventions had the reported effects, irrespective of the significant impact of the decision outcome.

6.4 | Conclusion

Algorithms are increasingly entering organisational HR processes. Our study offers a differentiated view of algorithmic fairness perceptions in personnel selection. We proposed that the rapid advancements in the development of

fair and objective algorithms have the potential to overcome the prevailing unequal treatment of applicants. Experimenting with a realistic application scenario, we provided a valuable foundation for understanding how justice dimensions proposed by organisational justice theory can contribute to explaining algorithmic fairness perceptions. We found that organisations can actively rely on interventions such as transparency and anthropomorphism to shape fairness perceptions. We further encouraged organisations to reflect on algorithmic decision-making regarding fairness perceptions, as these perceptions significantly affect how individuals see the respective organisation.

ACKNOWLEDGMENT

The authors thank the editor-in-chief, the senior editor, the anonymous associate editor, and the reviewers for their valuable feedback during the review process. Open Access funding enabled and organized by Projekt DEAL.

FUNDING INFORMATION

Parts of this research were supported by the Software Campus 2.0 Program funded by the Federal Ministry of Education and Research (BMBF) and the Bavarian State Ministry of Science and the Arts, in a program coordinated by the Bavarian Research Institute for Digital Transformation (bidt).

DATA AVAILABILITY STATEMENT

Data available on request due to privacy/ethical restrictions: The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

ORCID

Leonard Michels https://orcid.org/0000-0001-6597-498X

Sven Laumer https://orcid.org/0000-0002-9430-5774

REFERENCES

Ababneh, K. I., Hackett, R. D., & Schat, A. C. H. (2014). The role of attributions and fairness in understanding job applicant reactions to selection procedures and decisions. *Journal of Business and Psychology*, *29*(1), 111–129.

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

Adadi, A., & Berrada, M. (2018). Peeking inside the black-box: A survey on explainable artificial intelligence (XAI). *IEEE Access*, *6*, 52138–52160.

Ajunwa, I. (2020). The "black box" at work. *Big Data & Society*, *7*(2), 205395172096618.

Ajzen, I. (1982). Equity in attitude formation and change. In J. Greenberg & R. L. Cohen (Eds.), *Equity and justice in social behavior* (pp. 161–186). Elsevier.

Almås, I., Cappelen, A. W., Sørensen, E., & Tungodden, B. (2010). Fairness and the development of inequality acceptance. *Science*, *328*(5982), 1176–1178.

Annabi, H., & Lebovitz, S. (2018). Improving the retention of women in the IT workforce: An investigation of gender diversity interventions in the USA. *Information Systems Journal*, *28*(6), 1049–1081.

Avery, D. R., McKay, P. F., & Wilson, D. C. (2008). What are the odds? How demographic similarity affects the prevalence of perceived employment discrimination. *Journal of Applied Psychology*, *93*(2), 235–249.

Baehre, S., O'Dwyer, M., O'Malley, L., & Lee, N. (2022). The use of net promoter score (NPS) to predict sales growth: Insights from an empirical investigation. *Journal of the Academy of Marketing Science*, *50*(1), 67–84.

Bai, B., Dai, H., Zhang, D., Zhang, F., & Hu, H. (2021). The impacts of algorithmic work assignment on fairness perceptions and productivity: Evidence from field experiments. *AOM Annual Meeting Proceedings 2021*.

Bansal, G., & Zahedi, F. M. (2015). Trust violation and repair: The information privacy perspective. *Decision Support Systems*, *71*, 62–77.

Barredo Arrieta, A., Díaz-Rodríguez, N., Del Ser, J., Bennetot, A., Tabik, S., Barbado, A., Garcia, S., Gil-Lopez, S., Molina, D., Benjamins, R., Chatila, R., & Herrera, F. (2020). Explainable explainable artificial intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. *Information Fusion*, *58*, 82–115.

Bauer, K., & Gill, A. (2023). Mirror, mirror on the wall: Algorithmic assessments, transparency, and self-fulfilling prophecies. *Information Systems Research* (forthcoming). https://doi.org/10.1287/isre.2023.1217

Bauer, T. N., Truxillo, D. M., Sanchez, R. J., Craig, J. M., Ferrara, P., & Campion, M. A. (2001). Applicants reactions to selection: Development of the selection procedural justice scale (SPJS). *Personnel Psychology*, *54*(2), 387–419.

Belk, R. W. (1975). Situational variables and consumer behavior. *Journal of Consumer Research*, *2*(3), 157.

Benlian, A., Klumpe, J., & Hinz, O. (2019). Mitigating the intrusive effects of smart home assistants by using anthropomorphic design features: A multimethod investigation. *Information Systems Journal*, *30*(6), 1010–1042. https://doi.org/10.1111/isj.12243

Benlian, A., Klumpe, J., & Hinz, O. (2020). Mitigating the intrusive effects of smart home assistants by using anthropomorphic design features: A multimethod investigation. *Information Systems Journal*, *30*(6), 1010–1042.

Bertrand, M., & Mullainathan, S. (2004). Are Emily and Greg more employable than Lakisha and Jamal? A field experiment on labor market discrimination. *The American Economic Review*, *94*(4), 129–1013.

Bies, R. J., & Shapiro, D. L. (1987). Interactional fairness judgments: The influence of causal accounts. *Social Justice Research*, *1*(2), 199–218.

Binns, R., Van Kleek, M., Veale, M., Lyngs, U., Zhao, J., & Shadbolt, N. (2018). *It's reducing a human being to a percentage: Perceptions of justice in algorithmic decisions*. Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems, Paper *377*, 1–14.

Blut, M., Wang, C., Wünderlich, N. V., & Brock, C. (2021). Understanding anthropomorphism in service provision: A meta-analysis of physical robots, chatbots, and other AI. *Journal of the Academy of Marketing Science*, *49*(4), 632–658.

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

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

Carr, C. L. (2007). The FAIRSERV model: Consumer reactions to services based on a multidimensional evaluation of service fairness. *Decision Sciences*, *38*(1), 107–130.

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

Castelvecchi, D. (2016). Can we open the black box of AI? *Nature*, *538*(7623), 20–23.

Chapman, D. S., Uggerslev, K. L., & Webster, J. (2003). Applicant reactions to face-to-face and technology-mediated interviews: A field investigation. *Journal of Applied Psychology*, *88*(5), 944–953.

Chen, G., Gully, S. M., & Eden, D. (2001). Validation of a new general self-efficacy scale. *Organizational Research Methods*, *4*(1), 62–83.

Cheryan, S., & Markus, H. R. (2020). Masculine defaults: Identifying and mitigating hidden cultural biases. *Psychological Review*, *127*(6), 1022–1052.

Cohen-Charash, Y., & Spector, P. E. (2001). The role of justice in organizations: A meta-analysis. *Organizational Behavior and Human Decision Processes*, *86*(2), 278–321.

Colquitt, J. A. (2001). On the dimensionality of organizational justice: A construct validation of a measure. *Journal of Applied Psychology*, *86*(3), 386–400.

Colquitt, J. A., Wesson, M. J., Porter, C. O., Conlon, D. E., & Ng, K. Y. (2001). Justice at the millennium: A meta-analytic review of 25 years of organizational justice research. *Journal of Applied Psychology*, *86*(3), 425–445.

Colquitt, J. A., & Zipay, K. P. (2015). Justice, fairness, and employee reactions. *Annual Review of Organizational Psychology and Organizational Behavior*, *2*(1), 75–99.

Crolic, C., Thomaz, F., Hadi, R., & Stephen, A. T. (2022). Blame the bot: Anthropomorphism and anger in customer–chatbot interactions. *Journal of Marketing*, *86*(1), 132–148.

Cropanzano, R., Bowen, D. E., & Gilliland, S. W. (2007). The management of organizational justice. *Academy of Management Perspectives*, *21*(4), 34–48.

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

De-Arteaga, M., Feuerriegel, S., & Saar-Tsechansky, M. (2022). Algorithmic fairness in business analytics: Directions for research and practice. *arXiv:2207.10991 [cs.AI]*.

Di Stasio, V., & Heath, A. (2020). *Are employers in Britain discriminating against ethnic minorities? Technical Report 649255*. The Migration Observatory.

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

Dodge, J., Liao, Q. V., Zhang, Y., Bellamy, R. K. E., & Dugan, C. (2019). Explaining models: An empirical study of how explanations impact fairness judgment. In *Proceedings of the 24th international conference on intelligent user interfaces* (pp. 275–285). ACM.

Dolata, M., Feuerriegel, S., & Schwabe, G. (2022). A sociotechnical view of algorithmic fairness. *Information Systems Journal*, *32*(4), 754–818.

Doshi-Velez, F., & Kim, B. (2017). Towards a rigorous science of interpretable machine learning. *arXiv:1702.08608 [stat.ML]*.

Dressel, J., & Farid, H. (2018). The accuracy, fairness, and limits of predicting recidivism. *Science Advances*, *4*(1), eaao5580.

Dworkin, J. D., Linn, K. A., Teich, E. G., Zurn, P., Shinohara, R. T., & Bassett, D. S. (2020). The extent and drivers of gender imbalance in neuroscience reference lists. *Nature Neuroscience*, *23*(8), 918–926.

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

Eurostat. (2022). *Positions held by women in senior management positions*.

Felzmann, H., Villaronga, E. F., Lutz, C., & Tamò-Larrieux, A. (2019). Transparency you can trust: Transparency requirements for artificial intelligence between legal norms and contextual concerns. *Big Data & Society*, *6*(1), 1–14.

Feuerriegel, S., Dolata, M., & Schwabe, G. (2020). Fair AI: Challenges and opportunities. *Business & Information Systems Engineering*, *62*(4), 379–384.

Fumagalli, E., Rezaei, S., & Salomons, A. (2022). OK computer: Worker perceptions of algorithmic recruitment. *Research Policy*, *51*(2), 104420.

Gilliland, S. W. (1993). The perceived fairness of selection systems: An organizational justice perspective. *The Academy of Management Review*, *18*(4), 42.

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

Gönül, S., Önkal, D., & Goodwin, P. (2009). Expectations, use and judgmental adjustment of external financial and economic forecasts: An empirical investigation. *Journal of Forecasting*, *28*(1), 19–37.

Greenaway, K. E., Chan, Y. E., & Crossler, R. E. (2015). Company information privacy orientation: A conceptual framework: Company information privacy orientation. *Information Systems Journal*, *25*(6), 579–606.

Han, N. R., Baek, T. H., Yoon, S., & Kim, Y. (2019). Is that coffee mug smiling at me? How anthropomorphism impacts the effectiveness of desirability vs. feasibility appeals in sustainability advertising. *Journal of Retailing and Consumer Services*, *51*, 352–361.

Hausknecht, J. P. (2004). Applicant reactions to selection procedures: An updated model and meta-analysis. *Personell Psychology*, *57*, 639–683.

Hong, W., Chan, F. K. Y., Thong, J. Y. L., Chasalow, L. C., & Dhillon, G. (2014). A framework and guidelines for context-specific theorizing in information systems research. *Information Systems Research*, *25*(1), 111–136.

Jacoby, J. (2002). Stimulus-organism-response reconsidered: An evolutionary step in modeling (consumer) behavior. *Journal of Consumer Psychology*, *12*(1), 51–57.

Judge, T. A., & Colquitt, J. A. (2004). Organizational justice and stress: The mediating role of work-family conflict. *Journal of Applied Psychology*, *89*(3), 395–404.

Jussupow, E., Spohrer, K., Heinzl, A., & Gawlitza, J. (2021). Augmenting medical diagnosis decisions? An investigation into physicians' decision-making process with artificial intelligence. *Information Systems Research*, *32*(3), 713–735.

Keith, M. G., Harms, P., & Tay, L. (2019). Mechanical Turk and the gig economy: Exploring differences between gig workers. *Journal of Managerial Psychology*, *34*(4), 286–306.

Kizilcec, R. F. (2016). How much information? Effects of transparency on trust in an algorithmic interface. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (pp. 2390–2395). ACM.

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

Kordzadeh, N., & Ghasemaghaei, M. (2022). Algorithmic bias: Review, synthesis, and future research directions. *European Journal of Information Systems*, *31*(3), 388–409.

Kunin, T. (1955). The construction of a new type of attitude measure. *Personnel Psychology*, *8*(1), 65–77.

Kwak, H., Puzakova, M., & Rocereto, J. F. (2017). When brand anthropomorphism alters perceptions of justice: The moderating role of self-construal. *International Journal of Research in Marketing*, *34*(4), 851–871.

Langer, M., König, C. J., & Krause, K. (2017). Examining digital interviews for personnel selection: Applicant reactions and interviewer ratings. *International Journal of Selection and Assessment*, *25*(4), 371–382.

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

Leventhal, G. S. (1980). What should be done with equity theory? In K. J. Gergen, M. S. Greenberg, & R. H. Willis (Eds.), *Social exchange* (pp. 27–55). Springer US.

Li, H., Sarathy, R., Zhang, J., & Luo, X. (2014). Exploring the effects of organizational justice, personal ethics and sanction on internet use policy compliance: Determinants of IUP compliance. *Information Systems Journal*, *24*(6), 479–502.

Liew, T. W., & Tan, S.-M. (2021). Social cues and implications for designing expert and competent artificial agents: A systematic review. *Telematics and Informatics*, *65*, 101721.

Lin, J., Yin, J., Wei, K. K., Chan, H. C., & Teo, H. H. (2021). Comparing competing systems: An extension of the information systems continuance model. *MIS Quarterly*, *46*(4), 1851–1874.

Lukacik, E.-R., Bourdage, J. S., & Roulin, N. (2022). Into the void: A conceptual model and research agenda for the design and use of asynchronous video interviews. *Human Resource Management Review*, *32*(1), 100789.

Maedche, A., Legner, C., Benlian, A., Berger, B., Gimpel, H., Hess, T., Hinz, O., Morana, S., & Söllner, M. (2019). AI-based digital assistants: Opportunities, threats, and research perspectives. *Business & Information Systems Engineering*, *61*(4), 535–544.

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

Martin, K. D., Borah, A., & Palmatier, R. W. (2017). Data privacy: Effects on customer and firm performance. *Journal of Marketing*, *81*(1), 36–58.

Masterson, S. S., Lewis, K., Goldman, B. M., & Taylor, M. S. (2000). Integrating justice and social exchange: The differing effects of fair procedures and treatment on work relationships. *Academy of Management Journal*, *43*(4), 738–748.

McAuliffe, K., Blake, P. R., Steinbeis, N., & Warneken, F. (2017). The developmental foundations of human fairness. *Nature Human Behaviour*, *1*, 0042.

McCarthy, J. M., Bauer, T. N., Truxillo, D. M., Anderson, N. R., Costa, A. C., & Ahmed, S. M. (2017). Applicant perspectives during selection: A review addressing "so what?," "What's new?," and "where to next?". *Journal of Management*, *43*(6), 1693–1725.

Mehrabian, A., & Russell, J. A. (1974). *An approach to environmental psychology*. M.I.T. Press.

Miller, F. D. (1996). Aristotle and the origins of natural rights. *The Review of Metaphysics*, *49*(4), 873–907.

Mirchandani, D. A., & Lederer, A. L. (2014). Autonomy and procedural justice in strategic systems planning: Autonomy and procedural justice in strategic systems planning. *Information Systems Journal*, *24*(1), 29–59.

Moon, H., & Conlon, D. E. (2002). From acclaim to blame: Evidence of a person sensitivity decision bias. *Journal of Applied Psychology*, *87*(1), 33–42.

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

Morse, L., Teodorescu, M. H. M., Awwad, Y., & Kane, G. C. (2021). Do the ends justify the means? Variation in the distributive and procedural fairness of machine learning algorithms. *Journal of Business Ethics*, *181*, 1083–1095.

Moss-Racusin, C. A., Dovidio, J. F., Brescoll, V. L., Graham, M. J., & Handelsman, J. (2012). Science faculty's subtle gender biases favor male students. *Proceedings of the National Academy of Sciences*, *109*(41), 16474–16479.

Nass, C., & Moon, Y. (2000). Machines and mindlessness: Social responses to computers. *Journal of Social Issues*, *56*(1), 81–103.

Newlands, G. (2021). Algorithmic surveillance in the gig economy: The organization of work through Lefebvrian conceived space. *Organization Studies*, *42*(5), 719–737.

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

Obermeyer, Z., Powers, B., Vogeli, C., & Mullainathan, S. (2019). Dissecting racial bias in an algorithm used to manage the health of populations. *Science*, *366*(6464), 447–453.

Oetting, S. K., & Maier, G. W. (2018). The importance of procedural justice in human–machine interactions: Intelligent systems as new decision agents in organizations. *Computers in Human Behavior*, *89*, 27–39.

Peer, E., Rothschild, D., Gordon, A., Evernden, Z., & Damer, E. (2021). Data quality of platforms and panels for online behavioral research. *Behavior Research Methods*, *54*(4), 1643–1662.

Pfeuffer, N., Benlian, A., Gimpel, H., & Hinz, O. (2019). Anthropomorphic information systems. *Business & Information Systems Engineering*, *61*(4), 523–533.

Ployhart, R. E., & Ryan, A. M. (1997). Toward an explanation of applicant reactions: An examination of organizational justice and attribution frameworks. *Organizational Behavior and Human Decision Processes*, *72*(3), 308–335.

Puzakova, M., Kwak, H., & Rocereto, J. F. (2013). When humanizing brands goes wrong: The detrimental effect of brand anthropomorphization amid product wrongdoings. *Journal of Marketing*, *77*(3), 81–100.

Qiu, L., & Benbasat, I. (2009). Evaluating anthropomorphic product recommendation agents: A social relationship perspective to designing information systems. *Journal of Management Information Systems*, *25*(4), 145–182.

Quillian, L., Pager, D., Hexel, O., & Midtbøen, A. H. (2017). Meta-analysis of field experiments shows no change in racial discrimination in hiring over time. *Proceedings of the National Academy of Sciences of the United States of America*, *114*(41), 10870–10875.

Robert, L. P., Pierce, C., Marquis, L., Kim, S., & Alahmad, R. (2020). Designing fair AI for managing employees in organizations: A review, critique, and design agenda. *Human–Computer Interaction*, *35*(5–6), 545–575.

Saxena, N. A., Huang, K., DeFilippis, E., Radanovic, G., Parkes, D. C., & Liu, Y. (2020). How do fairness definitions fare? Testing public attitudes towards three algorithmic definitions of fairness in loan allocations. *Artificial Intelligence*, *283*, 103238.

Schanke, S., Burtch, G., & Ray, G. (2021). Estimating the impact of "humanizing" customer service chatbots. *Information Systems Research*, *32*(3), 736–751.
Schinkel, S., van Vianen, A. E. M., & Marie Ryan, A. (2016). Applicant reactions to selection events: Four studies into the role of attributional style and fairness perceptions: Attributional style and fairness perceptions. *International Journal of Selection and Assessment*, *24*(2), 107–118.
Schuler, H. (1993). *Social validity of selection situations: A concept and some empirical results*. Lawrence Erlbaum Associates, Inc.
Seeger, A.-M., Pfeiffer, J., & Heinzl, A. (2021). Texting with humanlike conversational agents: Designing for anthropomorphism. *Journal of the Association for Information Systems*, *22*(4), 931–967.
Shin, D., & Park, Y. J. (2019). Role of fairness, accountability, and transparency in algorithmic affordance. *Computers in Human Behavior*, *98*, 277–284.
Smith, M., & Robertson, I. (1989). *Advances in selection and assessment*. John Wiley & Sons.
Sonboli, N., Smith, J. J., Cabral Berenfus, F., Burke, R., & Fiesler, C. (2021). Fairness and transparency in recommendation: The users' perspective. In *Proceedings of the 29th ACM conference on user modeling, adaptation and personalization* (pp. 274–279). ACM.
Souza, M. (2020). *An introduction to unconscious bias – From industry experts*. https://www.brighttalk.com/webcast/17947/438643
Starke, C., Baleis, J., Keller, B., & Marcinkowski, F. (2021). Fairness perceptions of algorithmic decision-making: A systematic review of the empirical literature. *arXiv:2103.12016 [cs]*.
Suen, H.-Y., Chen, M. Y.-C., & Lu, S.-H. (2019). Does the use of synchrony and artificial intelligence in video interviews affect interview ratings and applicant attitudes? *Computers in Human Behavior*, *98*, 93–101.
Thielsch, M. T., Träumer, L., & Pytlik, L. (2012). E-recruiting and fairness: The applicant's point of view. *Information Technology and Management*, *13*(2), 59–67.
van den Broek, E., Sergeeva, A., & Huysman Vrije, M. (2021). When the machine meets the expert: An ethnography of developing AI for hiring. *MIS Quarterly*, *45*(3), 1557–1580.
Walker, H. J., Bauer, T. N., Cole, M. S., Bernerth, J. B., Feild, H. S., & Short, J. C. (2013). Is this how I will be treated? Reducing uncertainty through recruitment interactions. *Academy of Management Journal*, *56*(5), 1325–1347.
Wang, D., Yang, Q., Abdul, A., & Lim, B. Y. (2019). Designing theory-driven user-centric explainable AI. In *Proceedings of the 2019 CHI conference on human factors in computing systems CHI '19* (pp. 1–15). ACM Press.
Waytz, A., Cacioppo, J., & Epley, N. (2010). Who sees human? The stability and importance of individual differences in anthropomorphism. *Perspectives on Psychological Science*, *5*(3), 219–232.
Wei, S., Yin, J., & Chen, X. (2021). Paradox of supply chain integration and firm performance: The moderating roles of distributive and procedural justice. *Decision Sciences*, *52*(1), 78–108.
Willison, R., Warkentin, M., & Johnston, A. C. (2018). Examining employee computer abuse intentions: Insights from justice, deterrence and neutralization perspectives: Examining the influence of disgruntlement on computer abuse intentions. *Information Systems Journal*, *28*(2), 266–293.
Yalcin, G., Lim, S., Puntoni, S., & van Osselaer, S. M. (2022). Thumbs up or down: Consumer reactions to decisions by algorithms versus humans. *Journal of Marketing Research*, *59*(4), 696–717.
Ybarra, O. (2002). Naive causal understanding of valenced behaviors and its implications for social information processing. *Psychological Bulletin*, *128*(3), 421–441.
Yeomans, M., Shah, A., Mullainathan, S., & Kleinberg, J. (2019). Making sense of recommendations. *Journal of Behavioral Decision Making*, *32*(4), 403–414.
You, S., Yang, C. L., & Li, X. (2022). Algorithmic versus human advice: Does presenting prediction performance matter for algorithm appreciation? *Journal of Management Information Systems*, *39*(2), 336–365.
Zhu, J., Liapis, A., Risi, S., Bidarra, R., & Youngblood, M. (2018). Explainable AI for designers: A human-centered perspective on mixed-initiative co-creation. *Proceedings of the 2018 IEEE Conference on Computational Intelligence and Games (CIG'18)*.

AUTHOR BIOGRAPHIES

Jessica Ochmann (jessica.ochmann@fau.de) holds a PhD in Information Systems from Friedrich-Alexander-Universität Erlangen-Nürnberg. Her research interest lies in the use of algorithms in human decision-making. She investigates the behavioural consequences of generative technologies and aims to unlock the potential of interventions like anthropomorphism and transparency to make algorithms comprehensible and accessible for

humans. Thereby, she is especially interested in finding ways to introduce ethical and reflective algorithmic decision-making.

Leonard Michels (leonard.michels@fau.de) obtained his PhD in Information Systems at Friedrich-Alexander-Universität Erlangen-Nürnberg. His research centres on the behavioural effects of information technology, particularly focusing on digital choice architecture interventions. Employing a multidisciplinary approach that draws from behavioural economics, psychology, and information systems research, he leverages quantitative, experimental data to drive and assess behaviour change.

Verena Tiefenbeck (verena.tiefenbeck@fau.de) is the assistant professor for Digital Transformation and research group leader at the Friedrich-Alexander University Erlangen-Nuremberg, Germany. Prior to that, she led the Bits to Energy Lab at ETH Zurich, Switzerland, and held a part-time position as postdoctoral researcher at the Institute of Applied Microeconomics at the University of Bonn. Her empirical research combines digital technologies, data analytics, and behavioural research to foster behaviour change, with a focus on the sustainability and health context. Verena completed her Ph.D. at ETH Zurich, Switzerland after 3.5 years spent in Boston, USA, as visiting Ph.D. student at the Massachusetts Institute of Technology and as research assistant at the Fraunhofer Center for Sustainable Energy Systems. She holds a double-degree in Mechanical Engineering and Management from TU Munich and Ecole Centrale Paris. Her work has been published in top journals (incl. Management Science, ISR, Nature Energy, EJIS, Energy Policy, Global Environmental Change).

Christian Maier (christian.maier@uni-bamberg.de) is a professor at the University of Bamberg in Germany. His research interests include the IS use life cycle, especially the adoption, usage, and discontinuous usage of digital technologies in the private and organisational use contexts. His research has been published, among others, in MIS Quarterly and Information Systems Research. He was awarded the Schmalenbach prize, the Early Career Awards by the AIS and the ACM SIGMIS, and the Heinz Maier-Leibnitz price.

Sven Laumer (sven.laumer@fau.de) is the Schöller Endowed Professor and Chair of Information Systems in the School of Business, Economics and Society at Friedrich-Alexander-Universität Erlangen-Nürnberg, Head of the Dr. Theo and Friedl Schöller Research Lab 'Future of Work', and Director of the Dr. Theo und Friedl Schöller Research Center for Business and Society. His research focuses on digital transformation, responsible AI and the future of work. The results of his research have been published in all journals of the AIS Senior Scholars' Basket of Eight (incl. MISQ, ISR, JMIS). He serves on editorial boards of *Information Systems Journal* and *The DATA BASE for Advances in Information Systems*. He has been awarded the Young Talent Award 2018 by the German Academic Association of Business Research (VHB). When he does not research and teach digital work and life, he enjoys being a soccer referee, hiking in the Alps, and spending time with friends and family.

How to cite this article: Ochmann, J., Michels, L., Tiefenbeck, V., Maier, C., & Laumer, S. (2024). Perceived algorithmic fairness: An empirical study of transparency and anthropomorphism in algorithmic recruiting. *Information Systems Journal*, *34*(2), 384–414. https://doi.org/10.1111/isj.12482

APPENDIX A

A.1 | EXPERIMENTAL MANIPULATIONS AND STIMULI

In the following, we provide an overview of the experimental manipulations and stimuli implemented for the different experimental conditions.

A.1.1. | Decision maker

Human decision maker: In the human decision maker groups, the decision maker was referred to as a 'Skillistic HR employee' or 'our employee', depending on the textual context. No additional information on how the employee would reach their decision was given to the participants.

Algorithmic decision maker: In the algorithmic decision maker groups, the decision maker was referred to as 'Skillistic's algorithm' or 'our algorithm', depending on the textual context. No additional information on how the algorithm would reach its decision was given to the participants.

Transparent algorithmic decision maker: In the transparent algorithmic decision maker groups, the decision maker was referred to as 'Skillistic's algorithm' or 'our algorithm', depending on the textual context.

The participants further received the information that the algorithm would evaluate their application by reviewing their demographic criteria (e.g., age, gender, education), their mindset, and using prior data of successful Skillistic Insights applicants to assess their fit for the special participant pool.

TABLE A1 Measurement scales and items used in the experiment.

Construct	Item
Procedural justice	I was able to express my views and feelings during the procedures used to evaluate my application. I had influence over the [outcome] of my application resulting from the procedures used to evaluate my application. The procedures used to evaluate my application have been applied consistently. The procedures used to evaluate my application have been free of bias. The procedures used to evaluate my application have been based on accurate information. I was able to appeal the [outcome] of my application resulting from the procedures used to evaluate my application. The procedures used to evaluate my application upheld ethical and moral standards.
Distributive justice	The [outcome] of my application reflects the effort I have put into my application form. The [outcome] of my application is appropriate for the answers I have submitted. The [outcome] of my application reflects what I can contribute to Skillistic Insights. The [outcome] of my application is justified given my performance.
Interpersonal justice	[Decisionmaker] treated you in a polite manner. [Decisionmaker] treated you with dignity. [Decisionmaker] treated you with respect. [Decisionmaker] refrained from improper remarks or comments. [Decisionmaker] was candid in the communication with you. [Decisionmaker] explained the procedures thoroughly.
Informational justice	[Decisionmaker's] explanations regarding the procedures were reasonable. [Decisionmaker] communicated details in a timely manner. [Decisionmaker] seemed to tailor the communications to your specific needs.

(Continues)

TABLE A1 (Continued)

Construct	Item
Attitude	What is your overall evaluation of Skillistic Insights? I am pleased with [decisionmaker]. I am contented with [decisionmaker].
Satisfaction	I am delighted with [decisionmaker]. I am satisfied with [decisionmaker].
Recommendation	How likely are you to recommend Skillistic Insights to your friends or other Prolific participants?

Anthropomorphic algorithmic decision maker: In the anthropomorphic algorithmic decision maker groups, the decision maker was referred to as 'Emily, Skillistic's algorithm' or 'Emily, our algorithm, depending on the textual context'. Whenever the algorithm Emily was mentioned, a corresponding avatar of a female human person was added to the page. The avatar was frequently accompanied by speech bubbles, repeating the written information. In addition, multiple chat-like conversations were implemented in the survey, suggesting that participants were interacting with a human-like entity.

No additional information on how the algorithm would reach its decision was given to the participants.

A.1.2. | Decision outcome

Positive decision outcome: Congratulations, [*decision maker*] accepted your application!

We will add you to our special participant pool and inform you about our future surveys with bonus compensation.

Negative decision outcome: We are sorry, [*decision maker*] rejected your application!

We will not add you to our special participant pool and do not inform you about our future surveys with bonus compensation.