Enhancing Perceived Fairness of AI-Based Personnel Selection Procedures: The Role of AI Certification

Romina Liza Kleiner
RWTH Aachen University, kleiner@time.rwth-aachen.de

Nicole Janine Hartwich
RWTH Aachen University, hartwich@time.rwth-aachen.de

David Antons
RWTH Aachen University, antons@time.rwth-aachen.de

Enhancing Perceived Fairness of AI-Based Personnel Selection Procedures: The Role of AI Certification

Completed Research Paper

Romina L. Kleiner, Nicole J. Hartwich, David Antons
Institute for Technology and Innovation Management,
RWTH Aachen University
Kackertstr. 7, 52072 Aachen, Germany
{kleiner, hartwich, antons}@time.rwth-aachen.de

Abstract

Organizations increasingly integrate artificial intelligence (AI) into their selection procedures, for example, to analyze selection interviews. However, prior research indicates that applicants react negatively to an AI as decision agent in selection interviews. Only few studies have addressed possible mechanisms to mitigate negative effects of AI usage on applicant reactions, calling for increased research attention. The present experimental study (N=180) is among the first to test whether AI certification has the potential to positively influence people's fairness perceptions of AI usage. Consistent with previous research, our results show that people are less attracted to an organization that conducts AI-based (compared to human-based) selection interviews, partially due to lower procedural fairness perceptions. With regards to AI certification, we highlight the crucial role of the certification label's perceived trustworthiness in shaping applicants' procedural fairness perceptions and organizational attraction. Future research should investigate factors that enhance the trustworthiness of an AI certification label.

Keywords: AI-based selection, AI certification, perceived fairness, organizational attraction

Introduction

Artificial Intelligence (AI) is increasingly being used by human resource (HR) managers (Hmoud & Laszlo, 2019), ultimately affecting millions of applicants every year. A recent study by McKinsey identified the automation of HR processes and decisions as an increasingly important trend in the field of HR (Durth et al., 2022). With the implementation of technology-based selection tools, HR managers pursue the goals of increasing efficiency, reducing costs, standardizing systems, expanding the application pool, and promoting the organizational image (Chapman & Webster, 2003). AI has been proposed to be useful in three phases of the recruitment process, namely sourcing human resources, screening and short-listing, and selection (Hmoud & Laszlo, 2019). With regards to the selection phase, highly automated interviews gain more attention (Langer et al., 2019). Within these interviews, AI is used for collecting audio and video data, evaluating verbal and nonverbal behavior, choosing follow-up questions as well as ranking applicants (Langer et al., 2019). Some automated interview tools are already on the market, for example, HireVue or myInterview Intelligence. These tools analyze nonverbal behavior, choice of words, speaking patterns, and tone in order to draw conclusions about personality traits (Hmoud & Laszlo, 2019). One commonly used type of automated interviews is the digital interview, during which applicants record their interview responses, send them to the organization, and are automatically evaluated at a later point in time (Langer, König, Sanchez, et al., 2020).

Besides the numerous chances, major challenges are associated with the use of AI in personnel selection procedures. The example of Amazon's attempt to use an automated personnel selection tool that ended up discriminating against women illustrates that algorithms tend to replicate human biases and react strongly to unrepresentative and biased training data (Dastin, 2022). Another major problem that was identified by prior research is negative applicant reactions towards AI-based selection tools (e.g., Acikgoz et al., 2020; Langer, König, Sanchez, et al., 2020). In times of war for talent, it is crucial for organizations to keep qualified applicants attracted throughout the whole personnel selection procedure to increase the likelihood that they accept a job offer after the procedure (Chapman et al., 2005).

One influential dimension of applicant reactions is the perceived fairness of personnel selection procedures (Gilliland, 1993). Perceived fairness has been researched from an organizational justice perspective that distinguishes between distributive, procedural, and interactional fairness (Colquitt et al., 2005). Distributive fairness refers to the fairness of outcomes, procedural fairness to the fairness of decision-making procedures, and interactional fairness to the fairness of interpersonal treatment. So far, research on AI fairness in the organizational context has mainly focused on distributive fairness (Robert et al., 2020). To address the lacking research focus, we specifically look at procedural fairness perceptions in the present study. With regards to personnel selection procedures, Gilliland (1993) proposed ten procedural justice rules that people use to evaluate the fairness of selection procedures. These rules include, for example, the job-relatedness of the procedure and the opportunity for applicants to perform and show their skills. Procedural fairness perceptions have been shown to positively influence a variety of outcomes, for example, applicants' job pursuit intentions and job-organization attraction (Chapman et al., 2005). Compared to the other two fairness facets, procedural fairness was proposed to have the strongest impact on people's general fairness judgment (van den Bos et al., 2001) which in turn is expected to influence relevant attitudes like organizational attraction (Gilliland, 1993). With regards to AI usage, it might be crucial for organizations to understand (a) whether applicants perceive the personnel selection procedure as less procedurally fair when the decision agent is an AI compared with a human, and (b) whether there is a mechanism to mitigate potential negative reactions.

Prior research mostly points to negative applicant reactions towards AI usage in personnel selection procedures (e.g., Acikgoz et al., 2020; Langer, König, Sanchez, et al., 2020; Mirowska, 2020). When people believe that they are evaluated by an AI instead of a human during a selection interview, they are less attracted to the organization (Langer, König, Sanchez, et al., 2020) and show fewer intentions to apply for and pursue a job (Mirowska, 2020). Furthermore, automated personnel selection procedures result in more privacy concerns, less perceived controllability, and more creepy ambiguity (Langer et al., 2019). With regards to fairness perceptions, findings are rather mixed. While some studies report a negative effect of automated HR decisions on fairness perceptions (Acikgoz et al., 2020; Newman et al., 2020), other studies find no difference compared to traditional procedures conducted by a human (e.g., Langer, König, & Hemsing, 2020). One possible explanation for the mixed findings is that fairness perceptions were not measured consistently. Research that specifically looked at procedural fairness mostly found fairness perceptions to be lower when HR decisions are made by an AI instead of a human (e.g., Acikgoz et al., 2020; Newman et al., 2020). One empirically tested reason for lower ratings of procedural fairness is that people do not believe that an AI system has the capability to capture all relevant information and evaluate human skills accurately (Newman et al., 2020).

Despite the large number of studies that have shown negative effects of AI-based personnel selection procedures on applicants' fairness perceptions and organizational attraction, only a few studies have researched possible ways to mitigate these negative effects. In a recent review of organizational justice literature, Colquitt et al. (2023) highlighted the relevance to further investigate how transparency about AI influences fairness perceptions in the context of human resources decisions. Some research has examined the role of information about the algorithm as a strategy to improve applicant reactions towards AI usage. The results so far are inconclusive. Contrary to the assumption that more information about an algorithm positively affects reactions to automated HR decisions, Newman et al. (2020) found negative effects of providing more information about an algorithm on procedural fairness perceptions. While Langer et al. (2018) found a positive indirect effect of information about the algorithm on organizational attraction through the fairness dimensions open treatment and information known, they found a negative direct effect on organizational attraction. They concluded that information can act as a double-edged sword and that it might be of relevance what information is communicated to applicants.

One specific type of information to improve applicants' fairness perceptions of AI usage that has not been tested yet is AI certification. AI certification is currently discussed intensively by experts from business, research, society, and politics as one approach to fostering responsible use and enhancing the trustworthiness and acceptance of AI usage (Poretschkin et al., 2023). A certification that is issued by an independent party attests the adherence to standards and regulations that are currently under development for AI applications (Poretschkin et al., 2023). On 21 April 2021, the European Commission (2021) published the first draft of the AI Act which proposes a legal framework for AI usage and resulted in many national projects on AI certification, for example, the German AI.NRW flagship project "Certified AI".

Our study fills a current gap in research and tests assumptions about the positive effect of AI certification on people's perceptions and attitudes regarding AI usage in a highly relevant context, namely personnel selection procedures. Thereby, we contribute to the current debate on factors that could enhance the trustworthiness and acceptance of AI usage (Poretschkin et al., 2023). We conducted an experimental study in which 180 participants completed a simulated digital interview procedure and were either told that they were evaluated by an AI or by a human. Furthermore, either a certification label of the decision agent (AI vs. human) together with information about the certification was displayed or no information about a certification was given. After the interview procedure, participants rated the procedural fairness of the procedure and indicated how attracted they are to the organization that conducts this procedure. The experimental design allowed us to test causal relationships within a controlled setting.

With this study, we contribute to previous research in at least two meaningful ways. First, we advance the understanding of people's reactions to AI-based selection interviews using a novel experimental design that aims at simulating a realistic interview procedure and inducing involvement. We replicate and strengthen previous findings regarding the negative effect of AI-based selection interviews on applicants' procedural fairness perceptions and organizational attraction. Second, we highlight the important role of the perceived trustworthiness of an AI certification label in influencing people's perceptions and attitudes towards AI usage. Thereby, we identify a boundary condition for fairness heuristic theory, since our results suggest that a fairness signal has to be perceived as trustworthy to have a positive effect on people's general fairness judgment. This finding calls for increased research attention on factors that have an impact on the perceived trustworthiness of an AI certification label.

Conceptual Background and Hypotheses

The present study builds on three theories. First, signaling theory describes how signals, like AI usage and certifications, can be used by organizations to communicate unobservable qualities, thereby influencing people's attitudes. Second, fairness heuristic theory explains how applicants form their fairness judgment based on fairness signals like AI usage and certifications. Third, uncertainty management theory enables a deeper understanding of the circumstances under which people are most attentive towards fairness signals.

Signaling Theory

Signaling theory (Spence, 1973) is concerned with reducing information asymmetries between two parties by the use of signals (Connelly et al., 2011). Many markets are characterized by informational gaps, implying that information about an individual, organization, or product is not available to a person who would benefit from this information with regards to decision-making (Connelly et al., 2011). For example, applicants often lack information about organizations during personnel selection procedures when this information would be useful for them (Osburg et al., 2020). In order to overcome that information asymmetry, the party that possesses the relevant information (e.g., the organization) can use signals to communicate unobservable qualities to the party that would like to receive that information (e.g., the individual). Signals are actions that have the intention to communicate positive, unobservable qualities of the signaler to a receiver (Connelly et al., 2011). The receiver will then interpret the signal and draw inferences from it that can influence his or her decision-making. An example of an organizational signal that is used to signal quality and adherence to predefined standards is certification (Kimery & McCord, 2006). By obtaining certifications, organizations hope to influence organizational stakeholders' attitudes and behaviors, as well as to differentiate themselves from competitors (Kimery & McCord, 2006). Several studies on certifications in different contexts have built their hypotheses on signaling theory. For example, ethical certifications were tested as a signal for ethicality and good working conditions, resulting in higher talent attraction (Osburg et al., 2020). Within the recruiting context, it was found that the recruiter's behavior can represent a signal for unobservable organizational attributes (Rynes et al., 1991). When an AI instead of a human analyzes the selection interview, this might act as a signal and lead to inferences about the organization by the applicant. Organizations could aim to signal innovativeness, efficiency, and objectivity by conducting AI-based interviews (Mirowska, 2020). However, since signals are not always interpreted as intended, applicants could also interpret AI-based interviews as a lack of interest and appreciation (Acikgoz et al., 2020).

Fairness Heuristic Theory and Uncertainty Management Theory

Fairness heuristic theory (Lind, 2001; van den Bos et al., 2001) aims to explain when, why, and how people form fairness judgments in the workplace, and how they use these judgments to guide their attitudes, decisions, and behavior. With regards to the question of when and why fairness judgments are formed, the theory refers to the fundamental social dilemma that describes the decision of whether to engage in a group (e.g., an organization) or not (Lind, 2001). On the one hand, engaging as a group member leads to material rewards, while on the other hand, it involves the risk of being exploited. Furthermore, group membership strengthens the personal identity, but it also involves the risk of being excluded. Applicants face this dilemma when they have to decide whether to become part of the organization or not. Since these decisions are very complex, available information might be limited, and outcome contingencies are difficult to estimate, fairness heuristic theory claims that the fundamental social dilemma is resolved by using fairness perceptions as a decision heuristic (van den Bos et al., 2001). Especially when the trustworthiness of an organization is unknown, people become very attentive to fairness information and use it as a substitute for trust to guide decision-making (Lind, 2001). Turning to the question of how people form and use fairness judgments, fairness heuristic theory suggests that people form a general fairness judgment based on initial, available fairness signals and afterwards use it as a cognitive heuristic to interpret organizational experiences and guide behavior (Lind, 2001). New fairness information is assimilated to the relatively stable general fairness judgment, implying that early fairness information receives greater weight than later fairness information (Lind, 2001). With regards to AI-based selection interviews, applicants could interpret AI usage as an early fairness signal that leads to a negative fairness judgment and subsequently affects attitudes towards the organization. Certification, on the other hand, could act as an early positive fairness signal and improve the fairness judgment, leading to less negative attitudes towards the organization.

Uncertainty management theory (Lind & van den Bos, 2002) builds on fairness heuristic theory and extends it with regards to the question of when fairness signals matter most for the general fairness judgment and subsequent attitudes. The theory claims that a general fairness judgment is used as a psychological tool to manage uncertainty and can provide guidance on how to feel and behave in an uncertain (organizational) context (Lind & van den Bos, 2002). High uncertainty should enhance the attention and sensitivity towards fairness signals in order to quickly arrive at a certain fairness judgment that can guide attitudes and behavior (Lind & van den Bos, 2002). Uncertainty management theory can shed light on the circumstances under which fairness signals like AI usage or certifications might have the highest impact on fairness perceptions and organizational attraction. In personnel selection procedures, applicants often experience uncertainty as they usually have little knowledge about the organization which might lead them to pay particular attention to fairness signals like AI usage and certifications.

In the following, we will derive our hypotheses from the presented theories. Figure 1 provides an overview of the proposed research model.

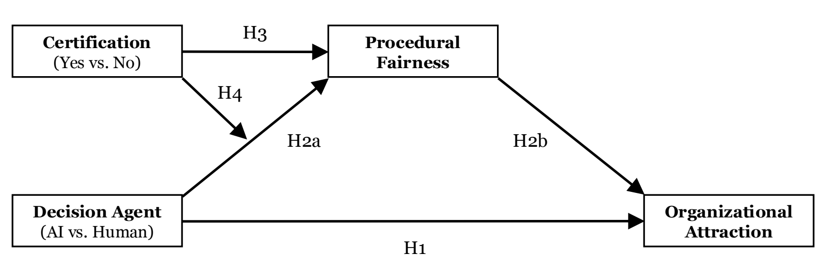

Figure 1. Proposed Research Model

The Effect of AI as Decision Agent on Organizational Attraction

Following signaling theory, applicants draw inferences about unknown organizational characteristics from signals that they receive during the personnel selection procedure (Rynes et al., 1991). These signals include, for example, recruiter and process characteristics and can influence applicant reactions and decisions (Rynes et al., 1991). By implementing AI-based selection procedures, organizations could aim to signal innovativeness, efficiency, and objectivity and anticipate favorable applicant reactions (Mirowska, 2020). This is in line with the results of a qualitative study by Chapman and Webster (2003) who identified promoting the organizational image as one motivation of HR managers for adopting technology-based selection tools. However, signals are sometimes interpreted differently than intended. Instead of interpreting AI-based selection interviews as innovative, applicants could infer that the organization does not value them as much as other organizations that conduct human-based interviews (Acikgoz et al., 2020). AI usage could be interpreted as a signal that the organization is more interested in saving time and money than in investing in applicants which will shape applicants' expectations of the organization. The perceived lack of appreciation might lead applicants to interpret this as a characteristic of the organization, resulting in less organizational attraction.

Prior research has mainly shown negative effects of AI-based selection procedures on organizational attraction (Acikgoz et al., 2020; Langer, König, Sanchez, et al., 2020). For example, participants were less attracted to an organization that conducts highly automated interviews compared with videoconference interviews (Langer, König, Sanchez, et al., 2020). Furthermore, participants reported lower intentions to apply when they were told that an AI would evaluate their recorded interview answers within a personnel selection procedure (Mirowska, 2020). In line with signaling theory and these findings, we predict that participants who are told that an AI analyzes their digital selection interview will infer negative organizational attributes and feel less attracted to the organization. We derive the following hypothesis:

Hypothesis 1: Applicants rate organizational attraction lower when the decision agent within the selection interview is an AI compared with a human.

The Mediating Effect of Procedural Fairness

Whether the decision agent within a personnel selection procedure is an AI or a human is not only expected to affect organizational attraction directly but also indirectly through procedural fairness perceptions. Fairness heuristic theory can provide a theoretical foundation for this assumption. The theory implies that the choice of decision agent (AI vs. human) by an organization can serve as an early fairness signal and therefore has a great impact on applicants' general fairness judgment of the personnel selection procedure. To quickly arrive at a fairness heuristic that can guide attitudes and behavior, applicants might rely on early and available fairness signals (i.e., choice of decision agent) to form their general fairness judgment.

Several studies found that AI-based personnel selection procedures are generally perceived as less fair than traditional procedures that involve a human (e.g., Acikgoz et al., 2020; Langer et al., 2019); however, other studies did not find that effect (e.g., Langer, König, & Hemsing, 2020). One explanation why an AI as decision agent in HR decisions could be perceived as less fair was suggested and successfully tested by Newman et al. (2020). They found that people do not believe that an AI can capture all relevant information and evaluate human skills accurately, leading people to perceive the process as reductionistic. With regards to procedural fairness, in particular, Acikgoz et al. (2020) showed mainly negative effects of AI-based selection interviews. Consistent with this finding, studies that looked at one subdimension of procedural fairness (i.e., change to perform) also found lower applicant ratings when they were evaluated by an AI compared with a human (e.g., Langer, König, & Hemsing, 2020). Therefore, we propose the following hypothesis:

Hypothesis 2a: Applicants perceive the personnel selection procedure as less procedurally fair when the decision agent within the selection interview is an AI compared with a human.

Fairness heuristic theory predicts that the general fairness judgment that people form acts as a heuristic for guiding attitudes, decisions, and behaviors in an otherwise uncertain organizational situation (van den Bos et al., 2001). In personnel selection procedures in which little is known about the organization, applicants are expected to rely on their fairness judgment to decide how to feel, how to behave, and what to think about the organization. A positive general fairness judgment might reduce fears of being exploited or excluded by an organization, leading to a cooperative orientation towards the organization (Lind, 2001). This cooperative orientation fosters the willingness to become part of an organization in order to benefit from the advantages of engaging in a group. A negative fairness judgment, on the other hand, is expected to amplify the fears of being exploited and excluded, thereby leading to lower organizational attraction.

Studies have shown that fairness perceptions of personnel selection procedures can influence applicants' attitudes, intentions, and behavior towards an organization (Chapman et al., 2005, for a meta-analytic review). Especially the effect of fairness perceptions on organizational attraction has received great attention and most studies have shown a positive relationship between the two constructs (e.g., Acikgoz et al., 2020; Bauer et al., 2001). Studies that have specifically examined procedural fairness (Acikgoz et al., 2020; Chapman et al., 2005) also found significant effects on organizational attraction. Thus, we derive the following hypothesis:

Hypothesis 2b: Higher procedural fairness perceptions of the personnel selection procedure are associated with higher organizational attraction.

The Effect of Certification on Procedural Fairness

Not only AI usage but also certification might act as a signal within personnel selection procedures. Building on signaling theory, research has shown that information received before the personnel selection procedure serves as a signal for organizational characteristics and affects applicants' attitudes towards the company (Osburg et al., 2020). One way to signal positive organizational characteristics and influence applicants' perceptions is certification (Kimery & McCord, 2006). The information that the decision agent (AI vs. human) within a selection interview procedure is certified by a third party according to pre-defined standards might lead applicants to perceive the procedure as fairer because it signals quality and trustworthiness (Kimery & McCord, 2006) as well as ethicality (Osburg et al., 2020). Another framework for explaining the effect of certification on procedural fairness perceptions is provided by fairness heuristic theory. Following fairness heuristic theory, certification can serve as a fairness signal that is used to form a general fairness judgment. Early fairness signals are proposed to have a higher impact on people's general fairness judgment than later fairness signals (Lind, 2001). This implies that informing applicants at the beginning of the personnel selection procedure about the decision agent's certification could strongly influence their general fairness judgment of the procedure. When situations are new or uncertain, when cognitive resources are limited, and when a judgment has to be made quickly, people are especially sensitive and attentive to fairness signals (Lind, 2001). All these circumstances are given in the context of personnel selection procedures so that the certification is expected to be used by applicants to arrive at a general fairness judgment of the personnel selection procedure.

A study that investigated the effect of third-party certifications on talent attraction found that ethical market signals increased the job-pursuit intentions of job seekers (Osburg et al., 2020). In other contexts that are characterized by asymmetric or incomplete information (e.g., e-commerce), third-party certifications have been shown to be effective for trust building (e.g., Aiken & Boush, 2006) and for influencing attitudes (e.g., Lowry et al., 2012). However, not all studies found an effect of certification on trust (e.g., McKnight et al., 2004), stressing the need for further research. So far, no study has investigated the effect of a decision agent's certification within a selection interview on procedural fairness perceptions. Based on signaling theory and fairness heuristic theory, we derive the following hypothesis:

Hypothesis 3: Applicants perceive the personnel selection procedure as procedurally fairer when information on the decision agent's certification is provided compared to when no information on certification is provided.

The Interacting Effect of Decision Agent and Certification on Procedural Fairness

Uncertainty management theory proposes that people are especially attentive and sensitive to fairness signals in situations that are characterized by high uncertainty in order to quickly build a general fairness judgment that can guide attitudes and behavior. The uncertainty in AI-based interviews might be even higher than in traditional interviews because people are not familiar with AI as decision agent (Acikgoz et al., 2020) and do not know how it arrives at a decision. Therefore, informational uncertainty might be particularly high. Furthermore, the situation is probably new to most people because AI usage within selection interviews is not that common yet and people might be uncertain about how to behave and how to feel, leading to personal uncertainty. Those assumptions are empirically supported by the results of Langer et al. (2019) who found that automated selection interviews led to more ambiguity and less perceived controllability. Since people want to overcome the state of uncertainty as quickly as possible, they rely on early fairness signals (e.g., certification) to build a certain and consistent fairness judgment that can serve as a heuristic for the interpretation of new information (Lind & van den Bos, 2002). Due to the higher uncertainty that is experienced as a reaction to AI-based interviews, applicants are expected to pay even more attention to the certification when the decision agent is an AI compared with a human. Therefore, certification might mitigate the negative effect of AI-based interviews on procedural fairness perceptions and reduce the difference between AI-based and human-based interviews in terms of procedural fairness perceptions.

So far, research has not investigated certification as one type of information that could mitigate the negative effect of AI-based selection interviews on fairness perceptions by reducing informational and personal uncertainty. However, other types of information were already investigated as a way to increase fairness perceptions of AI-based selection procedures. The results of these studies are mixed, as one study finds positive effects of information about the algorithm on fairness perceptions (Langer et al., 2018), while another study finds negative effects (Newman et al., 2020). In an attempt to explain the contradictory findings of information on fairness perceptions, Langer et al. (2021) found that process information (i.e., how the algorithm works) can have a detrimental effect, while process justification (i.e., why the algorithm was used) affects fairness perceptions positively. A certification is a different type of information that is, however, more similar to process justification than process information. Rather than providing technical details and explaining how the algorithm works, certification provides assurance through an independent, qualified third party and thereby justifies its usage. Third-party assurance might be a more effective fairness signal compared with technical information because the majority of people are not able to understand technical details. Therefore, AI certification is expected to reduce informational and personal uncertainty effectively. Based on uncertainty management theory and the presented research, it is expected that certification could be an effective instrument to mitigate negative effects of AI-based interviews on procedural fairness perceptions. Accordingly, we propose the following hypothesis:

Hypothesis 4: Information on the decision agent's certification mitigates the negative effect of AI as decision agent on applicants' procedural fairness perceptions of the personnel selection procedure.

Method

Sample

A G*Power analysis was conducted prior to data collection to calculate the necessary sample size. Based on the reported effect sizes of a study that examined the effect of AI usage in personnel selection procedures on fairness perceptions and organizational attraction (Acikgoz et al., 2020), small to medium effects were expected for this study. The analysis revealed a necessary sample size of N = 156 participants to find a small to medium effect (f^2 = 0.04) for a power of $1-\beta$ = .80 and α = .05. Participants were recruited at the Faculty of Human Sciences at a German University and received sweets or course credit for their participation. Data collection continued until 180 participants had completed the experiment to account for dropouts and exclusions. After data collection, data cleaning was performed to ensure high data quality. Participants were excluded who could not recall correctly if an AI or a human analyzes their digital selection interview (N = 13), who failed the attention check (N = 3), and who stated that they did not answer the questions conscientiously (N = 2). The attention check and the conscientiousness check were included to identify careless responses. Two additional participants were excluded because they did not indicate their age and gender, resulting in a final sample size of N = 159 participants. Table 1 provides an overview of the participants' characteristics.

Characteristic (%)					
Gender	Female (76.7)	Male (23.3)			
Occupation	Student (96.2)	Apprentice (0.6)	(Self-)Employed (1.9)	Job-Seeking (0.6)	Other (0.6)
Interview experience	No experience (19.5)	1-2 interviews (46.5)	3-5 interviews (25.8)	6-8 interviews (5.0)	> 8 interviews (3.1)
AI expertise	No knowledge (8.8)	Some knowledge (47.8)	Basic knowledge (37.7)	Advanced knowledge (5.7)	Proficient knowledge (0.0)

Table 1. Characteristics of Participants

Design and Procedure

The experiment took place in two rooms of a German university where participants filled out the survey on a computer and recorded audio files with a headset. We chose the laboratory setting to avoid technical difficulties with the audio recording and to ensure that participants take the interview seriously. We refrained from additionally conducting video recordings due to ethical concerns regarding participants' anonymity. The decision to only use audio recording is not expected to pose a threat to the external validity of our study since automated voice profiling of selection interviews is also common (Chamorro-Premuzic et al., 2016). Before participation, participants were told that the study deals with fairness perceptions of different personnel selection procedures. Neither AI nor certification was mentioned as the study subject to avoid biased responses. According to a 2 (certification: yes vs. no) x 2 (decision agent: AI vs. human) design, participants were randomly assigned to one of four conditions. After receiving information about data anonymity, data processing, and data protection, participants had to give consent for participation and audio recording. Regarding audio recording, participants were informed that their audio files will be saved for a maximum of five years and might be used for additional analyses only by the responsible research team. They were also informed about the possibility of deleting their audio files by generating an individual code at the end of the study. After giving consent, participants received information about the context of the study. They were asked to imagine that they applied for an attractive job and were invited to the first round of a selection procedure. As the first part of the selection procedure, they were instructed to undergo a digital interview procedure which includes replying to three interview questions via audio recording. Furthermore, they were told who will analyze the audio recordings (AI vs. human). This procedure was chosen to put people into a realistic interview situation, thereby enhancing their involvement compared with evaluating hypothetical vignettes or videos what has been done in previous studies (Acikgoz et al., 2020; Langer et al., 2019). On the next page, participants received further information about the decision agent. In the "certified" conditions, a certification label was shown at the top of the page and information about the source of certification and the certified dimensions was provided. The certification label was designed for this study and a reference was made to a fictitious European standard issued by a credible standardization institute. The well-known, independent standardization institute was mentioned as the issuer of certification since previous research has identified source credibility as an important factor for certifications' effectiveness (Kim & Benbasat, 2009). The dimensions on which the decision agent was stated to be certified, were taken from the AI Assessment Catalog (Poretschkin et al., 2023) to ensure a realistic content description of the certification. They include fairness, transparency, reliability, autonomy and control, security, and data protection. The dimensions' description was included to increase the understanding of the certification which has also been shown to be crucial for certifications' effectiveness (Lowry et al., 2012). In the "not-certified" conditions, participants were only told that the decision agent (AI vs. human) tries to consider those dimensions within the selection procedure so that the amount of text was held constant in all conditions.

Following, participants recorded the answers to three interview questions: (1) "Please describe a situation in which you had to work on a task as part of a team. How did you organize the teamwork? Which role did you take over? Did you encounter any difficulties and if so, how did you overcome them?"; (2) "Imagine you are starting a new job and are still in your first few weeks. How would you keep track of and prioritize your tasks? How would you structure your day?"; and (3) "Please name three of your strengths and describe in which situations they help you or have already helped you in the past". These interview questions were chosen in accordance with the questions used by Langer, König, and Hemsing (2020) who used a similar method to investigate applicant reactions to automatically evaluated job interviews. Furthermore, the choice of one situational, one patterned behavior description, and one general question should account for different reactions of applicants to different question types.

After the interview procedure, participants were asked to fill out the scales on procedural fairness and organizational attraction. In the "certified" conditions, the certification label was shown in the right corner of the survey pages to strengthen the manipulation. After filling out the scales, participants indicated their age, gender, and current occupation. They were asked whether they had an idea about the research question and if so, to specify this idea. To ensure that participants read the instructions thoroughly, they had to indicate who analyzes their audio recordings (AI vs. human) and were asked, if they answered all questions consciously. Furthermore, control variables were assessed, namely AI attitude, AI experience, and interview experience. Finally, participants in the "certified" conditions were asked to indicate how trustworthy they perceive the displayed certification label. In the end, participants were thanked and debriefed.

Measures

Procedural fairness. Procedural fairness was measured with the German translation of four subscales of the structure factor of the Selection Procedural Justice Scale (SPJS; Bauer et al., 2001). The structure factor includes formal characteristics of the procedure and can therefore be used to assess procedural fairness. The overall scale had a Cronbach's α = .80. Other studies that examined applicants' fairness perceptions of AI usage in personnel selection procedures have also used the SPJS (Acikgoz et al., 2020; Langer, König, & Hemsing, 2020). The four included subscales are job-relatedness (two items; Cronbach's α = .59), information known (three items; Cronbach's α = .76), chance to perform (four items; Cronbach's α = .84) and feedback (three items; Cronbach's α = .75). Reconsideration opportunity was excluded because participants had no chance to check their answers during the constructed interview procedure so that no variance was expected on this subscale. Participants were asked to respond to each item on a 5-point Likert scale ranging from 1 = "I strongly disagree" to 5 = "I strongly agree". A sample item for chance to perform was: "I could really show my skills and abilities through this test". Since we use the overall scale of procedural fairness for the main analyses, the low Cronbach's α of the subscale job-relatedness is negligible. With regards to the post-hoc analyses, the results for job-relatedness should be interpreted with caution.

Organizational attraction. Organizational attraction was measured with the German translation of two subscales of the Organizational Attraction Scale (OAS; Highhouse et al., 2003), namely company attractiveness (five items; Cronbach's α = .89) and intentions to pursue (five items; Cronbach's α = .86). The overall Cronbach's α of organizational attraction was α = .93. Participants were asked to respond to each item on a 5-point Likert scale ranging from 1 = "I strongly disagree" to 5 = "I strongly agree". The OAS has already been used in previous studies that investigated reactions to AI-based selection procedures (Acikgoz et al., 2020; Langer, König, Sanchez, et al., 2020). A sample item for organizational attractiveness was "For me, this company would be a good place to work".

AI attitude. AI attitude was measured with the Attitude towards Artificial Intelligence scale which is a short measure that consists of five items (Sindermann et al., 2021). Participants responded to the items on an 11-point Likert scale ranging from 1 = "I strongly disagree" to 11 = "I strongly agree". Cronbach's α for the scale was α = .63 and a sample item was "I trust artificial intelligence". Even though the reliability of the scale is lower than acceptable, it was decided to include AI attitude as a control variable as it did not influence the results in a significant way. However, the results regarding AI attitude should be interpreted with caution.

AI expertise. AI expertise was measured on a scale from 1 = "I don't have any knowledge about artificial intelligence" to 5 = "I have very profound knowledge about artificial intelligence".

Interview experience. Interview experience was measured by asking participants to indicate the number of job interviews that they had already completed in the past. Participants could select one of five options (zero job interviews; one to two job interviews; three to five job interviews; six to eight job interviews, more than eight job interviews). Even though interview experience was measured on an ordinal scale, it was treated as an interval variable for the following analyses. This method is reasonable to use because the distances between the categories are quite even and can be interpreted in a meaningful way (Labovitz, 1970).

Certification label's perceived trustworthiness. The certification label's perceived trustworthiness was measured with one item on a scale from 1 = "Not trustworthy at all" to 5 = "Completely trustworthy". This item was only answered by participants who were shown the certification label, and it was included to determine how trustworthy participants perceive the designed label.

Data Analyses

The analyses were conducted with IBM SPSS Statistics (Version 29). We used the SPSS macro PROCESS (Hayes, 2018) which allows for testing conditional process models that include moderation, mediation, or both. Before the analyses were conducted, the variables certification, decision agent, and gender were dummy-coded. Furthermore, the regression assumptions were tested and since none of them was violated, we proceeded with data analyses as planned. To test Hypotheses 1 and 2, we used PROCESS model number 4 and performed a simple mediation analysis with decision agent as the independent variable, procedural fairness as the mediator, and organizational attraction as the dependent variable. As advised, bootstrapping was used as a method to estimate the 95% confidence interval for the indirect effect (Hayes, 2018), and 5000 bootstrap samples were chosen. To test Hypotheses 3 and 4, we used PROCESS model number 7 and performed a moderated mediation analysis in which we added certification as a moderator.

Finally, for exploratory post-hoc analyses, two-sided t-tests were used to test the effect of AI-based interviews on the different procedural fairness subscales. Furthermore, hierarchical linear regression analyses were performed which entail block-wise entry of variables, to test the effect of the perceived trustworthiness of the certification label on procedural fairness perceptions and organizational attraction. In the first step, only the control variables were entered, followed by the independent variables (decision agent and certification label's perceived trustworthiness) in the second step. Finally, the interaction term (certification label's perceived trustworthiness x decision agent) was added in the third step.

Results

Descriptive Statistics

Table 2 shows the means, standard deviations, and Pearson's correlations of the independent variables, dependent variables, and control variables. Furthermore, Cronbach's α values are indicated on the diagonal. Most scholars recommend a reliability of at least .70 or .80 for scales in research (Nunnally, 1978). For both dependent variables, Cronbach's α fulfilled this requirement (Cronbach's $\alpha_{\text{Procedural Fairness}} = .80$; Cronbach's $\alpha_{\text{Organizational Attraction}} = .93$). A strong positive correlation was observed between organizational attraction and procedural fairness, $r = .48$, $p < .001$. Decision agent had a small negative correlation with procedural fairness, $r = -.17$, $p < .05$, and a moderate negative correlation with organizational attraction, $r = -.26$, $p < .001$, pointing in the hypothesized direction. Turning to the control variables, procedural fairness correlated positively with AI attitude, $r = .21$, $p < .01$, and gender, $r = .17$, $p < .05$, while it correlated negatively with interview experience, $r = -.17$, $p < .05$, and age, $r = -.20$, $p < .01$. Organizational attraction correlated positively with AI attitude, $r = .16$, $p < .05$, and gender, $r = .16$, $p < .05$.

	Scale	*M*	*SD*	1	2	3	4	5	6	7	8
1	Certification	.48	.50	-							
2	Decision Agent	.54	.50	.01	-						
3	Procedural Fairness	3.29	.55	-.04	-.17*	(.80)					
4	Organizational Attraction	3.23	.76	.02	-.26**	.48**	(.93)				
5	AI Attitude	6.72	1.53	.01	.04	.21**	.16*	(.63)			
6	AI Expertise	2.40	.73	.12	.02	.09	.13	.23**	-		
7	Interview Experience	2.26	.94	-.01	.15	-.17*	.02	-.01	.00	-	
8	Age	21.82	3.12	.03	-.08	-.20**	.10	.05	.07	.33**	-
9	Gender	1.77	.42	-.03	.00	.17*	.16*	.00	-.23**	-.12	-.28**

Note. N = 159. Certification was coded as 0 = No and 1 = Yes. Decision Agent was coded as 0 = Human and 1 = AI. Gender was coded as 0 = Female and 1 = Male. Numbers on the diagonal are internal consistency reliabilities.
*p < .05; **p < .01.

Table 2. Means, Standard Deviations, and Correlations among the Variables

Hypotheses Testing

Table 3 presents the results of the simple mediation analysis for Hypotheses 1 and 2. Both the model for procedural fairness, $F(6, 152) = 4.84$, $p < .001$, and the model for organizational attraction, $F(7, 151) = 10.99$, $p < .001$, were significant.

Hypothesis 1 proposed that applicants are less attracted to an organization if the decision agent within the selection interview is an AI compared with a human. Consistent with our hypothesis, the total effect of decision agent on organizational attraction was significant and in the expected direction, $c = -.40$, $t(152) = -3.50$, $p <.001$. Furthermore, the direct effect of decision agent on organizational attraction was significant as well which provides additional support for Hypothesis 1, $c' = -.28$, $t(151) = -2.62$, $p = .01$.

Hypothesis 2a stated that applicants perceive the personnel selection procedure as less procedurally fair when the decision agent is an AI compared with a human. The effect of decision agent on procedural fairness was significant and negative, providing support for Hypothesis 2a, $a = -.20$, $t(152) = -2.32$, $p = .02$. Hypothesis 2b suggested that higher procedural fairness perceptions of applicants are associated with higher organizational attraction. On average, participants who perceived the personnel selection procedure as procedurally fairer rated organizational attraction higher, which is consistent with Hypothesis 2b, $b = .62$, $t(151) = 5.79$, $p < .001$. Furthermore, the indirect effect of decision agent on organizational attraction through procedural fairness was significant, 95% CI [-.23, -.02]. This indicates that the relationship between decision agent (AI vs. human) and organizational attraction is partially mediated by applicants' procedural fairness perceptions.

	Consequent							
	M (Procedural Fairness)				*Y* (Organizational Attraction)			
Antecedent		Coeff.	*SE*	*p*		Coeff.	*SE*	*p*
(Constant)	i_M	3.53	.49	< .001	i_Y	-.08	.58	.90
Decision Agent (*X*)	a	-.20	.09	.02	c'	-.28	.11	.01
Procedural Fairness (*M*)					b	.62	.11	< .001
AI Attitude (C_1)	f_1	.07	.03	.02	g_1	.02	.04	.61
AI Expertise (C_2)	f_2	.07	.05	.19	g_2	.12	.07	.10
Interview Experience (C_3)	f_3	-.04	.05	.36	g_3	.07	.06	.20
Age (C_4)	f_4	-.03	.02	.14	g_4	.04	.02	.03
Gender (C_5)	f_5	-.17	.10	.08	g_5	-.30	.13	.02
	$R^2 = .15$, $F(6, 152) = 4.84$, $p < .001$				$R^2 = .33$, $F(7, 151) = 10.99$, $p < .001$			

Note. $N = 159$. Coefficients are not standardized. Certification was coded as 0 = No and 1 = Yes. Decision Agent was coded as 0 = Human and 1 = AI. Gender was coded as 0 = Female and 1 = Male.

Table 3. Results of the Simple Mediation Analysis

For testing Hypotheses 3 and 4, we added certification to our model as a moderator. The results of the moderated mediation analysis are depicted in Table 4.

Hypothesis 3 stated that applicants perceive the personnel selection procedure as procedurally fairer when the decision agent is certified compared to when no information about certification is provided. Contrary to our hypothesis, the decision agent's certification did not influence participants' procedural fairness perceptions significantly, $a_2 = -.09$, $t(150) = -.66$, $p = .51$. Thus, Hypothesis 3 was not supported.

Hypothesis 4 proposed that the decision agent's certification mitigates the negative effect of AI-based selection interviews on procedural fairness perceptions. However, the interaction between decision agent and certification was not found to be significant, $a_3 = .08$, $t(150) = .42$, $p = .68$. This implies that the difference between AI- and human-based selection interviews in terms of procedural fairness perceptions could not be reduced by certification. Hence, Hypothesis 4 was not supported. As the confidence interval for the index for moderated mediation also included zero, 95% CI [-.18, .25], the model that includes certification as moderator fits the data worse than the simple mediation model.

		Consequent						
		M (Procedural Fairness)				*Y* (Organizational Attraction)		
Antecedent		Coeff.	*SE*	*p*		Coeff.	*SE*	*p*
(Constant)	i_M	3.58	.51	< .001	i_Y	-.08	.58	.90
Decision Agent (*X*)	a_1	-.23	.11	.03	c'	-.28	.11	.01
Certification (*W*)	a_2	-.09	.13	.51				
Interaction (*XW*)	a_3	.08	.18	.68				
Procedural Fairness (*M*)					b	.62	.11	< .001
AI Attitude (C_1)	f_1	.07	.03	.02	g_1	.02	.04	.61
AI Expertise (C_2)	f_2	.07	.05	.18	g_2	.12	.07	.10
Interview Experience (C_3)	f_3	-.04	.05	.36	g_3	.07	.06	.20
Age (C_4)	f_4	-.03	.02	.14	g_4	.04	.02	.03
Gender (C_5)	f_5	-.17	.10	.08	g_5	-.30	.13	.02
		$R^2 = .15$, $F(8, 150) = 3.57$, $p < .001$				$R^2 = .33$, $F(7, 151) = 10.99$, $p < .001$		

Note. $N = 159$. Coefficients are not standardized. Certification was coded as 0 = No and 1 = Yes. Decision Agent was coded as 0 = Human and 1 = AI. Gender was coded as 0 = Female and 1 = Male.

Table 4. Results of the Moderated Mediation Analysis

Post-Hoc Analyses

Deeper Analysis of the Procedural Fairness Subscales

Additional two-sided *t*-tests were conducted, comparing the AI to the human conditions. The *t*-tests were conducted to identify differential effects of AI as decision agent on the four subscales of procedural fairness. The following results can guide future work and inspire other researchers' hypothesis development. However, since these post-hoc analyses were only exploratory, they should be interpreted with caution with regards to their generalizability. As can be seen in Table 5, significant mean differences between the AI and human conditions were found with regards to the subscale chance to perform. More precisely, people in the AI conditions perceived chance to perform significantly lower ($M = 2.58$, $SD = .77$) than people in the human conditions ($M = 2.85$, $SD = .85$), $t(157) = -2.11$, $p = .04$, $d = -.34$. The effect size can be interpreted as medium, indicating that especially chance to perform seems to be negatively affected by AI-based selection interviews.

	Human ($n = 73$)		AI ($n = 86$)		$t(157)$	*p*	Cohen's *d*
Subscale	*M*	*SD*	*M*	*SD*			
Job Relatedness	2.35	.76	2.16	.69	-1.67	.10	-.27
Information Known	3.70	.88	3.61	.89	-.64	.52	-.10
Chance to Perform	2.85	.85	2.58	.77	-2.11	.04	-.34
Feedback	4.51	.72	4.34	.82	-1.31	.19	-.21

Note. Two-sided *p*-values are reported.

Table 5. Results of the *t*-Tests Comparing AI and Human Conditions

Deeper Analysis of Certification: The Crucial Role of Perceived Trustworthiness

To deepen our understanding of the effect of certification, we investigated whether the perceived trustworthiness of the certification label can predict procedural fairness perceptions and organizational attraction. For this purpose, two multiple linear regression analyses were conducted for participants who had been shown the certification label ($N = 76$). A check of the assumptions revealed no violations, so we proceeded with the analysis as planned. We entered the control variables in the first step, the independent variables in the second step (i.e., certification label's perceived trustworthiness and decision agent), and the interaction variable in the third step. The model that included the independent variables was significant for procedural fairness, $F(7, 68) = 2.33$, $p = .03$, and for organizational attraction, $F(7,68) = 3.19$, $p = .01$ (see Table 6). The perceived trustworthiness of the certification label had a significant impact on procedural fairness perceptions, $\beta = .27$, $t(68) = 2.30$, $p = .02$, and organizational attraction, $\beta = .24$, $t(68) = 2.13$, $p = .04$. The higher participants rated the trustworthiness of the decision agent's certification label, the higher they rated procedural fairness of the selection procedure and organizational attraction. Including the interaction between the certification label's perceived trustworthiness and decision agent did neither improve the predictive quality of the model for procedural fairness nor for organizational attraction.

	Procedural Fairness			Organizational Attraction		
Antecedent	β	t	p	β	t	p
AI Attitude	.12	.98	.33	.04	.33	.75
AI Expertise	.21	1.71	.09	.37	3.10	<.001
Interview Experience	-.01	-.08	.93	-.03	-.29	.78
Age	-.23	-1.89	.06	.13	1.09	.28
Gender	-.08	-.63	.53	-.18	-1.51	.14
Decision Agent	-.18	-1.58	.12	-.31	-2.81	.01
Certification Label's Perceived Trustworthiness	.27	2.30	.02	.24	2.13	.04
	R^2 = .19, F(7,68) = 2.33, p = .03			R^2 = .25, F(7,68) = 3.19, p = .01		

Note. N = 76. Decision Agent was coded as 0 = Human and 1 = AI. Gender was coded as 0 = Female and 1 = Male.

Table 6. Results of the Post-Hoc Multiple Linear Regression Analyses

Discussion

The goal of the present study was to understand (a) whether applicants are less attracted to an organization and perceive the selection interview as less procedurally fair when the decision agent is an AI compared with a human, and (b) whether AI certification can mitigate potential negative reactions. We are among the first to test AI certification as an instrument to influence people's perceptions of AI usage. Based on signaling theory, fairness heuristic theory, and uncertainty management theory, we proposed that an AI as decision agent within selection interviews has a negative, direct effect on organizational attraction and a negative, indirect effect through procedural fairness perceptions. We further argued that the decision agent's certification has a positive, direct effect on applicants' procedural fairness perceptions as well as a moderating effect, mitigating the negative impact of AI as decision agent within the selection interview. To test our hypotheses, we conducted an experimental study in which 180 participants completed a digital interview procedure and afterwards rated procedural fairness and organizational attraction. As predicted, participants rated organizational attraction lower when the decision agent within the selection interview was an AI compared with a human. Furthermore, the relationship between decision agent and organizational attraction was partially mediated by participants' procedural fairness perceptions. Post-hoc analyses revealed that the negative effect of AI-based selection interviews on procedural fairness is especially strong on the subdimension chance to perform. These results are consistent with previous research (Acikgoz et al., 2020; Langer et al., 2019). It has been proposed that especially the low interactivity and social bandwidth of AI-based interviews result in negative applicant reactions (Langer, König, Sanchez, et al., 2020; Mirowska, 2020). The present study shows that this might not be the only reason why applicants have unfavorable attitudes towards AI-based selection interviews, as no interaction took place in any of the conditions. Regarding the impact of certification, the results show no effect of certification on procedural fairness perceptions and no moderating effect on the relationship between decision agent (AI vs. human) and procedural fairness perceptions. However, post-hoc analyses revealed that the perceived trustworthiness of the displayed certification label has an impact on participants' procedural fairness perceptions and organizational attraction. Previous research that investigated the effect of certifications on trust and attitudes in other contexts has pointed in a similar direction. For example, it has been shown that certifications need to be perceived as credible to have an impact on trust (Atkinson & Rosenthal, 2014). The results of the present study suggest that it is crucial to ensure that people perceive a certification label as trustworthy to influence their procedural fairness perceptions and organizational attraction.

Contributions to Theory

Two main theoretical contributions can be derived from the present study. First, utilizing signaling theory, fairness heuristic theory, and uncertainty management theory, we advance the understanding of people's reactions to AI-based selection interviews using a novel experimental design. Consistent with previous research (e.g., Langer, König, Sanchez, et al., 2020), we found people to be less attracted to an organization when the decision agent in the selection interview is an AI compared with a human. Furthermore, procedural fairness perceptions were identified as a significant mediator, indicating that the negative effect of AI-based selection interviews on organizational attraction can partially be explained by lower procedural fairness perceptions. Since our novel experimental design targeted on more closely representing a real selection interview and increasing the involvement of participants, we strengthen previous findings on the negative effect of AI-based selection interviews on applicant reactions.

Second, our results highlight the critical role of the perceived trustworthiness of an AI certification label in shaping fairness perceptions and attitudes (e.g., organizational attraction), thereby addressing the contradictory findings from prior research on the impact of certifications on trust and attitudes in e-commerce studies (e.g., Aiken & Boush, 2006; McKnight et al., 2004). Arguably, AI certification can positively influence people's perceptions of AI systems and increase the acceptance of AI usage (Poretschkin et al., 2023), yet, this assumption has not been addressed by prior research so far. We theorize and test that the certification of the decision agent (AI vs. human) within a selection interview influences applicants' procedural fairness perceptions positively and mitigates the negative effect of AI-based selection interviews. While our results do not support our initial assumptions, deeper analyses uncover that the perceived trustworthiness of the certification label has a direct, positive effect on procedural fairness perceptions and organizational attraction. This finding implies that a trustworthy certification has the potential to influence both procedural fairness perceptions and organizational attraction within a personnel selection procedure. Therefore, we identify a boundary condition for fairness heuristic theory, indicating that a fairness signal (e.g., certification) has to be perceived as trustworthy to have an impact on the general fairness judgment. Our findings highlight the need to design certification labels carefully, focusing on design factors such as trustworthiness. Otherwise, the intended effect of a certification label may not be achieved.

Contributions to Practice

The present study yields four important contributions to practice. First, when displaying a certification label to applicants, organizations have to ensure that applicants perceive the label as trustworthy to improve reactions towards the selection procedure and the organization. Previous research found that especially the source influences the credibility of a certification label, and thereby also trust (Atkinson & Rosenthal, 2014). More specifically, independent third-party certifications have been shown to induce more trust than self-proclaimed certifications (Kim & Benbasat, 2009). The results of the present study indicate that other parameters besides source are important to signal trustworthiness since the independent third-party certification that we used still did not lead to a general effect of certification. Specifically, with regards to an AI certification label that has not been developed yet, success factors should be identified that enhance perceptions of trustworthiness. Second, even though organizations might benefit from AI usage in personnel selection procedures in terms of saving money and time, they should be aware of possible negative applicant reactions. Personnel selection procedures often represent the first interaction between an organization and an applicant. The results of the present study indicate that AI usage might act as an early fairness signal that negatively influences applicants' procedural fairness perceptions and organizational attraction. In times of war for talent, it becomes crucial for organizations to keep applicants attracted to the organization throughout the personnel selection procedure (Chapman et al., 2005). Therefore, a procedure should be chosen that applicants perceive as fair, for example by including a human in the loop (Mosqueira-Rey et al., 2023) or by conducting a preliminary interview in which more information about AI usage is provided. The preliminary interview might lead applicants to build a positive fairness heuristic which then mitigates the influence of later negative fairness signals, like AI usage. Third, the results emphasize the importance for organizations to consider applicants' procedural fairness perceptions of the selection procedure, as our study could show that they have a significant, large effect on organizational attraction. It becomes apparent that besides the mitigation of bias and equal treatment of applicants, organizations should be sensitive to other factors that influence fairness perceptions. Finally, our study has an important policy implication. The more organizations become aware of negative applicant reactions towards AI usage within personnel selection procedures, the less they might be willing to disclose AI usage to applicants. That is why regulations like the EU AI Act are very important to ensure transparency so that applicants are informed about who evaluates their digital interviews (AI vs. human).

Limitations and Future Research

Some limitations have to be mentioned concerning the present study. First, the majority of participants were students (96.2%). People with higher levels of education could have different fairness perceptions of

AI usage and react differently to organizations that conduct AI-based selection interviews. However, Araujo et al. (2020) found no effect of education on fairness perceptions of automated decision-making, so this limitation might be neglectable. Second, as in most laboratory experiments, external validity might be a limitation. Even though the study aimed at simulating a realistic interview procedure, the involvement of participants cannot be compared to a real selection interview. Before the digital interview, participants were asked to name three adjectives that describe how they feel with regards to the upcoming interview. As the majority of participants listed adjectives like “excited” or “nervous”, this might indicate some involvement. Nonetheless, it would be beneficial to conduct a field study or to take other measures to increase involvement. Finally, even though we used participants’ perception of the certification label’s trustworthiness as a manipulation check, it cannot be claimed with certainty that participants paid attention to the certification label. Therefore, following studies should ensure that participants are aware of the certification label or let them indicate whether they paid attention to the label.

With regards to future research, two broad aspects should be explored further. First, more studies should address the question of how AI certification influences applicant reactions to AI-based selection procedures. Particular attention should be paid to factors that have an impact on the perceived trustworthiness of signals like AI certification labels. Relevant factors might include characteristics of (a) the signal, (b) the signaler, (c) the receiver, and (d) the signaling environment (Lins & Sunyaev, 2017). As the trustworthiness of an AI certification label might be influenced by different factors than other certification labels, it might be especially important to conduct additional research that focuses on AI certification labels, in particular. Furthermore, it should be investigated when applicants are most attentive to AI certification within the recruitment process. The present study only looked at the effect of AI certification on fairness perceptions after applicants have already entered the recruitment process. However, AI certification could also have an impact on applicants’ initial decision to apply to a company. The second aspect that future research should further address is the question of which factors might reinforce or mitigate negative reactions to AI-based selection interviews. For example, the negative effect of AI-based selection procedures might be weaker for jobs that mostly require technical skills than for jobs that mostly require social skills because an AI is not expected to have the capabilities to evaluate human skills adequately (Newman et al., 2020). It could also be investigated whether educating applicants about the chances and challenges of AI before the selection interview might mitigate the negative effect of AI-based selection interviews on fairness perceptions. Furthermore, a hybrid condition could be included in future studies to investigate how the collaboration between a human and an AI is perceived by applicants.

Conclusion

Personnel selection procedures are one field in which AI usage becomes more common, for example, within digital interview procedures. Therefore, it is crucial to understand applicant reactions towards AI-based selection interviews and to investigate factors that might improve potential negative reactions, for example, AI certification. The results of the present study show that applicants are less attracted to an organization that conducts AI-based (compared with human-based) selection interviews, partially due to lower procedural fairness perceptions. The certification of the decision agent (AI vs. human), had no overall effect on participants’ fairness perceptions and did not mitigate the negative effect of AI usage. However, if participants perceived the certification label to be trustworthy, the label did influence procedural fairness perceptions and organizational attraction. Hence, it can be concluded that certification might have the potential to act as an instrument to improve applicant reactions towards AI-based interview procedures, but only if the certification label is perceived as trustworthy which emphasizes the need for future research.

References

Acikgoz, Y., Davison, K. H., Compagnone, M., & Laske, M. (2020). Justice perceptions of artificial intelligence in selection. *International Journal of Selection and Assessment*, *28*(4), 399–416. https://doi.org/10.1111/ijsa.12306

Aiken, K. D., & Boush, D. M. (2006). Trustmarks, Objective-Source Ratings, and Implied Investments in Advertising: Investigating Online Trust and the Context-Specific Nature of Internet Signals. *Journal of the Academy of Marketing Science*, *34*(3), 308–323. https://doi.org/10.1177/0092070304271004

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

Atkinson, L., & Rosenthal, S. (2014). Signaling the Green Sell: The Influence of Eco-Label Source, Argument Specificity, and Product Involvement on Consumer Trust. *Journal of Advertising*, *43*(1), 33–45. https://doi.org/10.1080/00913367.2013.834803

Bauer, T. N., Truxillo, D. M., Sanchez, R. J., Craig, J. M., Ferrara, P., & Campion, M. A. (2001). Applicant Reactions to Selection: Development of the Selection Procedural Justice Scale (SPJS). *Personnel Psychology*, *54*(2), 387–419. https://doi.org/10.1111/j.1744-6570.2001.tb00097.x

Chamorro-Premuzic, T., Winsborough, D., Sherman, R. A., & Hogan, R. (2016). New Talent Signals: Shiny New Objects or a Brave New World? *Industrial and Organizational Psychology*, *9*(3), 621–640. https://doi.org/10.1017/iop.2016.6

Chapman, D. S., Uggerslev, K. L., Carroll, S. A., Piasentin, K. A., & Jones, D. A. (2005). Applicant Attraction to Organizations and Job Choice: A Meta-Analytic Review of the Correlates of Recruiting Outcomes. *Journal of Applied Psychology*, *90*(5), 928–944. https://doi.org/10.1037/0021-9010.90.5.928

Chapman, D. S., & Webster, J. (2003). The Use of Technologies in the Recruiting, Screening, and Selection Processes for Job Candidates. *International Journal of Selection and Assessment*, *11*(2–3), 113–120. https://doi.org/10.1111/1468-2389.00234

Colquitt, J. A., Greenberg, J., & Zapata-Phelan, C. P. (2005). What is organizational justice? A historical overview. In J. Greenberg & J. A. Colquitt (Eds.), *Handbook of organizational justice* (pp. 3–56). Lawrence Erlbaum Associates Publishers.

Colquitt, J. A., Hill, E. T., & De Cremer, D. (2023). Forever focused on fairness: 75 years of organizational justice in Personnel Psychology. *Personnel Psychology*, *76*(2), 413–435. https://doi.org/10.1111/peps.12556

Connelly, B. L., Certo, S. T., Ireland, R. D., & Reutzel, C. R. (2011). Signaling Theory: A Review and Assessment. *Journal of Management*, *37*(1), 39–67. https://doi.org/10.1177/0149206310388419

Dastin, J. (2022). Amazon Scraps Secret AI Recruiting Tool that Showed Bias against Women. In K. Martin (Ed.), *Ethics of Data and Analytics* (pp. 296–299). Auerbach Publications.

Durth, S., Gandhi, N., Komm, A., & Pollner, F. (2022). *HR's new operating model*. McKinsey&Company. https://www.mckinsey.com/capabilities/people-and-organizational-performance/our-insights/hrs-new-operating-model

European Commission. (2021). Proposal for a Regulation of the European Parliament and of the Council Laying down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act) and Amending Certain Union Legislative Acts (COM(2021) 206 final). https://artificialintelligenceact.eu/the-act/

Gilliland, S. W. (1993). The Perceived Fairness of Selection Systems: An Organizational Justice Perspective. *Academy of Management Review*, *18*(4), 694–734. https://doi.org/10.5465/amr.1993.9402210155

Hayes, A. F. (2018). Introduction to Mediation, Moderation, and Conditional Process Analysis: A Regression-Based Approach (2nd ed.). The Guilford Press.

Highhouse, S., Lievens, F., & Sinar, E. F. (2003). Measuring Attraction to Organizations. *Educational and Psychological Measurement*, *63*(6), 986–1001. https://doi.org/10.1177/0013164403258403

Hmoud, B., & Laszlo, V. (2019). Will artificial intelligence take over human resources recruitment and selection? *Network Intelligence Studies*, *7*(13), 21–30.

Kim, D., & Benbasat, I. (2009). Trust-Assuring Arguments in B2C E-commerce: Impact of Content, Source, and Price on Trust. *Journal of Management Information Systems*, *26*(3), 175–206. https://doi.org/10.2753/MIS0742-1222260306

Kimery, K. M., & McCord, M. (2006). Signals of Trustworthiness in E-Commerce: Consumer Understanding of Third-Party Assurance Seals. *Journal of Electronic Commerce in Organizations (JECO)*, *4*(4), 52–74. https://doi.org/10.4018/jeco.2006100103

Labovitz, S. (1970). The Assignment of Numbers to Rank Order Categories. *American Sociological Review*, *35*(3), 515–524. https://doi.org/10.2307/2092993

Langer, M., Baum, K., König, C. J., Hähne, V., Oster, D., & Speith, T. (2021). Spare me the details: How the type of information about automated interviews influences applicant reactions. *International Journal of Selection and Assessment*, *29*(2), 154–169. https://doi.org/10.1111/ijsa.12325

Langer, M., König, C., & Fitili, A. (2018). Information as a double-edged sword: The role of computer experience and information on applicant reactions towards novel technologies for personnel selection. *Computers in Human Behavior*, *81*, 19–30. https://doi.org/10.1016/j.chb.2017.11.036

Langer, M., König, C. J., & Hemsing, V. (2020). Is anybody listening? The impact of automatically evaluated job interviews on impression management and applicant reactions. *Journal of Managerial Psychology*, *35*(4), 271–284. https://doi.org/10.1108/JMP-03-2019-0156

Langer, M., König, C. J., & Papathanasiou, M. (2019). Highly automated job interviews: Acceptance under the influence of stakes. *International Journal of Selection and Assessment*, *27*(3), 217–234. https://doi.org/10.1111/ijsa.12246

Langer, M., König, C. J., Sanchez, D. R.-P., & Samadi, S. (2020). Highly automated interviews: Applicant reactions and the organizational context. *Journal of Managerial Psychology*, *35*(4), 301–314. https://doi.org/10.1108/JMP-09-2018-0402

Lind, E. A. (2001). Fairness Heuristic Theory: Justice Judgments as Pivotal Cognitions in Organizational Relations. In J. Greenberg & R. Cropanzano (Eds.), *Advances in Organizational Justice* (pp. 56–88). Stanford University Press.

Lind, E. A., & van den Bos, K. (2002). When fairness works: Toward a general theory of uncertainty management. *Research in Organizational Behavior*, *24*, 181–223. https://doi.org/10.1016/S0191-3085(02)24006-X

Lins, S., & Sunyaev, A. (2017, December 10). *Unblackboxing IT Certifications: A Theoretical Model Explaining IT Certification Effectiveness*. Thirty Eighth International Conference on Information Systems (ICIS 2017), South Korea. https://aisel.aisnet.org/icis2017/Security/Presentations/26

Lowry, P. B., Moody, G., Vance, A., Jensen, M., Jenkins, J., & Wells, T. (2012). Using an elaboration likelihood approach to better understand the persuasiveness of website privacy assurance cues for online consumers. *Journal of the American Society for Information Science and Technology*, *63*(4), 755–776. https://doi.org/10.1002/asi.21705

McKnight, D. H., Kacmar, C. J., & Choudhury, V. (2004). Shifting Factors and the Ineffectiveness of Third Party Assurance Seals: A Two-Stage Model of Initial Trust in a Web Business. *Electronic Markets*, *14*(3), 252–266. https://doi.org/10.1080/10196780420000245263

Mirowska, A. (2020). AI Evaluation in Selection: Effects on application and pursuit intentions. *Journal of Personnel Psychology*, *19*(3), 142–149. https://doi.org/10.1027/1866-5888/a000258

Mosqueira-Rey, E., Hernández-Pereira, E., Alonso-Ríos, D., Bobes-Bascarán, J., & Fernández-Leal, Á. (2023). Human-in-the-loop machine learning: A state of the art. *Artificial Intelligence Review*, *56*(4), 3005–3054. https://doi.org/10.1007/s10462-022-10246-w

Newman, D. T., Fast, N. J., & Harmon, D. J. (2020). When eliminating bias isn't fair: Algorithmic reductionism and procedural justice in human resource decisions. *Organizational Behavior and Human Decision Processes*, *160*, 149–167. https://doi.org/10.1016/j.obhdp.2020.03.008

Nunnally, J. C. (1978). *Psychometric theory* (2nd ed.). McGraw-Hill.

Osburg, V.-S., Yoganathan, V., Bartikowski, B., Liu, H., & Strack, M. (2020). Effects of Ethical Certification and Ethical eWoM on Talent Attraction. *Journal of Business Ethics*, *164*(3), 535–548. https://doi.org/10.1007/s10551-018-4018-8

Poretschkin, M., Schmitz, A., Akila, M., Adilova, L., Becker, D., Cremers, A. B., Hecker, D., Houben, S., Mock, M., Rosenzweig, J., Sicking, J., Schulz, E., Voss, A., & Wrobel, S. (2023). *Guideline for Designing Trustworthy Artificial Intelligence: AI Assessment Catalog*. https://www.iais.fraunhofer.de/en/research/artificial-intelligence/ai-assessment-catalog.html

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. https://doi.org/10.1080/07370024.2020.1735391

Rynes, S. L., Bretz Jr, R. D., & Gerhart, B. (1991). The Importance of Recruitment in Job Choice: A Different Way of Looking. *Personnel Psychology*, *44*(3), 487–521. https://doi.org/10.1111/j.1744-6570.1991.tb02402.x

Sindermann, C., Sha, P., Zhou, M., Wernicke, J., Schmitt, H. S., Li, M., Sariyska, R., Stavrou, M., Becker, B., & Montag, C. (2021). Assessing the Attitude Towards Artificial Intelligence: Introduction of a Short Measure in German, Chinese, and English Language. *KI - Künstliche Intelligenz*, *35*(1), 109–118. https://doi.org/10.1007/s13218-020-00689-0

Spence, M. (1973). Job Market Signaling. *The Quarterly Journal of Economics*, *87*(3), 355–374. https://doi.org/10.2307/1882010

van den Bos, K., Lind, E. A., & Wilke, H. A. M. (2001). The Psychology of Procedural and Distributive Justice Viewed From the Perspective of Fairness Heuristic Theory. In R. Cropanzano (Ed.), *Justice in the workplace: From theory to practice* (Vol. 2, pp. 49–66). Lawrence Erlbaum Associates Publishers.