Stockholm School of Economics
MSc in Business and Management
Master Thesis
Spring 2023

Fairness in Hire
What Talents Admire

Scrutinizing fairness perceptions of AI and transparency
in asynchronous video interviews

Authors
Shao-Fu Lu (42056)
Yining Zhang (42073)

Supervisor
Elmira van den Broek, Assistant Professor
Department of Entrepreneurship, Innovation and Technology

Date
May 15, 2023

Word count
17,735

Abstract

In pursuit of enhancing time- and cost-efficiency, organizations have been adopting asynchronous video interviews (AVIs) integrated with artificial intelligence (AI) to optimize and to automate recruitment processes. The roles of AI decision agents and transparency in shaping job applicants' fairness perceptions in AVIs remain largely unexplored despite the growing prominence of this novel interview format. The aim of this thesis is, therefore, to bridge the research gap by investigating applicants' procedural fairness perceptions in this technology-mediated interview assessment, taking into account the influences of different types of decision makers, the disclosure of decision makers, and the provision of explanations. An online scenario-based between-subject experiment was conducted and Gilliland's justice model was employed to measure applicants' fairness perceptions in the recruitment process. Based on the quantitative analysis of 288 observations, the findings revealed that (1) AVIs rated by humans are perceived procedurally fairer than those rated by AI; (2) transparency regarding decision makers does not significantly influence applicants' perceptions of procedural fairness; and (3) the provision of explanations has a moderating effect on procedural fairness perceptions only when the decision makers are humans. The research yields several important implications for organizations to mitigate applicants' potential negative perceptions and reactions and to facilitate a progressive transition from human to AI delegation in AVIs.

Keywords:
Artificial intelligence, transparency, fairness perception, procedural fairness, recruitment, asynchronous video interview

Acknowledgement

We would like to take this opportunity to express our heartfelt gratitude to everyone who has supported and encouraged us throughout the journey.

First of all, we want to thank Elmira van den Broek, our supervisor, for exceptional guidance, invaluable feedback, and unwavering support.

Also, we would like to thank all of the participants for spending time doing our survey. Your responses, voices, and thoughts are instrumental in this thesis and have added depth and richness to the research findings.

Finally, we appreciate our families and friends for all the support.

Once again, we are deeply grateful to everyone who contributed to our academic journey and made this achievement possible.

Table of contents

1. Introduction

With the advances in artificial intelligence (AI), the way businesses govern their recruitment processes and manage their human resources operations has been continuously transformed. The use of AI-based technologies has disrupted traditional personnel selection and been increasingly applied throughout recruitment processes (Hunkenschroer and Luetge, 2022; Kazim et al., 2021; Yarger et al., 2020).

As AI keeps evolving with new developments and new applications, it encompasses a wide range of technologies and has not been globally defined. However, the descriptions of AI are commonly associated with interpreting data; learning from experience; simulating human intelligence, such as perceiving, reasoning, and learning; and eventually performing human-like tasks (Berente et al., 2019; Duan et al., 2019; Glikson and Woolley, 2020; Kaplan and Haenlein, 2019; Longoni et al., 2019; Rai et al., 2019; von Krogh, 2018). Employing AI-based hiring software, companies can address the challenges in application surplus, perform standardized evaluations, and enhance the quality of the candidate pool with higher efficiency, in terms of both time and cost (Black and van Esch, 2020; Li et al., 2021; van Esch and Black, 2019).

The key technology that underlies AI is machine learning (ML), which has been in rapid progress over the last two decades (Berente et al., 2021). Practically, ML-based decision making involves training the algorithms with features extracted from historical data to construct a generalizable decision rule and make corresponding predictions for the future (Lipton, 2018; Tambe et al., 2019). In other words, ML models predict whether the applicants satisfy the qualifications for a role and select candidates by, for example, automatically screening information in their resumes or analyzing their personal traits in video assessments (Kazim et al., 2021).

There has been an emerging trend toward video interview analysis in recruitment (Hunkenschroer and Luetge, 2022), particularly web-based asynchronous video interviews (AVIs; Lukacik et al., 2022; Suen and Hung, 2023). AVI is a one-way, asynchronous, technology-mediated interview in which interviewees record themselves answering predefined questions in front of their webcam on an interview platform within a fixed response time (Brenner et al., 2016; Mejia and Torres, 2018). Importantly, the recorded responses are evaluated by either hiring practitioners or AI systems at a later time (Basch and Melchers, 2019; Levashina et al., 2014). The commonly recognized benefits of employing AVIs include, among others, reducing the burden of scheduling interviews, increasing the number of applicants being assessed, allowing multiple representatives to make a collective decision, and ensuring consistency of questioning across all interviews (Danieli et al., 2016; Gorman et al., 2018; Lukacik et al., 2022; Sellers, 2014). However, due to the relatively recent development and implementation of AVI tools, their effects on interview takers' reactions are still being empirically researched.

1.1 Problem discussion

Understanding applicants' reactions becomes increasingly important due to the competitive landscape for talent acquisition and the ethical considerations of algorithmic decision-making processes (Leicht-Deobald et al., 2019; Mujtaba and Mahapatra, 2019). Ensuring favorable perceptions among applicants is crucial as it not only is significant for the employers' reputation (Cable and Turban, 2003) but also substantially impacts applicants' actions thereafter (Acikgoz et al., 2020; Swider et al., 2015; Walker et al., 2013). Justice perceptions are generally considered vitally important for applicants' reactions (Acikgoz et al., 2020). Particularly, Gilliland's (1993) justice model states that fairness perceptions directly influence attitudes and behaviors; such as organizational attractiveness, recommendation intentions, and offer acceptance; during the selection process and after the hiring decision (Acikgoz et al., 2020; McCarthy et al., 2017a; Schinkel et al., 2016).

While the increasing use of algorithms optimizes human resource management (HRM) practices, whether AI-made decisions are fairer than human-made ones remains an open question (Lavanchy et al., 2023). Although research in computer science and AI ethics has been actively trying to embed notions of fairness into the design principles of algorithms (Arrieta et al., 2020; Lavanchy et al., 2023), the primary discussion around AI-based recruitment centers on if the outcome is biased or discriminatory (Köchling and Wehner, 2020). While some believe that the use of AI in hiring potentially reduces human bias, others have cautioned about the potential to reinforce existing biases (Li et al., 2021; Zielinski, 2020). Because algorithmic models are trained on historical data to perform prediction tasks for the future, biases can be embedded into algorithms through design principles, feature selection, and training data (Kleinberg and Mullainathan, 2019; Yarger et al., 2020). Moreover, AI systems are incapable of recognizing bias and cannot determine if they make discriminatory decisions (Beattie and Johnson, 2012; Black and van Esch, 2020; Danieli et al., 2016), thus possibly reinforcing discrimination (Tambe et al., 2019; Vasconcelos et al., 2018), exacerbating inequality (Yarger et al., 2020), and conflicting the societal expectations for making ethical decisions (Hunkenschroer and Luetge, 2022).

In addition, transparency plays an important role in enhancing applicants' fairness perceptions in AI-based recruitment processes. Rynes et al. (1991) state that insufficient information can cause negative fairness perceptions in the selection process as applicants may extrapolate signals from available information. While informational transparency is hindered in algorithmic hiring tools, applicants can merely make incomprehensive inferences about the firm and the process, thus intensifying the feeling of uncertainty and hastily forming heuristics and negative fairness perceptions (Acikgoz et al., 2020). The other concern for transparency resides in the decision-making process (i.e., the so-called "black box"). When algorithms rely on numerous data points to evaluate a candidate, it becomes

challenging to provide a clear explanation of the attributes that drive the decisions (Raghavan et al., 2020; Simbeck, 2019). A lack of explainability, which revolves around the extent to which the algorithmic system can be comprehended and the outcome can be elucidated (Kazim et al., 2021; Schumann et al., 2020), is problematic. Not only do the recruiters need to evaluate whether the rationale behind the decisions is legitimate and equitable (Kazim et al., 2021) but also the applicants need explanations for why they are (not) selected and feedback for potentially enhancing their candidacy (Dattner et al., 2019; Hunkenschroer and Kriebitz, 2022; van Esch and Black, 2019).

1.2 Research purpose and research questions

Despite the growing implementation of AI-based tools in recruitment, there is limited empirical research on applicants' responses to the utilization of AI in personnel selection and it remains a nascent topic in academic literature (Acikgoz et al., 2020; Hilliard et al., 2022; Langer et al., 2019; McCarthy et al., 2017a). Also, applicants' fairness perceptions of AI are inconsistent (Hunkenschroer and Luetge, 2022), especially in how people perceive decisions made by algorithms compared to those made by humans (e.g., Lee, 2018). Generally, prior studies indicate that, compared to human-only or AI-assisted human processes, AI-only recruitment processes are perceived as less procedurally fair regardless of whether the outcome is favorable (Dietvorst et al., 2014; Lavanchy et al., 2023; Newman et al., 2020).

Interestingly, job applicants' fairness perceptions are mixed when they are in different recruitment phases and when different assessment tools are used (Georgiou and Nikolaou, 2020; Hilliard et al., 2022; Suen et al., 2019). As AI-based AVIs have been employed recently and are inherently distinct from previous personnel selection formats (e.g., face-to-face interviews; Mejia and Torres, 2018), more empirical research is required to understand the perceptions in the setting of this new technology-mediated recruitment tool. Therefore, it is of our first research interest to investigate the differences in fairness perceptions between AI and humans as the decision makers in AVIs.

> ***RQ1:*** *To what extent do AI and human decision makers affect job applicants' perceptions of fairness in AVIs?*

Secondly, how transparency in AI-based hiring decisions relates to perceived fairness remains relatively unexplored (Hilliard et al., 2022). Extant literature mostly treats fairness perception and transparency as independent factors that are associated with job applicants' reactions (e.g., de Greeff et al., 2021; Lee and Cha, 2023). However, transparency can be considered one factor (van Esch and Black, 2019) or even the basis of fairness perceptions (Abdul et al., 2018). Moreover, more research is needed to provide human resources practitioners to comprehend the role of transparency

through the lens of perceived fairness and its influences in AI-based recruitment. Especially, while AI-based AVIs can process visual, verbal, and vocal data of the candidates to evaluate their suitability for the job (Suen and Hung, 2023), there are concerns about how these AI models work. Without direct interactions between the interviewer and the interviewee, the doubt of fairness can be amplified when AVI is employed (Mirowska and Mesnet, 2022).

In light of the aforementioned empirical research gap, this study aims to explore the role of transparency, investigate job applicants' perceptions of the new interview tool, and further derive more desirable practices in AI-based recruitment. Hence, the purposes lead us to the second research question as follows.

> ***RQ2:*** *How does transparency affect job applicants' perceptions of fairness in AVIs?*

1.3 Expected contribution

This study is expected to deliver three main contributions to the domain of AI recruitment. Firstly, it specifically examines job applicants' fairness perceptions in response to different decision-making agents in AVI scenarios. Secondly, this research explores the effects of transparency and of the interactions between explanations and different decision makers. Finally, this thesis aspires to provide practical implications to mitigate societal concerns about fairness in AI-based recruitment and to make traditionally broad transparency concepts more tangible for recruiters to incorporate in their work. Thus, organizations can formulate a virtuous circle in which they ensure applicants' positive reactions, build higher reputations, and attract more talents.

1.4 Delimitations

Firstly, this study focuses on the phase of interview assessment in the recruitment pipeline, particularly in AVIs. Typically, AI-enabled assessment software includes task-based, video-based, and game-based assessments (Li et al., 2021). Although employing distinct assessment formats could trigger different reactions among applicants (Suen et al., 2019), it is not feasible for the authors to simultaneously simulate task-based and game-based assessments in the experiment with the constraints on time and resources. Moreover, AI is inherently used in those two types of assessments, diminishing the relevance of exploring the influence of different decision-making agents. As a recent innovation that has attracted increasing attention and begun to displace traditional interview formats (Mejia and Torres,

2018; Rasipuram and Jayagopi, 2018; Torres and Mejia, 2017), AVI is chosen as the scenario in the research.

The second delimitation is that this study focuses on perceived fairness from job applicants' perspective. Such perceptions are crucial as they impact not only organizational attractiveness but also job applicants' intentions to accept the offer, recommend the organization to others, or withdraw from the selection process (Guchait et al., 2014; Truxillo et al., 2009). Additionally, the majority of job applicants do not possess expert knowledge in algorithms, directing the authors to study from a managerial perspective instead of delving into computational, factual fairness or algorithmic technicalities.

1.5 Thesis outline

This thesis comprises six chapters: (1) Introduction, (2) Literature review, (3) Methodology, (4) Results and analyses, (5) Discussion, and (6) Conclusion. In the next chapter, an extant literature and theory review that establishes the foundation for the hypotheses is presented. In chapter three, the research method and process are described and the research quality is discussed. Next, analyses of survey data and the results of hypothesis testing are presented in chapter four. Thereafter, chapter five includes discussions of the findings, practical implications, research limitations, and suggestions for future research. Finally, this study is concluded in chapter six.

2. Literature review

In this chapter, we present a review of extant literature that lays the foundation of the hypotheses. Initially, organizational justice theory and procedural fairness rules are presented. Then, how humans and AI as decision-making agents affect fairness perceptions is reasoned. Thereafter, a comprehensive review of transparency and its role in recruitment are elucidated; specifically, disclosure of decision makers and an extra provision of explanations. Next, how explanation (greater transparency) as a moderator interacts with different decision makers and influences procedural fairness perceptions is discussed. Lastly, the chapter is concluded with a visualized conceptual model and a summary of hypotheses.

2.1 Fairness

Firms should both admit their responsibilities and make strong commitments to organizational justice that requires treating different stakeholders with respect, equality, and fairness (Cropanzano et al., 2007; Demuijnck, 2009; Greenberg, 1990). In particular, firms are responsible for implementing a fair and just recruitment process for all job applicants (Gilliland, 1993). Fairness is regarded as a social construct within the realm of organizational science (Colquitt et al., 2001). In other words, an action is deemed fair if a majority of people perceive it as such (Cropanzano and Greenberg, 1997). Therefore, the notion of "what is fair" derived from previous study connects objective aspects of decision-making process and subjective perceptions of fairness (Colquitt et al., 2001).

The objective aspects refer to factual fairness that includes the objectively measurable features, whereas the subjective aspects pertain to perceived fairness, which is related to individuals' perceptions (Hooker, 2005; Marcinkowski et al., 2020; Pawlenka, 2005; Shulner-Tal et al., 2023). These two aspects are presumably correlated but are conceptually distinct (Marcinkowski et al., 2020). In the context of job applications, it is not only about how fair the procedure itself is but also about how fair the applicants perceive the whole process is (Köchling and Wehner, 2020). Furthermore, it is not enough for the hiring process to be factually fair if job applicants perceive it to be unfair and are dissatisfied (Marcinkowski et al., 2020). Therefore, perceived fairness must be treated as a pivotal value with respect to the recruitment process.

2.1.1 Organizational justice theory

Organizational justice deals with fairness perceptions in the workplace (Byrne and Cropanzano, 2001) and has served as a significant basis for research in the field of

job applicants' reactions (Ployhart et al., 2017). A great deal of research is based on Gilliland's (1993) theoretical model of applicant reactions to employment selection systems. Theoretically, justice and fairness are two different concepts as Aristotle, an ancient Greek philosopher, defined justice as the sum of lawfulness and fairness (Guest, 2017). However, empirical studies often inquire about individuals' perceptions of justice and fairness without explicitly investigating potential disparities between the two (Cugueró and Rosanas, 2011). In that sense, the labels used by empirical researchers are not consistent with those proposed by Aristotle, and the concepts of justice and fairness have been used interchangeably (Cugueró and Rosanas, 2011).

Organizational justice literature suggests using a multidimensional construct of fairness (Gilliland, 1993; Greenberg, 1987). Used as a framework for examining applicants' reactions to selection situations, Gilliland's (1993) organizational justice theory focuses on two dimensions, namely distributive fairness and procedural fairness. Distributive fairness is the perceived fairness of the hiring decision (i.e., the selection outcome) and procedural fairness is the perceived fairness of the selection activities (i.e., the processes employed to achieve the outcome; Leventhal, 1980; Skarlicki and Folger, 1997). Other research (e.g., Bies, 2005; Greenberg, 1990; Greenberg and Cropanzano, 1993) further distinguishes procedural fairness from interactional fairness, which has been divided into informational fairness and interpersonal fairness (Acikgoz et al., 2020). The former pertains to the sufficiency of procedural information and justifications, and the latter refers to the treatment of politeness, dignity, and respect (Bell et al., 2006; Colquitt et al., 2001; Rupp et al., 2017). These two dimensions serve as additional cues to the enactment of procedural fairness (Mirowska and Mesnet, 2022) and are found influential in the recruitment processes (Gilliland and Hale, 2005).

The issues of procedural fairness in recruitment have been extensively highlighted in extant literature. The impact of perceived procedural fairness extends to a range of attitudinal outcome (e.g., organizational trust, attractiveness, and commitment) and behavioral outcome (e.g., offer acceptance and recommendation intentions; Cohen-Charash and Spector, 2001; Colquitt et al., 2001; McCarthy et al., 2017; Ötting and Maier, 2018; Schinkel et al., 2016). While both distributive and procedural fairness are important in shaping fairness perceptions, research (Folger and Konovsky, 1989; Tyler et al., 1985) has indicated that variables associated with procedural fairness tend to explain a greater amount of variance in fairness judgments compared to those related to distributive fairness (van den Bos et al., 2001). In other words, procedural fairness is a stronger predictor of overall fairness judgment than distributive fairness (Morse et al., 2022). For example, people tend to lean on procedural fairness when they are uncertain about how trustworthy the decision maker is (van den Bos et al., 1998). This tendency is particularly relevant in the context of AI systems that are often perceived to be opaque and difficult to understand (Glikson and Woolley, 2020). However, problems of procedural fairness

receive far less attention than issues of distributive fairness in AI research (Marcinkowski et al., 2020). Previous research on fairness has concentrated on achieving a fair distribution of hiring outcome, with little attention paid to the decision-making process in which the outcome is generated (Grgić-Hlača et al., 2018). Also, the designs of fairness metrics focus almost exclusively on distributive fairness (Saxena et al., 2020; Selbst et al., 2019). Therefore, procedural fairness is an important topic to be discussed.

2.1.2 Procedural fairness rules

Procedural fairness is associated with fairness perceptions in the decision-making process and perceptions of the way that individuals are treated (Gilliland, 1993; Greenberg, 1990). Gilliland's (1993) model of applicant reactions delineates ten procedural fairness rules that fall under three broader categories, including formal characteristics of the selection process, explanations offered in the process, and interpersonal treatment. Formal characteristics encompass job relatedness, opportunity to perform, opportunity for reconsideration, and consistency of administration (Gilliland, 1993). Explanation offered during the selection process, the second category, consists of feedback, selection information, and honesty in treatment (Gilliland, 1993). The third category relates to interpersonal treatments given to applicants during the process, including interpersonal effectiveness of the administrator, two-way communication, and propriety of questions (Gilliland, 1993). The level of satisfaction or violation of these rules in the decision-making processes determines the degree of procedural fairness perceptions (Zhang et al., 2020).

Specifically, the rule of propriety of questions is derived from and in line with Leventhal's (1980) concept of bias suppression, which states that a decision-making procedure should ensure impartiality and prevent any favoritism (Morse et al., 2022). The rule refers to the degree to which "questions avoid personal bias, invasion of privacy, and illegality, and are deemed fair and appropriate" (Bauer et al., 2001). However, with the development of technologies and the evolution of the recruitment processes, the main bias sources do not merely exist in the questions per se. Given that algorithms typically generate predictions from past data and if there has been certain biases associated with high performers who serve as benchmarks, the algorithms will learn those patterns and perpetuate the biases (Lee and Shin, 2020; Li et al., 2021; Zhang et al., 2020). Instead, propriety of decision criteria, which refers to "the appropriateness of the basis for decision making, including biased standards and procedures" (Zhang et al., 2020), is considered a more important dimension of bias suppression.

The definitions of procedural fairness rules are presented in Table 2.1.

Formal characteristics of the selection process
Job relatedness refers to the extent to which a test either appears to measure content relevant to the job situation or appears to be valid.
Opportunity to perform pertains to having sufficient chance to demonstrate one's knowledge, skills, and abilities in the testing contexts.
Reconsideration opportunity refers to the opportunity to challenge or modify the decision-making process and the opportunity to review scores and scoring.
Consistency of administration refers to ensuring that decision procedures are consistent across people and over time.
Explanations offered during the selection process
Feedback pertains to whether informative feedback is provided in a timely manner.
Selection information (i.e., "information known"; Bauer et al., 2001) refers to whether information, communication, and explanation about the selection process are provided prior to testing.
Honesty (i.e., "openness"; Bauer et al., 2001) is defined as the extent to which communications are perceived as being honest, sincere, truthful, and open.
Interpersonal treatment
Interpersonal effectiveness of administrators (i.e., "treatment"; Bauer et al., 2001) refers to the degree to which applicants are treated with warmth and respect.
Two-way communication refers to the opportunity for applicants to offer input or to have their views considered during the test or in the selection process.
Bias suppression (derived from "propriety of questions") pertains to the extent to which questions and decision criteria are appropriate.

Table 2.1 Procedural fairness rules
(adapted from Bauer et al., 2001; Gilliland, 1993; Zhang et al., 2020)

2.2 Human-rated versus AI-rated AVIs

Current research findings on applicants' reactions to human-based interviews and AI-based ones are discrepant. On the one hand, job applicants perceive that employers value equality and novelty (Acikgoz et al., 2020; van Esch et al., 2021) as AI-based systems offer objective and consistent evaluations without involving personal biases (Black and van Esch, 2020). On the other hand, an organization that conducts AI-enabled interviews may signal placing low value on future employees

compared to an organization that spends time, money, and effort on human-led interviews (Acikgoz et al., 2020). Owing to the physical absence of a subject to interact with, it is intriguing to investigate the impact of additional absence of human factors in the decision-making process in asynchronous video interviews (AVIs). The primary differences in perceptions between human-rated and AI-rated AVIs relate to four dimensions; i.e., opportunity to perform, reconsideration opportunity, bias suppression, and selection information.

Firstly, the absence of human interactions and social presence can reduce applicants' perceptions of opportunity to perform. Impression management theory explores how individuals shape others' perceptions by highlighting positive qualities, concealing flaws, or creating false impressions (Leary and Kowalski, 1990). When human evaluations are completely excluded throughout the process, applicants will feel restricted in using impression management tactics to influence human judgment (Bonaccio et al., 2016; Levashina et al., 2014) and thus will perceive the interviews mediated by technology as less fair (Blacksmith et al., 2016). Also, there is a common belief that, when making judgments, the factors humans take into account are both more intuitive and more superficial than those utilized by algorithms (Hilliard et al., 2022). As a result, applicants may perceive that they are less able to manipulate how algorithms judge them and have less opportunity to perform when being evaluated by AI (Hilliard et al., 2022), consequently leading to lower fairness perceptions in AI-based situations.

Secondly, the perceptions of reconsideration opportunities can be diminished when the decision makers are not humans. Applicants may believe that AI is not capable of recognizing their uniqueness, and thus favor humans as the ones who make decisions (Kaibel et al., 2019; Lavanchy et al., 2023). Similarly, Lee (2018) suggests that applicants feel that algorithms are not able to discern good candidates because they can neither measure qualitative data nor make exceptions whereas humans can, thus resulting in distrust and the feeling of unfairness.

Thirdly, the employment of AI decision makers in AVIs is not necessarily associated with a lower perceptive level of bias. Generally, algorithms enhance the standardization of procedures in decision making so that the procedures can potentially be more objective, consistent, and less biased (Kaibel et al., 2019). However, AI-based decision-making process is not free of bias because machine learning algorithms can also replicate human biases if the software developers fail to pay close attention to the data or properly train and validate their algorithmic models (Caliskan et al., 2017; Zhang et al., 2020). Furthermore, Mirowska and Mesnet (2022) find that applicants are aware that biases present in traditional hiring contexts can be reproduced in AI-based evaluation processes. Therefore, employing AI to evaluate candidates and to make hiring decisions may not always be perceived as less biased.

Lastly, a lower perception of selection information can be caused by the employment of AI as the decision maker in AVIs. The majority of AI-based systems are proprietary and the algorithms that construct the systems are not available to the public (Scherer, 2015). Due to the lack of information, concerns about black box and explainability can intensify and impact user acceptance and fairness perceptions in a negative way (Mirowska and Mesnet, 2022). Moreover, as full AI delegation in personnel selection is still novel, the use of AI in AVIs is likely to increase uncertainty (Acikgoz et al., 2020) and anxiety about selection information, therefore decreasing applicants' fairness perceptions. Accordingly, the first hypothesis is proposed:

> *H1: AVIs rated by humans are perceived procedurally fairer than those rated by AI.*

2.3 Transparency

Despite the general skepticism surrounding technology-mediated interviews (Langer et al., 2017; Guchait et al., 2014), research (e.g., Truxillo et al., 2009) has shown that it is important to consider that various forms of explanations can significantly impact applicants' reactions to the selection process. Although full transparency in AI algorithms is still a significant challenge to achieve (Ananny and Crawford, 2018), providing partial explanations and the rationale behind algorithmic decisions or recommendations can enhance perceived transparency among users (Hunkenschroer and Kriebitz, 2022), leading to increased perceptions of legitimacy (Shin, 2021), higher level of trust (Glikson and Woolley, 2020; Kizilcec, 2016), and eventually higher level of fairness perceptions (Ribeiro et al., 2016; Shulner-Tal et al., 2023).

2.3.1 Types of transparency

The concept of transparency holds significant importance across various contexts and has been widely discussed in academic literature. However, a universally agreed definition of transparency in AI has not been established yet. Transparency is often associated with interpretability and explainability (Köchling and Wehner, 2020), and can be seen as the degree to which individuals understand how and why AI assesses and decides something and follows human rules and logic (Hoff and Bashir, 2015).

Interpretability may be more relevant for the recruiters than for the applicants. Specifically, the concept comprises three distinct levels, including simulatability, decomposability, and algorithmic transparency (Lipton, 2018). On a holistic level, simulatability pertains to the extent to which the entire model can be thoroughly understood and simulated. Decomposability refers to the ability of the components, such as parameters and computation of a model, to be intuitively explicated. Algorithmic transparency denotes the visibility of the factors that impact the learning

algorithms to individuals who utilize and regulate those algorithms. Fairness perception is affected by these three levels of interpretability as demands for fairness often lead to demands for understanding the processes, and subsequently interpretable algorithmic models (Lipton, 2018; Shulner-Tal et al., 2023).

Explainability and explainable results, in the context of recruitment, mainly focus on knowing the attributes that drive algorithmic decisions (De Fine Licht et al., 2014; Glikson and Woolley, 2020; Tambe et al., 2019; Yu and Li, 2022). To build equitable algorithms and AI applications, organizations not only need to focus on the model form (i.e., understand and explain how their AI models operate) but also have to carefully select samples (i.e., what data and criteria are used in making decisions; Corbett-Davies and Goel, 2018; Schumann et al., 2020). Low explainability often associates AI-involved decision-making processes with "black box". Black-box AI models are considered opaque, less transparent, and their inner operations cannot be explained, making it difficult for users to understand why the algorithms make the decisions in a certain way (Hunkenschroer and Kriebitz, 2022).

Nevertheless, extant literature indicates that explainability is often absent from the complex methods underlying state-of-the-art prediction algorithms (Tambe et al., 2019). Disclosing the conditions and providing qualitative explanations for each attribute and algorithmic decision are challenging given that complex algorithms learn from millions of data points and become too intricate to be entirely understood and explained, even by those who created them (Hunkenschroer and Luetge, 2022; Raghavan et al., 2020; Simbeck, 2019). Additionally, while the use of more data and more sophisticated algorithms increases predictive power and accuracy, it also becomes more arduous for people to understand and explain. At the current status of AI technologies, it remains not only unclear how to balance this trade-off (Tambe et al., 2019) but also technically challenging how to produce explainable results (Hunkenschroer and Kriebitz, 2022).

2.3.2 Transparency in decision maker

While the feasibility of establishing full transparency in AI is concerned (Hunkenschroer and Kriebitz, 2022), disclosing the decision maker is a practical action that organizations can take to increase informational transparency. In line with Gilliland's (1993) rule of selection information, information asymmetry can lead to lower perceptions of procedural fairness. However, hiring organizations usually do not explicitly disclose the use of AI during the recruitment process (van Esch et al., 2019) in fear of alienating potential candidates. Köchling et al. (2022) find that open communication about the use of AI in recruitment diminishes affective responses, which are individuals' evaluations, emotions, and attitudes to a stimulus (Zhang, 2013), especially in the early phases of the selection process. Similarly, Langer et al. (2021) discover that negative perceptions occur when candidates are informed about

the use of algorithmic tools due to privacy concerns. Nevertheless, a recent field study (Suen and Hung, 2023) finds that transparently informing the use of AI in digital interviews can increase applicants' cognitive trust, which reflects their beliefs about the reliability and dependability (McAllister, 1995). Cognition-based trust is determined by rational thinking (Glikson and Woolley, 2020) and can be activated by informing that AI algorithms are employed while simultaneously conveying objectivity (Suen and Hung, 2023). Taken together, the inconsistencies between the theories and empirical findings and the growing applications of AI in interview assessments make decision-maker transparency a vital topic needed to be discussed.

In addition to Gilliland's (1993) fairness model, signaling theory (Spence, 1973) and fairness heuristic theory (Lind, 2001) can offer additional perspectives to comprehend job applicants' reactions to the information about decision makers. Signaling theory suggests that, through recruitment activities, applicants form impressions from available information that serves as signals of latent attributes for them to make inferences about the organizations (Celani and Singh, 2011). For example, organizations that honestly and explicitly disclose the identity of the raters in the hiring processes may signal to the applicants that they place higher values on potential employees, compared with organizations that do not proactively communicate. Fairness heuristic theory explains how people utilize the notion of fairness as a cognitive shortcut to form a sense of security in social interactions and to steer their behaviors toward collaboration and compliance (Lind et al., 2001). The heuristic, shaped by early and significant events, serves as a lens through which individuals perceive and interpret organizational actions (McCarthy et al., 2017a). For example, the absence of information about the decision maker is likely to amplify the sense of uncertainty experienced by job applicants, speeding up fairness heuristics formation and a consequent negative impact on fairness perceptions. Accordingly, the following hypothesis is proposed:

> *H2: Transparency in decision makers positively impacts perceptions of procedural fairness in AVIs.*

2.3.3 Moderating effects of explanations

It is not surprising, according to some of the researchers (e.g., Shaw et al., 2003), that the provision of any kind of explanations are related to fairness, based on the common ground that the majority of them focus on increasing job applicants' perceptions (Truxillo et al., 2009). Moreover, to a significant extent, the provision of explanations is in accordance with Gilliland's (1993) procedural fairness rules. Despite the obvious direct effect of explanations on perceptions and indirect effect on applicants' actions thereafter, only scant research exists on how explanations interact with different decision makers in affecting applicants' fairness perceptions.

Transparency for the decision-making processes (i.e., explainability and explainable results) are arguably the most important from the applicants' standpoint, particularly when the applicants might get rejected based on unexplained or unknown reasons or based on criteria that are not sufficiently validated regarding job performance or relevant qualifications (Chamorro-Premuzic et al., 2016; Dattner et al., 2019; Kim, 2016; Raghavan et al., 2020). In a substantial amount of studies (e.g., Gilliland et al., 2001; Truxillo et al., 2002), such explanations not only focus on the procedure itself (i.e., what steps will be involved or what can be expected) but also emphasize how the selection procedure is related to the job (Truxillo et al., 2009).

As AVI by its nature is not synchronized (i.e., no interviewer to directly interact with and no back-and-forth communication; Lukacik et al., 2022), the candidates may feel that they have fewer chances to present themselves or to demonstrate their impression management tactics (Blacksmith et al., 2016; Köchling et al., 2022). Such feelings may be stronger in fully automated, no-human-involved processes, especially with lower perceptions of opportunity to perform and to manipulate the interview toward a positive outcome (Lee, 2018).

However, more detailed explanations may moderate these relationships. As Shaw et al. (2003) noted, explanations reveal "the reason for, or the cause of, some event that is not immediately obvious or entirely known." The level of knowledge about what the algorithms use to make judgments and the chance of appropriately explaining how AI evaluates applicants' performance will be increased (Acikgoz et al., 2020; Cheng and Hackett, 2021). Thus, interview takers have more insights into how to accordingly adjust their tactics while answering the questions. Also, applicants can benchmark their performance to the selection criteria and are thus able to have a more convincing rationale when trying to justify or challenge the processes.

In an empirical study, Basch and Melchers (2019) find that the perceived fairness is higher when messages emphasizing standardization are provided. Such messages relate to the rule of consistency and imply exclusion of biases and equivalent opportunities for each applicant to show their qualifications (Basch and Melchers, 2019). Although the use of AVI is essentially to standardize the processes and to increase the consistency across the interviews over time, neither consistency nor unbiasedness can be guaranteed when humans are the ones who make the decisions. While consistency may be possible in AI-based conditions, free of bias is not the case. Consequently, transparently revealing the selection criteria is believed to be as a means to show honesty, integrity, truthfulness, and openness (Jaser et al., 2022; Sánchez-Monedero et al., 2020), thus potentially easing the concerns and increasing fairness perceptions.

In sum, when the black-box issues are seemingly unsolved and understandable AI models are not available for the mass of job applicants (Zielinski, 2020), increasing transparency level and proactively providing more explanations of the "internal

operations", such as what attributes are driving the decisions (Raghavan et al., 2020; Simbeck, 2019; Tambe et al., 2019), are likely to have a moderating effect. Thus, the third hypothesis is proposed:

> *H3: In AVIs, the provision of explanations moderates the relationships between decision makers and procedural fairness perceptions.*

2.4 Summary of hypotheses

Figure 2.1 below provides an overview of the conceptual model.

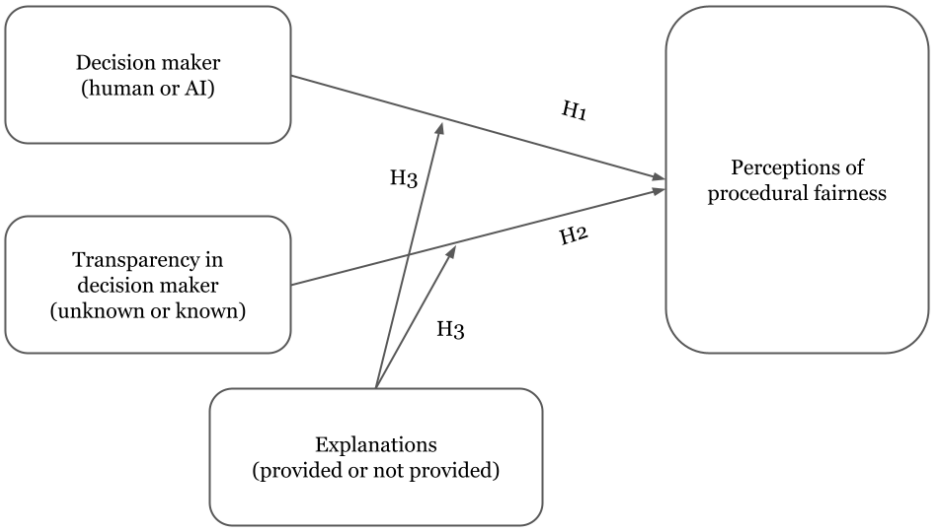


Figure 2.1 Conceptual model

To conclude, three hypotheses are generated to lead the empirical study and analysis, each related to and developed from previous literature.

Hypotheses
Hypothesis 1 *AVIs rated by humans are perceived procedurally fairer than those rated by AI.*
Hypothesis 2 *Transparency in decision makers positively impacts perceptions of procedural fairness in AVIs.*
Hypothesis 3 *In AVIs, the provision of explanations moderates the relationships between decision makers and procedural fairness perceptions.*

Table 2.2 Hypotheses overview

3. Methodology

In this chapter, the methods of the empirical research are elaborated. To begin with, the scientific approach to test the hypotheses is described. Next, a thorough description of the research design and the components in it are discussed. Thereafter, the process of the main study is presented, followed by discussions of measure reliability, validity, and replicability of the research.

3.1 Research approach

Research for job applicants' reactions has emerged since the 1980s (McCarthy et al., 2017a). The proliferation of solid theories, rigorous methods, and comprehensive measurement tools has established fundamental frameworks for subsequent researchers to explore the field. However, novel technologies in recruitment have continuously been employed and changed the processes of talent acquisition. To test the hypotheses associated with the state-of-the-art technologies, the authors considered it appropriate and valid to adopt a quantitative approach in this research. As Bell et al. (2019, p.35) describe, quantitative research is "a research strategy that emphasizes quantification in the collection and analysis of data and that entails a deductive approach to the relationship between theory and research, in which the emphasis is on testing the theories." In particular, Gilliland's (1993) organizational justice model, a classic and dominant model, serves as the primary model to study job applicants' fairness perceptions in this research.

3.2 Research design

3.2.1 Between-subject design

In the main study, the authors conducted an online between-subject experiment in which each participant was randomly assigned to only one treatment group. Particularly, simple random assignment was adopted, in which the probability of being assigned to the treatment group was identical for all subjects, thereby ensuring that the treatment status was statistically independent of the subjects' potential outcome and their background attributes (Gerber and Green, 2012, p.32). The aim was to infer the causality of the manipulations (independent variables, IVs) on the outcome (dependent variables, DVs). Under random group assignment, the effects of the manipulations can be established by comparing variations in the level of the DVs among the treatment groups (Bell et al., 2019, p.49; Charness et al., 2012).

In this research, the manipulations included the disclosure of decision makers (i.e., recruiting managers, AI, or unknown; Appendix A.3 and A.4) and binary conditions

of whether or not explanatory details were provided (Appendix A.5). Consequently, six treatment groups were created (Table 3.1). Each participant (N=318) was randomly assigned to one of the groups and answered the questionnaire. The results of our interests were fairness perceptions, which were measured with a seven-point Likert scale to investigate the clusters of attitudes.

One of the main reasons the authors preferred a between-subject to a within-subject design was to minimize the "demand effect", a spurious effect caused by the respondents' expectations to act "in accord with some pattern, or attempting to provide answers to satisfy their perceptions of the experimenter's expectations" (Charness et al., 2012). When the participants interpret the experimenter's intentions as to compare their perceptions between AI and humans as the decision maker, they may change their behavior accordingly, either consciously or unconsciously (Charness et al., 2012; Rosenthal, 1976; White, 1977). Moreover, the authors argue that the results are more generalizable in a between-subject design because job applicants are highly unlikely to be exposed to more than one of the conditions in the real world when they take an asynchronous video interview (AVI).

	Decision maker		
	Unknown	**Human**	**AI**
Explanations not provided	Treatment group 1	Treatment group 3	Treatment group 5
Explanations provided	Treatment group 2	Treatment group 4	Treatment group 6

Table 3.1 Manipulations and treatment groups

3.2.2 Questionnaire

The format adopted was an online self-completion questionnaire, in which the respondents answered questions by completing the questionnaire themselves (Bell et al., 2019, p.232). To a certain extent, this format ensured participants' anonymity and allowed them to answer the questionnaire without being observed at site, thus reducing socially desirable responses and behaviors (Steenkamp et al., 2010).

3.2.2.1 Material

The authors started the survey by portraying a scenario in which the respondents received an invitation to take an interview. Next, all participants were provided with

the same description of the interview guidance and privacy rules. Subsequently, relevant instructions were given based on the manipulations for different treatment groups. Precisely, the disclosure of the decision maker was firstly manipulated, followed by the second manipulation of whether explanations were provided. Rather than indicating specific technical skills, the explanations for selection criteria were deliberately described in general terms (e.g., leadership competencies) both to avoid restricting participants to certain roles and to increase generalizability. Lastly, a simulation of AVI interface was presented. Such scenario-based approach has been commonly employed in social psychology and ethics research to investigate individuals' beliefs, attitudes, and opinions (Lee, 2018; Petrinovich et al., 1993). The material is included in Appendix A.

Additionally, participants in the AI-based conditions (treatment group 5 and 6) were presented with the definition of AI at the beginning of the survey. The definition was adopted from Kaplan and Haenlein (2019):

> *"The ability of a machine to perform cognitive functions that we associate with human minds, such as perceiving, reasoning, learning, interacting with the environment, problem-solving, decision-making, and even demonstrating creativity."*

The definition was provided to ensure that participants had a similar understanding in mind and to reduce the potential impact of discrepancies in their perceptions of AI on the manipulations being studied.

3.2.2.2 Scale

Firstly, to measure overall fairness perceptions on the simulated interview process, the item "I think the shown procedure was fair", derived from Warszta (2012), was presented in the questionnaire. As aforementioned, Gilliland's (1993) model consists of various dimensions that could influence the overall justice perception; however, some researchers (e.g., Greenberg, 2001; Jones and Martens, 2009) indicate that the overall perception is formed as an aggregation of these dimensions and impacts the recruitment outcome (e.g., applicants' tendency to accept the offer or to recommend the firm to others) more than a single justice dimension (Warszta, 2012). Following the overall fairness item, narratives in accordance with the fairness rules (e.g., "I could really show my skills and abilities through this interview" for opportunity to perform) were given. The inclusion of fairness subscales allowed for the possibilities to elucidate the underlying factors associated with the disparities in the overall procedural fairness perceptions. The narrative items were derived from the Selection Procedural Justice Scale (SPJS; Bauer et al., 2001), a commonly used scale that was designed to measure respondents' fairness perceptions based on Gilliland's model. The participants were asked to select from 1 (strongly disagree) to 7 (strongly agree).

While keeping the maximal validity and applicability of SPJS, the authors made some minor adjustments. Firstly, all the words "test" in the questionnaire were replaced with "interview" to fit the context of this study. Secondly, while some narratives were designed with room for customization (e.g., "A person who scored well on this test will be a good [*insert job title*]"), the authors aimed to avoid the pre-existing attitudes toward a specific job and therefore revised the narrative in a neutral manner (e.g., "A person who scored well in this interview will perform well in his/her role").

The authors deliberately excluded the procedural fairness rules concerning interpersonal treatments (i.e., treatment at the site and two-way communication) and feedback to reflect on the factual absence of these components in the AVI setting. Additionally, due to the evolution of recruitment technologies, the concept of bias in the recruitment process is no longer limited to the propriety of questions but including the propriety of decision criteria (Zhang et al., 2020). Therefore, the procedural fairness rule of "propriety of questions" was conceptually reframed as "bias suppression" and the corresponding self-developed item "The selection process was objective and without bias" was included in the questionnaire (Appendix B.1).

To increase the readability and minimize the dropout rate, the narratives were intentionally selected to be short and concise, with the longest description being 15 words. Also, one attention check, where the respondents were asked to select a designated option, was included in the questionnaire to filter out inattentive respondents and to assure the quality of our data.

Key demographic information, including age, gender, and education level, about the participants was collected. These three variables have been continuously investigated and proved to be associated with different perceptions in recent AI research (e.g., Araujo et al., 2020; Van Berkel et al., 2021; Wang et al., 2020). Finally, the participants in the AI-based conditions were asked to self-identify their knowledge of AI, from 1 (not knowledgeable at all) to 5 (extremely knowledgeable) at the end of the questionnaire. Domain knowledge in AI was deemed interesting and relevant as extant research has found contrary results. A higher level of AI and programming knowledge could be associated with both lower levels of perceived fairness (e.g., Lee and Baykal, 2017) and higher levels of perceived fairness (e.g., Araujo et al., 2020). The demographic items are shown in Appendix B.2.

3.2.3 Measurement

Concepts either provide an explanation of a certain aspect of the real world or represent the things that we seek to explain (Bell et al., 2019, p.168). In this quantitative research, fairness perceptions are the key concepts employed and thus

have to be measured. In scientific research, measures refer to things that can be "relatively unambiguously counted" (i.e., quantities; Bell et al., 2019, p.169), such as age or salary. However, fairness perceptions are less directly quantifiable. Therefore, indicators are needed to stand for this concept (Bell et al., 2019, p.168). In order to validly measure the intangible concept, a 7-point Likert scale was adopted as the indicator in the self-completion questionnaire in this research, thus creating the possibility to analyze the effects of the manipulations on the outcome.

Prior research suggests that applicant reactions are not static and can vary depending on when they are measured in the selection process (McCarthy et al., 2017a). To eliminate the effects of measurement timing, the authors decided to contrive the questionnaire to post-AVI conditions. Therefore, the validity of the measurement, including the representation of the 7-point scale to the non-quantifiable concept of fairness and the timing of measurement, toward the results of our interest is guarded.

3.2.4 Pilot test

To minimize ambiguities and assure the understandability of the survey, the authors conducted a pilot test before publishing the survey. Six students at Stockholm School of Economics (SSE) were randomly approached and, after consent, were shown the scenarios and the questionnaire. Also, they were encouraged to identify anything unclear or not understandable. All of the samples had taken AVIs and thus were able to understand the scenarios portrayed. Some according changes were made afterward with the care taken not to affect the validity or generalizability of the scale. For instance, given the fact that the authors simulated the interview process instead of conducting a field study, one of the narratives in past tense (i.e., Applicants were able to have their test results reviewed if they wanted) was revised to a scenario-based presumption (i.e., Applicants would be able to have their interview results reviewed if they wanted) in accordance with our research method.

3.3 Main study

3.3.1 Data collection

Qualtrics, a web-based survey tool, was used as the primary tool for collecting responses. The participants were recruited both from digital social platforms such as Facebook and from offline places, mainly in SSE atrium. The surveys were completed in 253 seconds on average. The data was collected for the time span of eleven days, specifically from March 27, 2023 to April 7, 2023.

Ten days after the survey was published (i.e., on April 6, 2023), 316 responses were collected. A power analysis (Magnusson, n.d.) was subsequently performed to determine how many more participants were needed based on the effect sizes (Cohen's d) at that time. Given the effect exists, power is the ability to detect it. Given a predetermined power, more samples are needed to detect a smaller effect size. Likewise, given the same sample size and significance level, the power is higher when the effect size is bigger and is easier to be detected. At the typical significance level of 0.05, to have 80% power to correctly detect an effect size of 0.33, the largest effect size among the subscale items at that time, a minimum sample size of 72 was needed to be reached. Namely, we needed to recruit at least 116 more participants (72 * 6 groups - 316 samples we already had), on the condition that the effect size would not decrease when more responses were collected. However, the authors decided not to keep recruiting respondents due to the limited timeframe of the study and the uncertainties in whether the effect size would remain. Hence, the survey was closed the next day.

3.3.2 Participants

318 people responded. After participants who failed the attention check (N=30) were omitted, 288 participants remained in the sample. The participants were fairly equally distributed in the treatment groups (Table 3.2). Their ages ranged from 18 to 47, with an average age of 24.53 (SD=3.46). The majority of the participants were female (61.8%), followed by male (36.5%). To a remarkable degree, they were fairly educated as 95.2% of the participants had their education at either bachelor's or master's level. In the AI-based conditions (N=97), the mean of self-identified AI knowledge score was 2.76 (SD=0.70). The demographic distribution of the participants in this study is presented in Table 3.3.

Treatment group	N
1. Unknown - Explanations not provided	45
2. Unknown - Explanations provided	53
3. Human - Explanations not provided	47
4. Human - Explanations provided	46
5. AI - Explanations not provided	47
6. AI - Explanations provided	50

Table 3.2 Treatment group distribution

Gender		Education level		AI knowledge	
Male	N=105 (36.5%)	High school	N=8 (2.8%)	1	N=2 (2.1%)
Female	N=178 (61.8%)	Bachelor's degree	N=90 (31.3%)	2	N=32 (33%)
Non-binary	N=2 (0.7%)	Master's degree	N=184 (63.9%)	3	N=50 (51.5%)
Prefer not to say	N=3 (1%)	Other	N=5 (1.7%)	4	N=13 (13.4%)
		Prefer not to say	N=1 (0.3%)	5	N=0 (0%)

Note:
AI knowledge: 1=Not knowledgeable at all, 2=Slightly knowledgeable, 3=Moderately knowledgeable, 4=Very knowledgeable, 5=Extremely knowledgeable.

Table 3.3 Demographics distribution

3.3.3 Analytical tools

The use of web-based surveys allowed the answers to be automatically programmed and to be downloaded into a database, an advantage that eliminated the daunting coding of questionnaires (Bell et al., 2019, p.241) and potential manual errors. Subsequently, the database was retrieved and programmed in RStudio, where the data was processed with programming language R. The analyses were performed with independent samples *t*-tests, one-way analysis of variance, Pearson's chi-squared tests, and moderated hierarchical multiple regressions. The confidence interval of 95% was used for all tests. For the purposes of examination and result replication, the R scripts (without output) are appended at the end of this thesis. The full R scripts and codes including the output are provided in a separate file and can be assessed via an URL in Appendix D.

3.4 Research quality

3.4.1 Reliability of measures

Reliability deals with the question of whether the results of a study are repeatable (Bell et al., 2019, p.46) and is fundamentally concerned with stability and consistency of measures (Bell et al., 2019, p.172). Specifically, three types of reliability, including stability, internal reliability, and inter-rater reliability, are discussed.

Firstly, stability entails the question of whether or not a measure is stable over time so that the results relating to the measure for the sample do not fluctuate (Bell et al., 2019, p.172). Although the test-retest method is the most obvious way to examine the stability of a measure (Bell et al., 2019, p.173), the method is not feasible to be taken within the time span of the study. Instead, the authors employed Bauer et al.'s (2001) SPJS, a comprehensive measure of fairness perceptions, as the measurement in this research. The scale was developed under psychometric procedures for scale development (Hinkin, 1998). Moreover, SPJS is considered the "gold standard" for assessing candidate reactions and is believed to be versatile enough in any personnel selection setting without the need for major adjustments (Butucescu et al., 2019).

Secondly, internal reliability determines whether a multiple-item measure is coherent and measures the same intended variable (Bell et al., 2019, p.173). A Cronbach's alpha coefficient of 0.7 is suggested to be an efficient level of internal reliability (Schutte et al., 2000). All measures developed by Bauer et al. (2001) showed adequate internal reliability, among which the minimum was 0.73 (job-relatedness content) and the maximum was 0.92 (treatment). Nonetheless, not all the items in SPJS fitted the scenario in this research and, in fact, only one item in each subscale was adopted in the questionnaire. As a result, each fairness rule was measured and represented by one item, thus making the correlation level of each item not available and internal reliability tests not applicable.

Lastly, inter-rater reliability entails the inconsistency issue when more than one rater is involved in the recording of subjective judgment observations or translations (Bell et al., 2019, p.172). The issue did not exist in this research as a pre-developed questionnaire was employed to measure and record the outcome.

3.4.2 Validity

Validity is concerned with the integrity of conclusions that are generated from the research (Bell et al., 2019, p.46). In particular, four main types of validity are discussed, including measurement validity, internal validity, external validity, and ecological validity (Bell et al., 2019, pp.46-47).

Firstly, measurement validity refers to the extent to which a measure captures the concept that it is intended to capture (Bell et al., 2019, p.46). As aforementioned, the measurements in this study were adopted from well-established SPJS, which has served as the foundation for a wide range of studies and is extensively used as an instrument to test Gilliland's (1993) procedural justice rules (McCarthy et al., 2017a). As Hinkin (1998) discussed, convergent and divergent validity are both important to establish measurement validity. Both convergent and divergent validity of the instruments were psychometrically tested and the validity evidence was

demonstrated in the development of SPJS (Bauer et al., 2001), indicating that the use of the SPJS factors are related to the measurement of Gilliland's justice model.

Secondly, internal validity mainly refers to causality, which deals with the question of whether IV is responsible for the variation in DV rather than something else (Bell et al., 2019, p.47). In a between-subject experimental design, where participants are assigned to only one treatment group, it is crucial to protect the participants against interference, which means whether a subject is treated only depending on its own treatment group condition and the assignments have no bearing effects on whether the other subjects receive the treatment (Imbens and Rubin, 2015). The survey tool used in this study entirely randomized the participants in different treatment groups and the self-completed questionnaire assured the responses were independent of one another. Internal validity was increased in these non-interfered conditions in the experiment, increasing the confidence that the manipulations were responsible for the variation in the outcome.

Thirdly, external validity refers to whether the findings of the research can be generalized beyond the context where the research was conducted and beyond the cases that make up the sample (Bell et al., 2019, p.177). One of the common questions regarding generalizability requires the sample to be as representative as possible (Bell et al., 2019, p.177). Although the data was collected mainly based on convenience sampling method that may constrain the generalizability of this study, the sample with mean age of 24.53 (SD=3.46) is arguably generalizable to the job applicants who need to take video interviews for relatively junior roles.

Finally, ecological validity is concerned with whether the research findings can be applied to people's everyday, natural, and social settings (Bell et al., 2019, p.47). In this research, the unnaturalness of having to answer the questionnaire rather than being naturally observed can potentially limit ecological validity of the findings. Such constraint inherently exists in the method of using a questionnaire. However, the authors simulated the video interview process by providing visualization of an interview interface in the survey, based on real-world situations, with the intention to mimic the scenarios in the real world, thus arguably increasing ecological validity.

3.4.3 Replicability

Replicability is considered an important quality factor in quantitative research (Bell et al., 2019, p.180). The concept of replicability revolves around the explicitness of methods used, the procedures taken to generate the findings, and the possibility to replicate a piece of research (Bell et al., 2019, p.178). In accordance with replicability, the theories, methods, instruments, materials, procedures, and measurements are explicitly described in this research. A clear description of data analysis procedures is included. Statistical inference, on which the results depend, as well as uncertainties

of them are discussed. All in all, replication is allowed by following the empirical steps and the replicability of the study is ensured.

4. Results and analyses

In this chapter, the results of this research are presented, starting with the data of the experimental study. Next, both the analytical steps and the statistical methods are described, followed by the results of hypothesis testing. At the end, a further analysis is conducted to gain more insights into the sample and the results.

4.1 Descriptive statistics

The means, standard deviations, and correlations among the study variables are presented in Table 4.1. It can be seen that the correlations between the overall fairness item and the procedural fairness subscales were all highly significant and in the same direction. Also, among procedural fairness perception facets, all the significant correlations were positive (i.e., two variables move in the same direction), indicating that these subscales could potentially be used as a multi-item measure to coherently measure the same intended variable (i.e., overall fairness perception). However, the Cronbach's alpha of the seven subscale items was 0.638, which did not reach the efficient internal consistency threshold of 0.7 (Schutte et al., 2000). The result indicated that the subscales could not be computed into an index to represent participants' overall perceptions of fairness. Therefore, one overall and seven subscale items were independently employed in the following hypothesis testing steps and analyses.

In addition, the demographic variables (i.e., gender, age, education level, and AI knowledge) appeared to be significantly correlated with part of the items. Correlations between age and overall fairness perceptions ($r=-0.15$, $p=0.01$), job relatedness ($r=-0.14$, $p=0.02$), honesty ($r=-0.23$, $p<0.001$), and bias suppression ($r=-0.20$, $p<0.001$) were significant and all in negative directions. In contrast, AI knowledge was observed to be positively correlated with reconsideration opportunity ($r=0.20$, $p=0.05$) and consistency of administration ($r=0.22$, $p=0.03$). However, the correlations of gender and education level with the items were not considered valid because these two variables were dummy coded (see note in Table 4.1 and their distribution in Table 3.3). Further analysis and discussions about the participants' characteristics are included in Section 4.5.

Variables	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1. Known	-														
2. Decision maker	0.87	-													
3. Explanation	-0.03	-0.02	-												
4. Overall procedural fairness	-0.02	-0.05	0.00	-											
5. Job relatedness	0.00	-0.04	0.01	0.42***	-										
6. Opportunity to perform	0.02	0.00	0.03	0.33***	0.51***	-									
7. Reconsideration opportunity	0.00	0.02	-0.03	0.13*	0.12*	0.16**	-								
8. Consistency of administration	0.03	0.03	0.07	0.21***	0.06	0.09	-0.03	-							
9. Selection information	-0.04	0.00	-0.02	0.16**	0.17**	0.27***	-0.11	0.16**	-						
10. Honesty	-0.07	-0.08	0.00	0.39***	0.43***	0.39***	0.07	0.08	0.30***	-					
11. Bias suppression	0.05	0.07	0.02	0.33***	0.40***	0.32***	0.10	0.25***	0.10	0.41***	-				
12. Gender	0.03	0.03	-0.05	-0.15*	-0.15*	-0.08	0.12*	-0.07	-0.06	0.00	0.02	-			
13. Age	0.01	0.00	0.00	-0.15*	-0.14*	-0.09	0.01	0.04	-0.01	-0.23***	-0.20***	-0.09	-		
14. Education level	-0.05	-0.05	0.00	-0.04*	-0.05	-0.06	0.12*	-0.01	-0.04	-0.03	-0.11	0.07	0.16**	-	
15. AI knowledge	-	-	0.03	0.12	-0.13	-0.08	0.20*	0.22*	0.06	-0.02	-0.03	0.07	0.03	0.08	-
Mean (SD)	-	-	-	5.00 (1.40)	3.86 (1.49)	3.77 (1.50)	5.02 (1.62)	5.52 (1.25)	4.99 (1.43)	4.94 (1.38)	4.01 (1.56)	-	24.53 (3.46)	-	2.76 (0.70)

Note:
"Known" is coded as 0=unknown (treatment groups 1, 2) and 1=known (treatment groups 3, 4, 5, 6).
"Decision maker" is coded as 0=unknown (treatment groups 1, 2), 1=human (treatment groups 3, 4), and 2=AI (treatment groups 5, 6).
"Explanation" is coded as 0=explanations not provided (treatment groups 1, 3, 5) and 1=explanations provided (treatment groups 2, 4, 6).
"Gender" is coded as 1=male, 2=female, 3=non-binary, and 4=prefer not to say.
"Education level" is coded as 1=high school, 2=bachelor's degree, 3=master's degree, 4=other, and 5=prefer not to say.
Significance levels: * $p<0.05$, ** $p<0.01$, *** $p<0.001$.

Table 4.1 Means, standard deviations, and correlations among the variables

4.2 Potential covariate analysis

Differences in perceptions due to participants' demographics among treatment groups cannot be interpreted as causal effects (Suen et al., 2019). Therefore, prior to testing the hypotheses, a series of analysis of variance (ANOVA) and chi-squared tests were conducted to evaluate the comparability of the experimental groups and to examine whether there were demographic differences that should be controlled in the analyses. The results are shown in Table 4.2. These analyses did not find any statistically significant factors across the treatment groups. As a result, these demographics were not treated as covariates in the subsequent analyses.

	Gender	**Age**	**Education level**	**AI knowledge**
One-way ANOVA test	F(5,282)=0.39, p=0.86	F(5,281)=0.13, p=0.99	F(5,282)=0.22, p=0.95	F(1,95)=0.06, p=0.81
Pearson's chi-squared test	X^2=8.665, df=15, p=0.89	X^2=100.39, df=95, p=0.33	X^2=13.282, df=20, p=0.86	X^2=4.048, df=3, p=0.26

Significance levels: * p<0.05, ** p<0.01, *** p<0.001.

Table 4.2 Results of potential demographic covariate analysis

4.3 Hypothesis testing

4.3.1 Types of decision maker

The first hypothesis stated that asynchronous video interviews (AVIs) rated by humans would be perceived procedurally fairer than those rated by AI. To test the hypothesis, independent samples *t*-tests (Welch's *t*-tests) were conducted to compare treatment group 3 (human decision maker and explanations not provided) and treatment group 5 (AI decision maker and explanations not provided) in terms of both overall and subscales of procedural fairness perceptions.

As shown in Table 4.3, there were no significant differences observed in the seven procedural fairness subscale items. Although AVIs rated by humans were favored over those rated by AI with regard to job relatedness (t(92)=0.82, p=0.41, d=0.17), opportunity to perform (t(92)=0.54, p=0.59, d=0.11), consistency of administration (t(85)=0.70, p=0.48, d=0.14), honesty (t(90)=0.54, p=0.59, d=0.11), the differences were not statistically significant. In terms of selection information (t(84)=-1.67, p=0.10, d=0.34) and bias suppression (t(92)=-0.06, p=0.95, d=0.01), AI raters were perceived as more desirable than human raters in AVIs, but the differences were still not statistically significant. Furthermore, these two types of decision makers were

considered to provide insignificantly equivalent reconsideration opportunities (t(86)=0, p=1.00, d=0.00) in an AVI scenario.

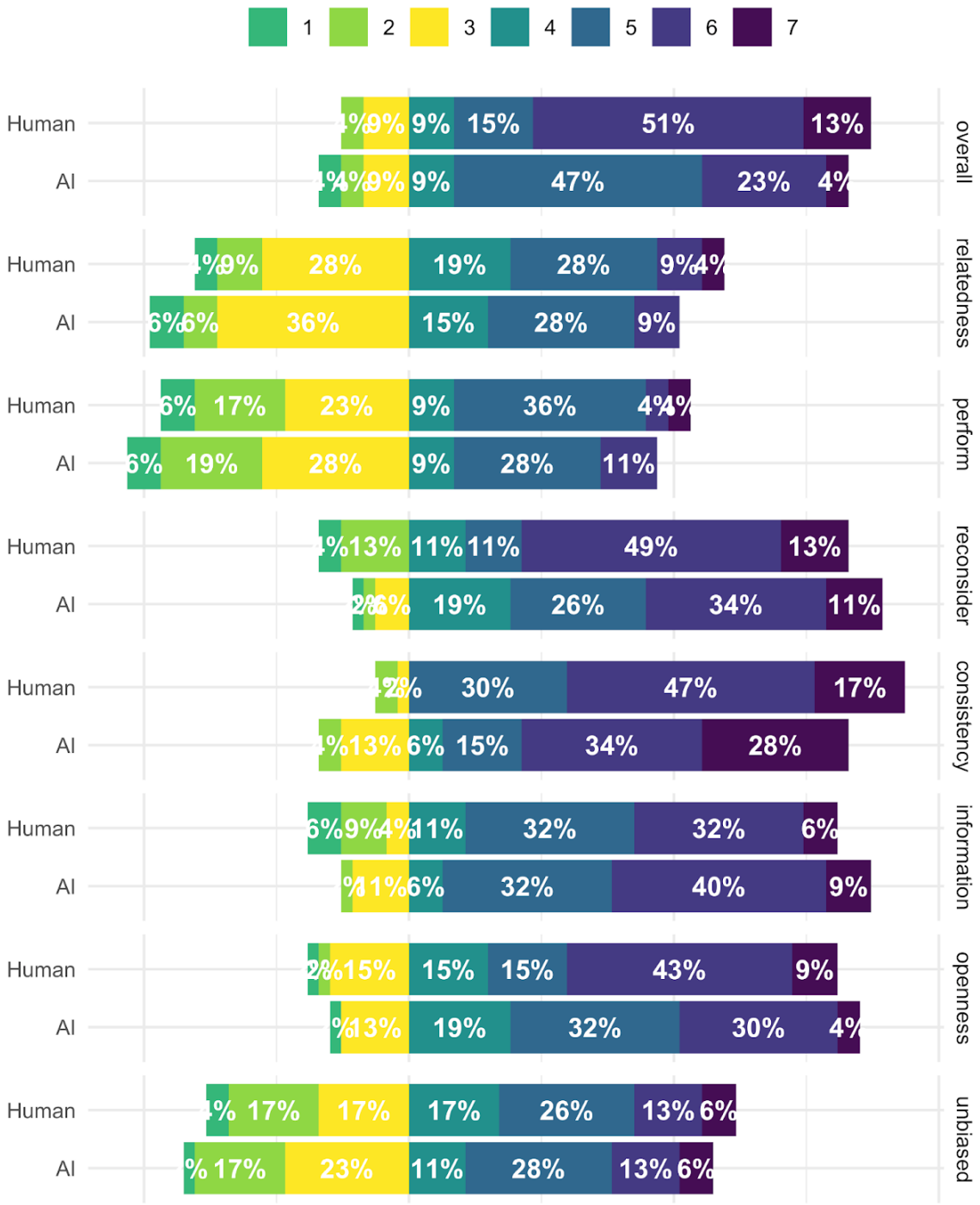
However, the *t*-test analysis demonstrated that, for the overall fairness item, there was a significant difference favoring human-rated AVIs (M=5.38, SD=1.31) over AI-rated ones (M=4.77, SD=1.37), t(92)=2.23, p=0.03, d=0.46. The inconsistency in the results of overall and individual procedural fairness measures may suggest that the procedural fairness subscales developed by Gilliland (1993) cannot fully explain applicants' perception discrepancies between AI and human in modern context (more details will be discussed in the next chapter). Therefore, H1 was still supported.

	Types of decision maker (explanations not provided)		
Measures	**Human** (N=47) Mean (SD)	**AI** (N=47) Mean (SD)	
Overall procedural fairness	5.38 (1.31)	4.77 (1.37)	t(92)=2.23, p=0.03*, d=0.46
Job relatedness	4.00 (1.43)	3.77 (1.34)	t(92)=0.82, p=0.41, d=0.17
Opportunity to perform	3.81 (1.56)	3.64 (1.50)	t(92)=0.54, p=0.59, d=0.11
Reconsideration opportunity	5.09 (1.73)	5.09 (1.33)	t(86)=0, p=1.00, d=0.00
Consistency of administration	5.64 (1.11)	5.45 (1.50)	t(85)=0.70, p=0.48, d=0.14
Selection information	4.74 (1.62)	5.23 (1.18)	t(84)=-1.67, p=0.10, d=0.34
Honesty	5.00 (1.44)	4.85 (1.23)	t(90)=0.54, p=0.59, d=0.11
Bias suppression	4.06 (1.62)	4.09 (1.59)	t(92)=-0.06, p=0.95, d=0.01

Significance levels: * p<0.05, ** p<0.01, *** p<0.001.

Table 4.3 People's perceptions of human versus AI decision maker

A visualized result distribution of these two groups (i.e., human versus AI decision maker) is provided in Figure 4.1.



Note:
Overall=Overall procedural fairness. Relatedness=Job relatedness. Perform=Opportunity to perform. Reconsider=Reconsideration opportunity. Consistency=Consistency of administration. Information=Selection information. Openness=Honesty. Unbiased=Bias suppression.

Figure 4.1 Distribution of survey results (human versus AI)

4.3.2 Transparency in decision maker

The second hypothesis proposed that transparency in the decision maker would positively impact perceptions of procedural fairness in AVIs. In order to test this hypothesis, independent samples *t*-tests (Welch's *t*-tests) were conducted to compare the conditions when the decision makers were unknown (treatment group 1) with the ones when the decision makers were known (treatment groups 3 and 5), both with no additional provision of explanations, across all the items. The results (Table 4.4) showed that, although conditions in which the decision makers were disclosed (M=5.07, SD=1.37) were perceived fairer than non-disclosed conditions (M=4.84, SD=1.54), the difference was not statistically significant, t(78)=-0.85, p=0.40, d=0.16.

Furthermore, neither were significant differences found in the procedural fairness subscales. Given no statistically significant differences, decision-maker transparency was favored in applicants' perceptions of job relatedness (t(75)=-0.53, p=0.60, d=0.10), reconsideration opportunity (t(83)=-0.14, p=0.89, d=0.03), consistency of administration (t(85)=-1.32, p=0.19, d=0.24), and bias suppression (t(89)=-1.05, p=0.30, d=0.19). In contrast, in the subscales of opportunity to perform (t(87)=0.12, p=0.91, d=0.02), selection information (t(95)=0.41, p=0.68, d=0.07), and honesty (t(98)=0.14, p=0.89, d=0.02), the conditions without specifying the decision makers were slightly more favorable in the AVI context, but these differences were still not statistically significant.

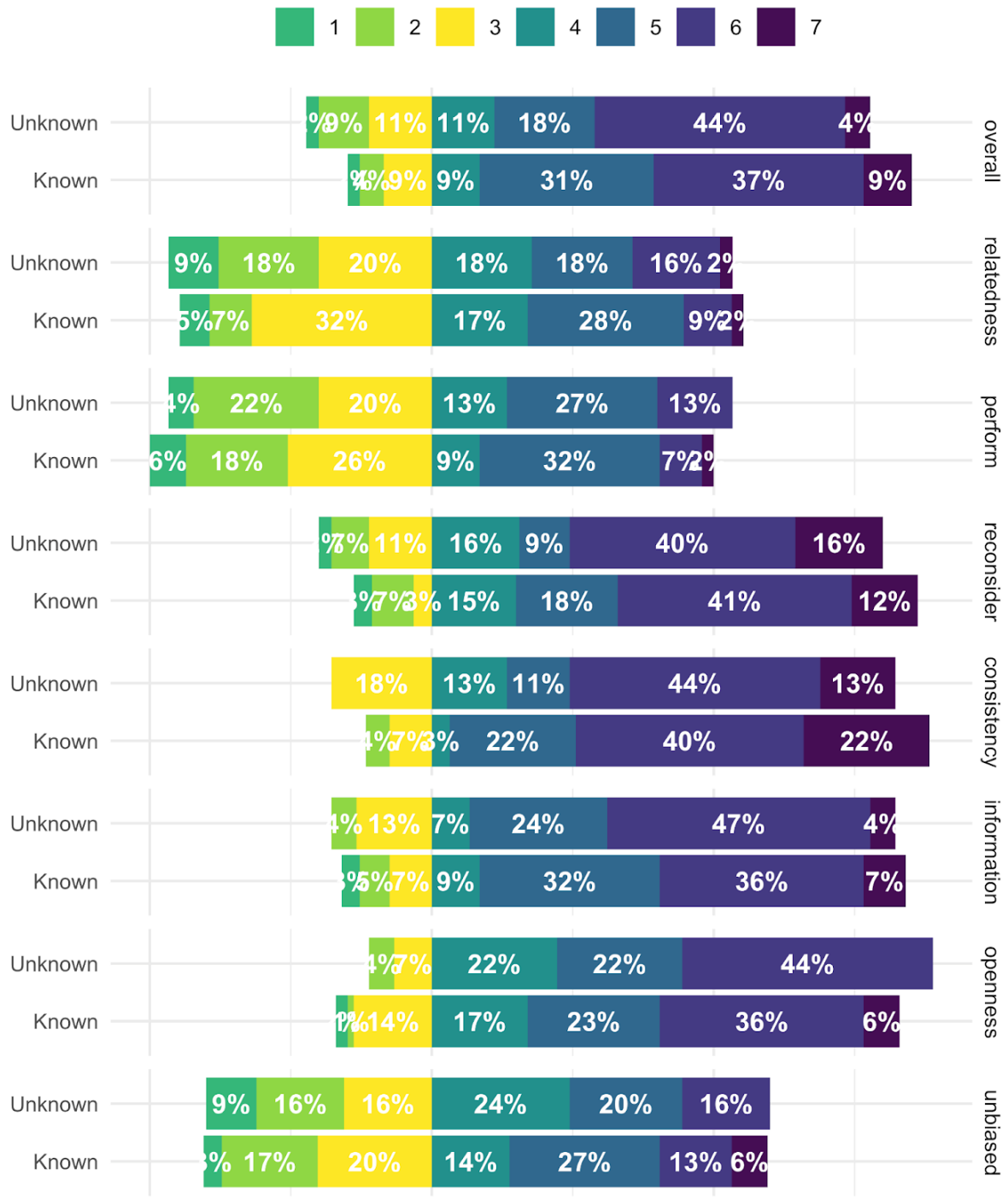
Overall, the results indicated that, in the context of AVI, there were no significant differences in perceived procedural fairness between the conditions in which the decision makers were disclosed or were not disclosed. In other words, transparent decision-maker identity did not affect job applicants' procedural fairness perceptions. Therefore, H2 was not supported.

	Types of decision maker (explanations not provided)		
Measures	**Unknown** (N=45) Mean (SD)	**Known** (N=94) Mean (SD)	
Overall procedural fairness	4.84 (1.54)	5.07 (1.37)	t(78)=-0.85, p=0.40, d=0.16
Job relatedness	3.73 (1.64)	3.88 (1.38)	t(75)=-0.53, p=0.60, d=0.10
Opportunity to perform	3.76 (1.51)	3.72 (1.52)	t(87)=0.12, p=0.91, d=0.02
Reconsideration opportunity	5.04 (1.62)	5.09 (1.54)	t(83)=-0.14, p=0.89, d=0.03
Consistency of administration	5.22 (1.35)	5.54 (1.32)	t(85)=-1.32, p=0.19, d=0.24
Selection information	5.09 (1.29)	4.99 (1.43)	t(95)=0.41, p=0.68, d=0.07
Honesty	4.96 (1.17)	4.93 (1.34)	t(98)=0.14, p=0.89, d=0.02
Bias suppression	3.78 (1.55)	4.07 (1.59)	t(89)=-1.05, p=0.30, d=0.19

Significance levels: * p<0.05, ** p<0.01, *** p<0.001.

Table 4.4 People's perceptions of unknown versus known decision maker

Figure 4.2 shows the distribution of survey data of these two conditions (i.e., unknown versus known decision maker).



Note:
Overall=Overall procedural fairness. Relatedness=Job relatedness. Perform=Opportunity to perform. Reconsider=Reconsideration opportunity. Consistency=Consistency of administration. Information=Selection information. Openness=Honesty. Unbiased=Bias suppression.

Figure 4.2 Distribution of survey results (unknown versus known)

4.3.3 Moderating effects of explanations

Hypothesis 3 suggested that the provision of explanations (moderating variable, MV) would moderate the relationships between the decision makers and procedural fairness perceptions. To test this hypothesis and examine the moderating effects, moderated hierarchical multiple regressions were employed. In the first step, procedural fairness measures were defined as the dependent variables (DVs) and the decision maker and the provision of explanations as the independent variables (IVs). In the second step, the respective cross-products (i.e., interaction term) of decision makers and explanations were added to fit the categorical by categorical interaction model. Both IV and MV were dummy coded because they were categorical variables. Also, treatment group 1 (unknown decision maker without explanations) was set as the reference group (i.e., baseline) in these regressions. The results are presented in Table 4.5.

In the first step, the provision of explanations was not a significant predictor of the procedural fairness variables in any regression model. In the second step, however, the interaction term of human and explanation (human:explain) was significant in predicting the overall fairness perception item (β=-0.92, p=0.02) and was marginally significant in predicting the consistency of administration item (β=-0.70, p=0.05). Specifically, the interaction terms inferred the moderating effects including (1) the human effect (human - unknown) in the explained condition versus the human effect in the unexplained condition, and (2) the explain effect (explained - unexplained) for human decision maker versus the explain effect for unknown decision maker. These negative coefficients indicated that the perceptions of overall fairness and the perceptions of consistency were lower when the decision maker was human with the explanations provided. Visualization of moderating effects is shown in Figure 4.3.

Also, marginally significant simple effects of human decision makers in predicting overall fairness perceptions (β=0.54, p=0.07) and of explanations in predicting the consistency item (β=0.48, p=0.06) were observed. The former referred to the predicted difference in overall fairness perceptions between human and unknown decision makers in unexplained conditions, and the latter indicated that in a scenario when the decision maker was unknown, providing explanations could lead to higher perceptions of administrative consistency.

Except for the ones mentioned, all the other interaction terms were not statistically significant in predicting either overall procedural fairness perceptions or any other subscales, indicating that the provision of extra explanations did not moderate the relationships between the decision makers and applicants' procedural fairness perceptions. Thus, hypothesis 3 was only partially supported.

	Overall procedural fairness			**Job relatedness**			**Opportunity to perform**			**Reconsideration opportunity**		
	β	R^2	ΔR^2	β	R^2	ΔR^2	β	R^2	ΔR^2	β	R^2	ΔR^2
Step1		0.005			0.006			0.004			0.004	
(Intercept)	5.03***			3.84***			3.68***			5.08***		
AI	-0.17			-0.14			-0.01			0.09		
Human	0.07			0.12			0.17			-0.11		
Explain	-0.01			0.05			0.08			-0.11		
Step2		0.025	0.020		0.008	0.002		0.005	0.001		0.006	0.002
(Intercept)	4.84***			3.73***			3.76***			5.04***		
AI	-0.08			0.03			-0.12			0.04		
Human	0.54†			0.27			0.05			0.04		
Explain	0.34			0.25			-0.06			-0.04		
AI:explain	-0.15			-0.33			0.20			0.10		
Human:explain	-0.92*			-0.27			0.23			-0.30		
	Consistency of administration			**Selection information**			**Honesty**			**Bias suppression**		
	β	R^2	ΔR^2	β	R^2	ΔR^2	β	R^2	ΔR^2	β	R^2	ΔR^2
Step1		0.005			0.006			0.006			0.005	
(Intercept)	5.39***			5.10***			5.08***			3.86***		
AI	0.09			-0.00			-0.26			0.26		
Human	0.06			-0.23			-0.17			0.10		
Explain	0.17			-0.07			0.01			0.07		
Step2		0.019	0.014		0.012	0.006		0.010	0.004		0.008	0.003
(Intercept)	5.22***			5.09***			4.96***			3.78***		
AI	0.22			0.15			-0.10			0.31		
Human	0.42			-0.34			0.04			0.29		
Explain	0.48†			-0.05			0.23			0.22		
AI:explain	-0.24			-0.28			-0.28			-0.09		
Human:explain	-0.70†			0.24			-0.41			-0.37		

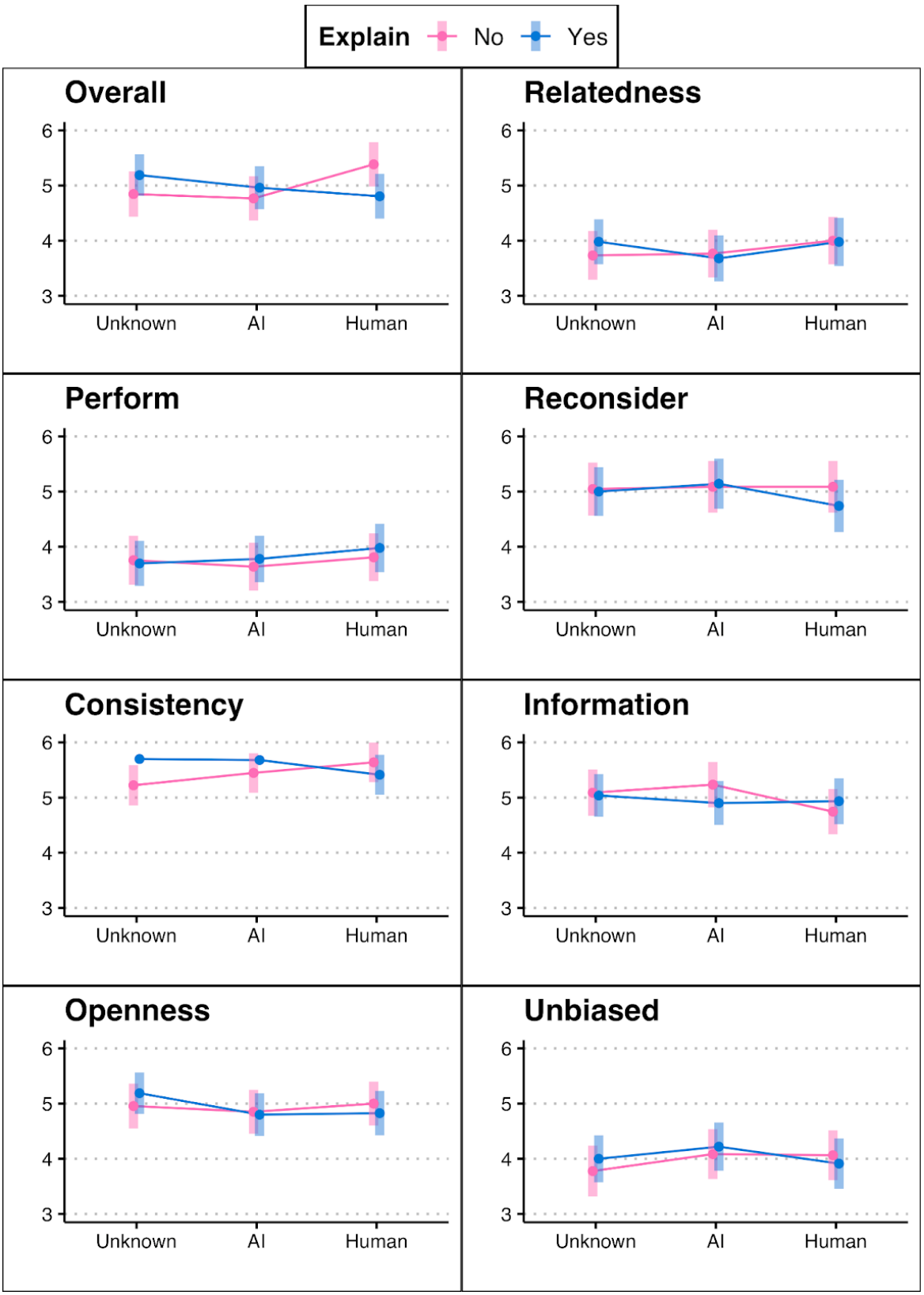
Note:

Group unknown is coded as the baseline. Explain was coded as 0=Explanation not provided and 1=Explanation provided.

df=(3, 284) in Step 1, and df=(5, 282) in Step 2.

† p<0.1, * p<0.05, ** p<0.01, *** p<0.001.

Table 4.5 Moderated hierarchical multiple regression analysis



Note:
Overall=Overall procedural fairness. Relatedness=Job relatedness. Perform=Opportunity to perform. Reconsider=Reconsideration opportunity. Consistency=Consistency of administration. Information=Selection information. Openness=Honesty. Unbiased=Bias suppression.

Figure 4.3 Moderating effects of explanations

4.4 Summary of hypothesis testing

In sum, the results of hypothesis testing are shown in the table below.

Hypotheses	Results
Hypothesis 1 *AVIs rated by humans are perceived procedurally fairer than those rated by AI.*	**Supported**
Hypothesis 2 *Transparency in decision makers positively impacts perceptions of procedural fairness in AVIs.*	**Not supported**
Hypothesis 3 *In AVIs, the provision of explanations moderates the relationships between decision makers and procedural fairness perceptions.*	**Partially supported**

Table 4.6 Summary of hypothesis testing

4.5 Subgroup analysis

In addition to evaluating the comparability among the experimental groups (i.e., Section 4.2), independent samples *t*-tests between several subgroups were performed to analyze in depth the potential variances caused by the characteristics of the participants. They were divided into subgroups such as male (N=105) and female (N=178), older (N=116) and younger (N=171) split by the mean age (24.5) of the participants, and bachelor's (N=90) and master's (N=184), regardless of the treatment groups they were assigned to. Figure 4.4 illustrates the overall distribution of perceptions among these subgroups.

As presented in Table 4.7, some statistically significant differences existed among the subgroups, most noticeably between older and younger participants. The results of independent samples *t*-tests indicated that, in terms of job relatedness, there was a significant difference between older (M=3.61, SD=1.44) and younger participants (M=4.04, SD=1.50), t(254)=-2.40, p=0.01, d=0.29. Also, a superior perception of opportunity to perform was observed for the younger (M=3.93, SD=1.49) than for the older subgroups (M=3.54, SD=1.50), t(245)=-2.15, p=0.03, d=0.26. Moreover, statistical significance was found in the items of honesty and bias suppression between these two subgroups. The younger participants had better perceptions in honesty (M=5.15, SD=1.30) in comparison to the older (M=4.63, SD=1.44), t(230)=-3.13, p<0.01, d=0.38. The item of bias suppression was viewed as

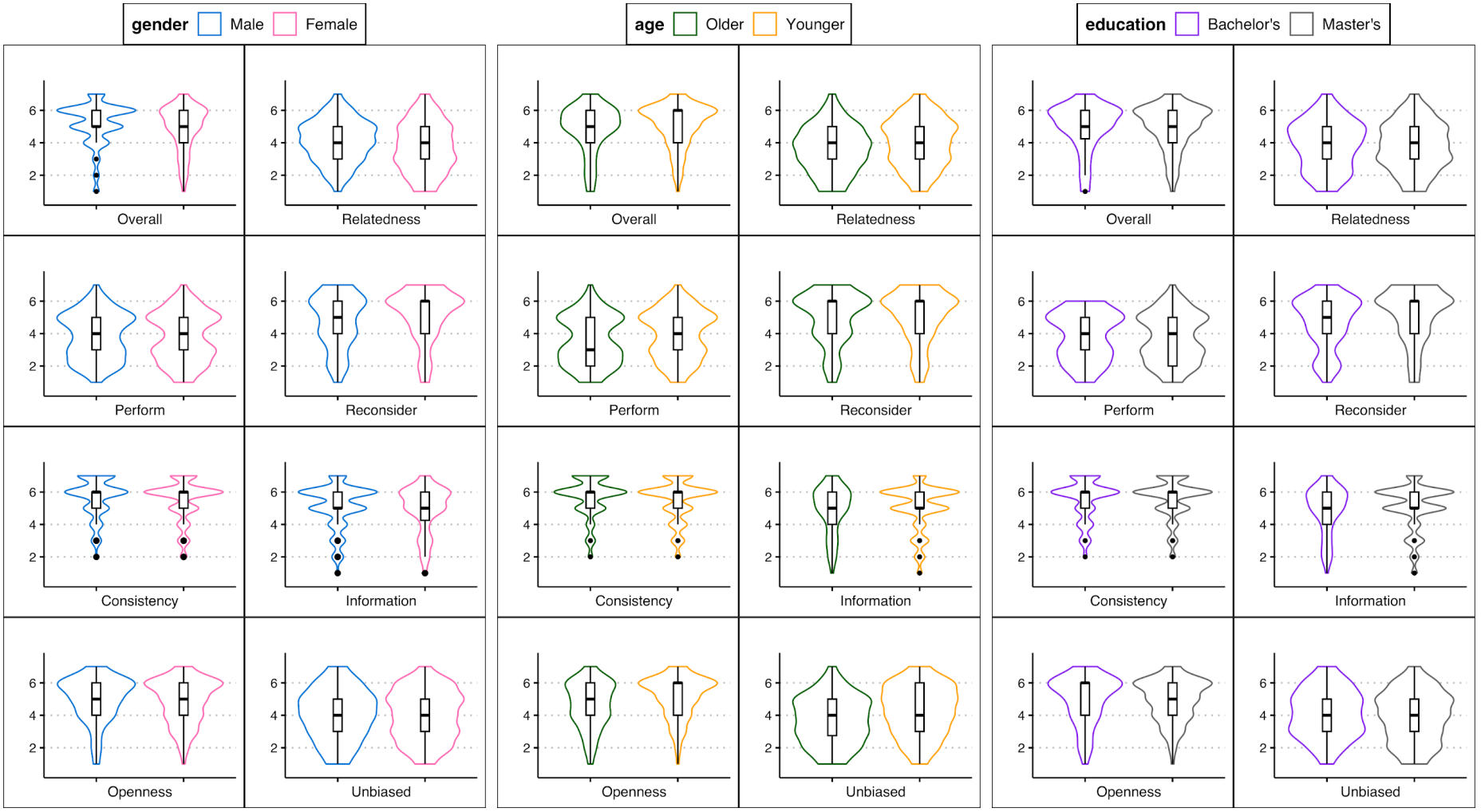
significantly superior by the younger (M=4.27, SD=1.56) than by the older (M=3.66, SD=1.49), $t(255)=-3.36$, $p<0.001$, $d=0.40$. In addition to age, there was a significant difference in the job relatedness item between males (M=4.11, SD=1.30) and females (M=3.73, SD=1.55), $t(249)=2.23$, $p=0.03$, $d=0.26$. Finally, the participants who have master's education demonstrated significantly better perceptions in reconsideration opportunities (M=5.16, SD=1.64) compared to those who have bachelor's education (M=4.69, SD=1.62), $t(179)=-2.27$, $p=0.02$, $d=0.29$.

These subgroup analyses provided in depth insights into how different characteristics of participants could be associated with statistical significance in a few items in the questionnaire. The results indicated that the characteristics may be potential explanatory variables for some of the fairness perception items despite the treatments. However, these demographic variables were not included in the prior analyses for hypothesis testing because of random assignment of treatment groups and the insignificant differences in potential covariate analysis (Table 4.2). Still, the findings of subgroup analyses were deemed valuable and helpful for the discussions in the next chapter.

Measures	**Male** (N=105) Mean (SD)	**Female** (N=178) Mean (SD)		**Older** (N=116) Mean (SD)	**Younger** (N=171) Mean (SD)		**Bachelor** (N=90) Mean (SD)	**Master** (N=184) Mean (SD)	
Overall	5.17 (1.32)	4.94 (1.39)	t(228)=1.41, p=0.16, d=0.17	4.85 (1.49)	5.10 (1.34)	t(229)=-1.43, p=0.15, d=0.18	4.98 (1.49)	5.00 (1.37)	t(164)=-0.12, p=0.91, d=0.02
Job relatedness	4.11 (1.30)	3.73 (1.55)	t(249)=2.23, p=0.03*, d=0.26	3.61 (1.44)	4.04 (1.50)	t(254)=-2.40, p=0.01*, d=0.29	3.89 (1.51)	3.78 (1.49)	t(175)=0.55, p=0.58, d=0.07
Opportunity to perform	3.85 (1.41)	3.76 (1.52)	t(231)=0.47, p=0.64, d=0.06	3.54 (1.50)	3.93 (1.49)	t(245)=-2.15, p=0.03*, d=0.26	3.84 (1.50)	3.71 (1.51)	t(178)=0.69, p=0.49, d=0.09
Reconsideration opportunity	4.76 (1.67)	5.16 (1.60)	t(210)=-1.95, p=0.05†, d=0.24	4.99 (1.63)	5.04 (1.63)	t(247)=-0.25, p=0.80, d=0.03	4.69 (1.62)	5.16 (1.64)	t(179)=-2.27, p=0.02*, d=0.29
Consistency of administration	5.67 (1.19)	5.44 (1.28)	t(231)=1.48, p=0.14, d=0.18	5.65 (1.21)	5.44 (1.27)	t(255)=1.36, p=0.17, d=0.16	5.52 (1.27)	5.54 (1.23)	t(171)=-0.13, p=0.90, d=0.02
Selection information	5.10 (1.34)	4.94 (1.45)	t(232)=0.98, p=0.33, d=0.12	4.93 (1.47)	5.03 (1.40)	t(238)=-0.57, p=0.57, d=0.07	4.91 (1.44)	5.04 (1.41)	t(173)=-0.72, p=0.47, d=0.09
Honesty	5.00 (1.39)	4.89 (1.38)	t(216)=0.66, p=0.51, d=0.08	4.63 (1.44)	5.15 (1.30)	t(230)=-3.13, p=0.00**, d=0.38	4.98 (1.50)	4.95 (1.30)	t(157)=0.17, p=0.86, d=0.02
Bias suppression	4.04 (1.59)	3.98 (1.55)	t(215)=0.31, p=0.76, d=0.04	3.66 (1.49)	4.27 (1.56)	t(255)=-3.36, p=0.00***, d=0.40	4.10 (1.50)	3.90 (1.58)	t(185)=1.01, p=0.31, d=0.13

Significance levels: † p<0.1, * p<0.05, ** p<0.01, *** p<0.001.

Table 4.7 Results of *t*-tests for the subgroups



Note:
Overall=Overall procedural fairness. Relatedness=Job relatedness. Perform=Opportunity to perform. Reconsider=Reconsideration opportunity. Consistency=Consistency of administration. Information=Selection information. Openness=Honesty. Unbiased=Bias suppression.

Figure 4.4 Survey result distribution of the subgroups

5. Discussion

In this chapter, we discuss the findings of our online scenario-based experimental study and the managerial implications derived from them. Furthermore, limitations of this research and suggestions for future research are discussed at the end.

5.1 Different decision makers in AVIs

One of the main purposes of this study is to investigate the extent to which artificial intelligence (AI) and algorithmic prediction models that are gradually employed in asynchronous video interviews (AVIs) influence job applicants' fairness perceptions. Hence, this section will start by discussing the first research question:

> *To what extent do AI and human decision makers affect job applicants' perceptions of fairness in AVIs?*

In support of hypothesis 1, the first findings indicate that AVIs rated by humans were perceived fairer overall than those rated by AI. Interestingly, no significant differences were found between these two decision makers when examined with SPJS subscales. Similar discrepancy is also found in some previous research. For example, Langer et al. (2017) find no differences in the subscales of applicants' perceptions but in a three-item index to measure the overall perceptions of procedural fairness. This overall index is also used in other research (e.g., Langer et al., 2021), and is further employed but with minor alteration by Salley (2022) who, likewise, discovers a paradoxical result that procedural fairness perceptions toward human and AI decision makers are significantly different when measured with an overall fairness index but not with any of the SPJS subscale items.

The superior overall fairness perceptions of humans to AI can be attributed to the following reasons. Firstly, the restriction of impression management (Blacksmith et al., 2016; Roulin et al., 2014) could explain the lower preference of AI than of humans as the decision maker. Although AVIs remove the immediate interactions between the two parties in an interview, the interview takers may still think that they have higher chances to successfully deliver nonverbal messages like nodding or smiling (Frauendorfer et al., 2014) when their interview videos are watched by humans. Secondly, emotional creepiness, an irrational feelings defined as "a queasy feeling paired with uncertainty about how to behave or how to judge a situation" (Langer et al., 2019), can occur when the interviewees have unfamiliar interactions with technologies (Langer et al., 2017; Tene and Polonetsky, 2013). In the manipulations of decision makers in this research, a human avatar was used to represent a human recruiting manager and a robot with neural circuits was used to represent an AI evaluation system (Appendix A.3 and A.4). The latter may elicit

participants' feeling of creepiness due to unfamiliarity with AI or the new technology involved, thus creating a negative affective reaction and decreasing their fairness perceptions. Thirdly, candidates might feel that they deserve individual evaluations and expect that the evaluator will spend time reviewing their videos rather than simply being sorted by an automated system within seconds, as if they were just an ordinary one among many in a candidate pool.

Although humans were perceived fairer than AI as the decision maker in the overall fairness item, there were no significant differences in the seven subscale items adopted from SPJS. The first potential explanation to this inconsistent result concerns the design of SPJS items. Unlike the overall item that inquires into a broader perception of whether the whole process is fair, the subscale items focus on more specific and detailed aspects of procedural justice rules. When perceptions are measured in a more granular manner, the participants' general affective response to a novel and unfamiliar technology may be more diluted (Salley, 2022). In addition, SPJS was originally designed to measure applicants' fairness perceptions toward traditional human-led recruitment. However, the hiring process evolution and the increasing use of AI technologies have led to shifts in the dimensions of procedural fairness perceptions. When the interview processes have been massively changed by new technologies, new factors that associate with fairness perceptions may not be fully captured by the items in SPJS, therefore leading to the finding of no differences. Additionally, the characteristics of the interview format could potentially explain the finding. When the interview is asynchronous, the interviewees have less pressure on answering the questions immediately. Longer preparation time allows the interview takers to prepare their answers and reasonably leads to a higher level of perceived behavioral control (Langer et al., 2017). Moreover, in our experiment, the instructions informed the participants that they could record their answers up to three times (Appendix A.2). Compared with other interview methods, particularly synchronous ones, that expect instant response and allow no more than one chance (e.g., face-to-face interviews or phone calls), AVIs could inferrably decrease the perceptive differences in the decision makers.

5.2 The role of transparency in AVIs

In this section, the discussions revolve around transparency, in terms of both the disclosure of decision makers and the provision of additional explanations, grounded on the second research question:

> *How does transparency affect job applicants' perceptions of fairness in AVIs?*

5.2.1 Transparency in decision maker

The findings from hypothesis 2 indicate that transparency in the identity of decision makers did not influence participants' procedural fairness perceptions. The result is in line with Salley's (2022) finding that there are no significant effects on procedural fairness whether or not the information about the decision maker is provided. One possible explanation is that, since personnel selection processes have become lengthy and complicated nowadays, job applicants normally do not expect excessive information disclosure in the first place. In other words, as AVIs are commonly employed as the first step in interview assessment, candidates typically do not expect a high degree of transparency at this early stage of the process. Furthermore, the results could also be ascribed to the limited visibility job applicants often have into how organizations make hiring decisions. Even though information asymmetry exists in recruitment processes, applicants may worry that requesting more information will negatively affect their applications and pose them in a more disadvantaged position than others who do not request. Consequently, they have little choice but passively accept an opaque recruitment process.

Although the disclosure of decision makers did not appear to influence fairness perception in this study, hiding the use of AI evokes ethical, legal, and applicant reaction concerns, and can potentially harm the organization. Firstly, from a business ethics perspective, job applicants have the right to know whether it is AI or human who handles their applications. There are two major perspectives concerning the responsibility for disclosure, as Hunkenschroer and Luetge (2022) summarize: Utilitarians may not prioritize transparent processes as long as the best candidates are hired, whereas the deontologists argue that violating individuals' rights cannot be justified by the greater good for the majority. To ensure an ethical use of AI in recruitment, one of the standards that organizations can implement is to proactively inform applicants with whom they are communicating and by whom their data will be processed and analyzed (Simbeck, 2019). Secondly, from a legal perspective, concealing information about the use of AI in recruitment processes may be considered illegal, as regulators are becoming more concerned about privacy protection and the legal right to disclosure has been progressively legislated. For example, the European Union General Data Protection Regulation (GDPR), which has been in force since May 2018, requires the organizations to disclose the use of AI if it is employed to process personal data (Seizov and Wulf, 2020). By regulating the collection and storage and by mandating consent before processing any personal data, GDPR safeguards the rights and privacy of European Union citizens (Hunkenschroer and Luetge, 2022). On another continent, the Artificial Intelligence Video Interview Act also facilitates greater informed consent for the use of AI-based tools in the United States (Hilliard et al., 2022). Thirdly, from an applicant reaction perspective, disguising the use of AI can result in distrust and suspicion if job applicants discover from other sources of information. Such detrimental perceptions consequently lead to a decrease in the attractiveness of the company and result in

unfavorable applicant reactions, such as withdrawing the application or declining the job offer (Köchling et al., 2022). Furthermore, the negative applicant reactions are likely to persist even after the recruitment process ends (Köchling et al., 2022).

5.2.2 Explanations

The findings from hypothesis 3 suggest that the moderating effects of explanations only partially existed when the decision makers were humans. Despite that prior research (e.g., Basch and Melchers, 2019) has examined how explanations directly affect job applicants' perceptions and their actions afterward, to the best of our knowledge, this is the first research investigating the provision of explanation as a moderator between various types of decision makers and job applicants' fairness perceptions in digital interviews.

In this study, the provision of additional explanations prior to the interview had negative moderating effects on the relationships between human decision makers and participants' procedural fairness perceptions. Also, the perception of consistency toward human-rated AVIs decreased when extra explanations were provided. One of the underlying reasons may be that these explanations remind applicants of human biases associated with inconsistency in personnel selection. Indeed, the way humans process information is less systematic than that of algorithms and can lead to contradictory or insufficiently evidence-based results (Woods et al., 2020). Even though the selection criteria have been predetermined, interpretation discrepancies among individual human raters inevitably exist, indicating that the same recorded interview video can get different results when evaluated by different human raters. Hence, specifying the selection criteria and providing more detailed explanations may serve as a reminder of potential inconsistencies and thus trigger feelings of unfairness among applicants.

Intriguingly, the provision of more explanations did not moderate the relationships between AI decision makers and participants' procedural fairness perceptions. The result could potentially be explained by two reasons, one of which is that the participants in this study are relatively young (M=24.5) and may have a higher level of technology acceptance. Most of the participants can be seen as digital natives, a label used to describe people who have been engrossed in a networked world and have been surrounded by information technologies for their entire lives (Kesharwani, 2020), and thus are more likely to be familiar and comfortable with new technologies and new selection tools. This explanation is supported by the subgroup analysis in Section 4.5, which showed that younger participants had superior perceptions in the items of job relatedness, opportunity to perform, honesty, and bias suppression. Participants' familiarities, comforts, and confidences in taking technology-mediated interviews may consequently lead them to pay less attention or even ignore the additional information provided in the guidance. The other underlying reason could

be that participants' preconceived beliefs about the decision agent in AVIs and about the characteristics of AI merely elicit subtle reactions that are difficult to detect in the research. Due to the lack of any interaction and the seamless integration of technology in AVIs, participants may assume that the whole process, from asking questions, evaluating responses, to ultimately making decisions, is fully automated. Essentially, AI is likely to be viewed as an invisible force that operates behind the digital interviews. Furthermore, the explanations provided in the survey materials may, to a large extent, align with the participants' acknowledgements of how AI works. As a result, the participants are neither surprised nor motivated to engage in sense-making, and thus do not have any notable reactions to be discerned.

To sum up, although transparency and explanations were expected to increase procedural fairness perceptions in AVIs, they could lead to unintended consequences. Specifically, neither transparently revealing the identity of decision makers nor proactively providing more detailed explanations in AI-rated AVIs increases applicants' procedural fairness perceptions. Nonetheless, adverse effects may occur when human recruiters are involved in the decision-making processes.

5.3 Practical implications

By manipulating different decision-making agents and simulating an AVI interface for candidate assessments, we observe that applicants still prefer humans to AI in terms of fairness perceptions. Also, providing additional explanations may negatively moderate the perceptions when the decision makers are humans. Based on the findings, there are five practical implications that human resources practitioners can take away from this study.

First and foremost, organizations should carefully consider what information to disclose and to what extent when employing AVIs and AI decision agents. Research has shown that, when employing novel technologies in personnel selection, applicants' adverse reactions can be mitigated cost-efficiently by providing information (Lahuis et al., 2003; McCarthy et al., 2017b; Truxillo et al., 2009). Nevertheless, this study finds that disclosing the decision makers and providing more explanations in AI-rated AVIs do not significantly affect participants' fairness perceptions. Moreover, extra explanations can have a negative impact when the interviewees are informed that a human will watch their videos and evaluate their performance. Langer et al. (2018) also discover that excessive process justification could lead to unfavorable perceptions when novel technologies are involved. Altogether, providing information can potentially bring both benefits and risks. Therefore, increasing transparency with the aim to reduce adverse reactions should be approached with caution, as it may have the opposite effect and exacerbate negative perceptions.

In addition to increasing transparency, organizations could also increase the level of humanization in the interface to reduce the perceptive gap between humans and AI, thus potentially mitigating the negative effects of AI in recruitment. Automated interview interfaces can elicit negative emotional responses such as the feeling of creepiness, which are likely to reduce applicants' affective trust (Langer et al., 2017; Langer et al., 2019). In an attempt to counterbalance the negative emotions, one practical method is to increase tangibility, which refers to the ability of AI to be physically perceived by humans (Liu and London, 2016), and embody AI with a more amiable image. Yet, merely a few research investigates the effects of increasing tangibility of AI, one of which is Suen and Hung (2023) who find that creating a human-like interface by using an attractive virtual agent or avatar can increase applicants' affective trust in automated digital interviews. The findings further indicate that incorporating humanized features in the web-based interview interface can enhance applicants' trust by simulating a social presence and human interaction. In line with this claim, some researchers (e.g., Mirowska and Mesnet, 2022) suggest that interview takers conceivably feel more comfortable with an anthropomorphized interface and tend to respond to AI as they would do to human interviewers.

Thirdly, offering informative and creative user tutorials can be an effective approach to engage applicants and alleviate negative reactions toward the use of AI. Although it is impossible to ensure that all candidates are knowledgeable of advanced technologies and are familiar with AI-based interview processes, organizations can at least increase the perceived ease of use toward the systems by offering options for practices and guidance on how to prepare for and record responses. Furthermore, offering interesting but informative user tutorials is also an action that firms can take to engage the users and enhance both their knowledge and their perceptions toward the use of new tools. For the tutorials to be "interesting", the attention can be paid to the design, style, and presentation. A remarkable example is the pre-flight safety demonstrations. Recognizing the limited effectiveness of presenting information in dull, plain, and straightforward formats, airline companies have begun to replace their strait-laced safety videos with more engaging contents by adding more pop-culture elements, humor, and appealing visual designs (Schneider, 2017). Inferably, user tutorials that are specially designed and creatively presented not only can boost their effectiveness and interviewee engagement but also can signal an organization's dedication to improving the application experience. Moreover, the tutorials can be used as a means to attract similar-minded candidates, subsequently increasing organizational attractiveness and shaping a more vibrant organizational climate (Gilliland, 1993).

Fourthly, utilizing a hybrid decision-making structure and incorporating both humans and algorithms in the recruitment processes can gradually increase the acceptance of AI in the loop. While exploiting the benefits like cost-efficiency and time-efficiency of AI, a hybrid system can simultaneously satisfy applicants' preference for humans in making decisions and serve as a progressive step toward

the full human-to-AI delegation. There are two possible approaches to implement the hybrid systems. One approach is to divide the decision-making process based on the strengths of each decision maker. For example, Lavanchy et al. (2023) suggest that organizations can use algorithms to examine quantitative elements and keep employing humans to evaluate qualitative factors. The other approach is to adopt a sequential decision-making structure. In situations with extensive alternative sets, such as hiring processes, an AI-to-human sequential decision-making structure is deemed beneficial (Shrestha et al., 2019). In this hybrid system, AI is employed in the initial stage of the decision-making process as a filter to eliminate most of the unsuitable or unqualified candidates, and humans in the subsequent stage select the candidates from the remaining pool. This human-involved sequential decision making structure also increases the interpretability of the ultimate hiring decisions (Shrestha et al., 2019).

Finally, organizations should pay attention to whether negative procedural fairness perceptions toward AI as a decision-making agent decrease over time. It is highly likely that job applicants will be more familiar with new selection tools as the use of AI systems in HRM becomes more prevalent. This familiarity may lead to a greater acceptance of AI in the future (Köchling et al., 2022). This trend is consistent with the findings that younger participants have superior perceptions on part of the subscales. However, given that employing humans to watch recorded interview videos entails higher costs of labor and time (Torres and Gregory, 2018), organizations are still likely to delegate preliminary hiring decisions to AI and algorithmic systems despite the current perceptions that human-made decisions are fairer. Nonetheless, organizations should be mindful of three things. First, whether the advantages of fully delegating decision-making authority to AI outweigh the risks has not been concluded, especially in the context of emerging assessment tools. Second, as employers normally have greater power than the job applicants, human resources practitioners should deliberate upon the ethical issues of exploiting the power asymmetry in the job market. Thirdly, even if the decisions are fully delegated to AI models, some employees should still stay in the loop as more and more laws and regulations protect job applicants' "right to explanations" or, more strictly, right to require "human in the loop" (Sánchez-Monedero et al., 2020).

5.4 Limitations and suggestions for future research

5.4.1 Scale

Despite the dominant use of SPJS in measuring procedural fairness, it was developed in an era when interviews were mostly conducted in the format of either face-to-face or telephone. However, the process for job applications and the format of interviews have significantly changed since 2001, the year when SPJS was published. These changes may decrease the applicability of the items in the scale. For example, when

applicants are asked to record themselves answering questions and upload the videos to an evaluation system, they do not have the chance to interact with anyone but their electronic devices, eliminating the chance to lead the conversation flow. To reflect on the current status of technology-mediated interviews, the authors chose to either exclude or revise items that are not applicable in the AVI context and added an overall fairness item. Even so, the questionnaire was not customized enough to catch up with the technological advances. Moreover, recruitment processes have been transformed when algorithms are in play, creating new features, such as tangibility, that can be associated with fairness perceptions. Hereby, the authors call for the need to update a scale tailored for measuring fairness perceptions in the modern context.

Another scale-related issue that may constrain the results in this research is that the authors only picked one item, the most applicable one, from each procedural fairness rule with the intention to limit the time needed for the participants to complete the survey and thus reduce dropout rates. Yet, each item in the scale was developed to be used altogether, thus making examinations of internal scale reliability (i.e., Cronbach alpha) possible. Merely picking one item from each subscale hinders the internal reliability examination within each subscale. Uncertain representativeness of the item to the concept might be one of the reasons causing the relatively insignificant results in this research.

5.4.2 Sample

To the best of the authors' knowledge, an adequate sample size is roughly 50 in each treatment group in similar experimental studies (e.g., Basch et al., 2021; Köchling et al., 2022; Langer et al., 2017; Langer et al., 2019). Therefore, the goal of total number of participants had been set at 300. However, as discussed in Section 3.3.1, one of the limitations was that the number of samples was only enough for confidently detecting a few of the effect sizes with 80% power at a 0.05 significance level. The other recognized limitation concerns sampling; in particular, convenience sampling. A convenience sample refers to one that is directly accessible to the researcher (Bell et al., 2019, p.197). The participants in this study were recruited either at Stockholm School of Economics or from the authors' social networks. The sampling method may cause lower variance in the participants and thus lead to lower representation and generalizability of the results to the mass population. Furthermore, effects of explanations are found weaker in studies that consist of student samples (Truxillo et al., 2009). Still, our sample is arguably a representative one because job applicants around 24 or 25 years old are the ones who are most likely to experience such an interview process. The longer people work, the more likely they will utilize other channels, such as via professional networks or referrals, when applying for a job.

5.4.3 Research method

The third recognized limitation concerns the research method. In particular, field experiments may be a more suitable approach to studying job applicants' reactions. Although the findings could be informative in a laboratory experimental design (Lukacik et al., 2022) and true-to-life vignettes could be useful for studying the perceptions, feelings, and attitudes of real-life situations (Taylor, 2006), the results generated from a field research are likely to have higher generalizability, especially when the manipulations are meaningful for the participants (Acikgoz et al., 2020). Blacksmith et al.'s (2016) meta-analysis of technology-mediated interviews report a much larger negative effect in field research than in laboratory studies. Similarly, the effects of providing job applicants with explanations are also found stronger in field settings than in simulations (Truxillo et al., 2009); that is, simulated interviews may still lack the emotional and cognitive fidelity of real ones (Posthuma et al., 2002), leading to weaker perceptions among the participants and thus lowering the chances to detect them.

While acknowledging the lack of fidelity as a limitation, the authors found it difficult to perform a field experiment for several reasons. Firstly, with limited resources in this thesis in terms of both time and funding, a field study was nearly impossible. It would be more ideal for future researchers to work with a middle-to-large-scale firm that receives enough applications for a real job opening. Moreover, ethical issues should also be taken into consideration, especially when the outcome truly affects job applicants' lives. For instance, is it fair that some of them are provided with more explanations while others are not? Is it ethical to intervene in a real hiring process? Is it moral to analyze private information of the applicants beyond job application purposes? More dilemmas can be envisioned when the research is designed as a field experiment. Hence, ethical issues seem inevitable and the authors suggest that future researchers should thoroughly contemplate them before conducting field studies.

5.4.4 Materials

Upon reviewing previous research, the authors were aware that the attitudes toward processes involving technology-based decisions could differ based on the context. For example, Lee (2018) compares perceptions of human versus algorithmic managerial decisions based on tasks requiring either mechanical or human skills. Yet, one crucial variable in this study is transparency, including both (non) disclosure of the decision maker and the provision of extra explanations, rather than different types of tasks. Consequently, when designing the materials for the survey, the authors aimed to avoid respondents' prejudices about a specific role and thus decided not to specify it. Likewise, the explanations regarding the selection criteria provided to treatment groups 2, 4, and 6 were intentionally described in a general manner to minimize the chance of inferring specific roles. However, such general information could be

presumed to catch less attention and cause proportionally general results, leading to smaller effect sizes and fewer statistically significant findings.

6. Conclusion

New technologies have been increasingly used in recruitment processes. Not only has asynchronous video interview (AVI) been adopted as a means to screen and evaluate candidates but AI and algorithms have been integrated and take on tasks that used to be done by humans. This research investigated how different decision makers, greater transparency, and their interactions affect job applicants' perceptions of fairness in an AVI setting. The results suggest that people have superior fairness perceptions of human recruiters than of AI in less transparent conditions. However, when more explanations are provided, applicants may be reminded of the inherent existence of human bias and the level of perceived fairness decreases. Interestingly, the moderating effect does not exist when the interview videos are evaluated and when the decisions are made by AI. Finally, although the study reaffirms that new technologies like AI have not been seen as perfect substitutes for humans in personnel selection, there are still actions human resources practitioners can take to mitigate the negative perceptions and lead the applicants to favorable post-interview actions in the future.

7. References

Abdul, A. et al. (2018) 'Trends and trajectories for explainable, accountable and intelligible systems: an HCI research agenda', *Proceedings of the 2018 CHI conference on human factors in computing systems,* Montréal, Canada, 21–26 April.

Acikgoz, Y. et al. (2020) Justice perceptions of artificial intelligence in selection. *International journal of selection and assessment*. 28 (4), 399–416.

Ananny, M. and Crawford, K. (2018) Seeing without knowing: limitations of the transparency ideal and its application to algorithmic accountability. *New media & society*. 20 (3), 973–989.

Araujo, T. et al. (2020) In AI we trust? Perceptions about automated decision-making by artificial intelligence. *AI & society*. 35 (3), 611–623.

Arrieta, A. B. et al. (2020) Explainable artificial intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. *Information fusion*. 58, 82–115.

Basch, J. and Melchers, K. (2019) Fair and flexible?! Explanations can improve applicant reactions toward asynchronous video interviews. *Personnel assessment and decisions*. 5 (3), 2.

Basch, J. M. et al. (2021) It takes more than a good camera: which factors contribute to differences between face-to-face interviews and videoconference interviews regarding performance ratings and interviewee perceptions? *Journal of business and psychology*. 36 (5), 921–940.

Bauer, T. N. et al. (2001) Applicant reactions to selection: development of the selection procedural justice scale (SPJS). *Personnel psychology*. 54 (2), 387–419.

Beattie, G. and Johnson, P. (2012) Possible unconscious bias in recruitment and promotion and the need to promote equality. *Perspectives : policy and practice in higher education*. 16 (1), 7–13.

Bell, B. S. et al. (2006) Consequences of organizational justice expectations in a selection system. *Journal of applied psychology*. 91 (2), 455–466.

Bell, E. et al. (2019) *Business research methods*. Fifth edition. Oxford: Oxford University Press.

Berente, N. et al. (2019) Managing AI. *MIS quarterly*, 1–5.

Berente, N. et al. (2021) Managing artificial intelligence. *MIS quarterly*. 45 (3), 1433–1450.

Bies, R. J. (2005) 'Are procedural justice and interactional justice conceptually distinct', in Greenberg, J. and Colquitt, J. A. (ed.) *Handbook of organizational justice*. Lawrence Erlbaum Associates Publishers. 85–112.

Black, J. S. and van Esch, P. (2020) AI-enabled recruiting: what is it and how should a manager use it? *Business horizons*. 63 (2), 215–226.

Blacksmith, N. et al. (2016) Technology in the employment interview: a meta-analysis and future research agenda. *Personnel assessment and decisions*. 2 (1), 2.

Bonaccio, S. et al. (2016) Nonverbal behavior and communication in the workplace: a review and an agenda for research. *Journal of management*. 42 (5), 1044–1074.

Brenner, F. S. et al. (2016) Asynchronous video interviewing as a new technology in personnel selection: the applicant's point of view. *Frontiers in psychology*. 7, 863.

Butucescu, A. et al. (2019) The generalizability of reactions to assessment: an application of the selection procedural justice scale (SPJS) in academic settings. *International journal of selection and assessment*. 27 (4), 406–410.

Byrne, Z. S. and Cropanzano, R. (2001) The history of organizational justice: the founders speak. *Justice in the workplace: From theory to practice*. 2(1), 3–26.

Cable, D. M. and Turban, D. B. (2003) The value of organizational reputation in the recruitment context: a brand-equity perspective. *Journal of applied social psychology*. 33 (11), 2244–2266.

Caliskan, A. et al. (2017) Semantics derived automatically from language corpora contain human-like biases. *Science*. 356 (6334), 183–186.

Celani, A. and Singh, P. (2011) Signaling theory and applicant attraction outcomes. *Personnel review*. 40 (2), 222–238.

Chamorro-Premuzic, T. et al. (2016) New talent signals: shiny new objects or a brave new world? *Industrial and organizational psychology*. 9 (3), 621–640.

Charness, G. et al. (2012) Experimental methods: between-subject and within-subject design. *Journal of economic behavior & organization*. 81 (1), 1–8.

Cheng, M. M. and Hackett, R. D. (2021) A critical review of algorithms in HRM: definition, theory, and practice. *Human resource management review*. 31 (1), 100698.

Cohen-Charash, Y. and 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. et al. (2001) Justice at the millennium: a meta-analytic review of 25 years of organizational justice research. *Journal of applied psychology*. 86 (3), 425–445.

Corbett-Davies, S. and Goel, S. (2018) The measure and mismeasure of fairness: a critical review of fair machine learning. *arXiv preprint arXiv*. 1808.00023.

Cropanzano, R. et al. (2007) The management of organizational justice. *Academy of management perspectives*. 21 (4), 34–48.

Cropanzano, R. and Greenberg, J. (1997) 'Progress in organizational justice: tunneling through the maze', In Cooper, C. and Robertson, I. (ed.) *International review of industrial and organizational psychology*. New York: Wiley, 317-372.

Cugueró, N. and Rosanas, J. M. (2011) Fairness, justice, subjectivity, objectivity and goal congruence in management control systems. *IESE Business School working paper No. 891*.

Danieli, O. et al. (2016) How to hire with algorithms. *Harvard business review*. 17.

Dattner, B. et al. (2019) The legal and ethical implications of using AI in hiring. *Harvard Business Review*. 25.

De Fine Licht, J. et al. (2014) When does transparency generate legitimacy? Experimenting on a context-bound relationship. *Governance (Oxford)*. 27 (1), 111–134.

de Greeff, J. et al. (2021) 'The FATE system: FAir, transparent and explainable decision making', *Proceedings of the AAAI 2021 spring symposium on combining machine learning and knowledge engineering*, Palo Alto, USA, 22–24 March.

Demuijnck, G. (2009) Non-discrimination in human resources management as a moral obligation. *Journal of business ethics*. 88 (1), 83–101.

Dietvorst, B. J. et al (2014) Algorithm aversion: people erroneously avoid algorithms after seeing them err. *Journal of experimental psychology: general*. 143 (6), 1–13.

Duan, Y. et al. (2019) Artificial intelligence for decision making in the era of big data – evolution, challenges and research agenda. *International journal of information management*. 48, 63–71.

Folger, R. and Konovsky, M. A. (1989) Effects of procedural and distributive justice on reactions to pay raise decisions. *Academy of management journal*. 32 (1), 115–130.

Frauendorfer, D. et al. (2014) Nonverbal social sensing in action: unobtrusive recording and extracting of nonverbal behavior in social interactions illustrated with a research example. *Journal of nonverbal behavior*. 38 (2), 231–245.

Georgiou, K. and Nikolaou, I. (2020) Are applicants in favor of traditional or gamified assessment methods? Exploring applicant reactions towards a gamified selection method. *Computers in human behavior*. 109, 106356.

Gerber, A. S. and Green, D. P. (2012) *Field experiments : design, analysis, and interpretation*. New York: W. W. Norton.

Gilliland, S. W. (1993) The perceived fairness of selection systems: an organizational justice perspective. *The academy of management review*. 18 (4), 694–734.

Gilliland, S. W. et al. (2001) Improving applicants' reactions to rejection letters: an application of fairness theory. *Personnel psychology*. 54 (3), 669–703.

Gilliland, S. W. and Hale, J. M. (2005) 'How can justice be used to improve employee selection practices', in Greenberg, J. and Colquitt, J. A. (ed.) *Handbook of organizational justice*. Lawrence Erlbaum Associates Publishers, 411–438.

Glikson, E. and Woolley, A. W. (2020) Human trust in artificial intelligence: review of empirical research. *The academy of management annals*. 14 (2), 627–660.

Grgić-Hlača, N. et al. (2018) 'Beyond distributive fairness in algorithmic decision making: feature selection for procedurally fair learning', *Proceedings of the 32nd AAAI conference on artificial intelligence*, New Orleans, USA, 2–7 February.

Greenberg, J. (1987) A taxonomy of organizational justice theories. *The academy of management review*. 12 (1), 9–22.

Greenberg, J. (1990) Organizational justice: yesterday, today, and tomorrow. *Journal of management*. 16 (2), 399–432.

Greenberg, J. (2001) Setting the justice agenda: seven unanswered questions about 'what, why, and how'. *Journal of vocational behavior*. 58 (2), 210–219.

Greenberg, J. and Cropanzano, R. (1993) 'The social side of fairness: interpersonal and informational classes of organizational justice', in Cropanzano, R. (ed.) *Justice in the workplace: approaching fairness in human resource management*. Lawrence Erlbaum Associates, 79–103.

Guchait, P. et al. (2014) Video interviewing: a potential selection tool for hospitality managers – a study to understand applicant perspective. *International journal of hospitality management*. 36, 90–100.

Guest, J. W. (2017) Justice as lawfulness and equity as a virtue in Aristotle's Nicomachean Ethics. *The review of politics*. 79 (1), 1–22.

Gorman, C. A. et al. (2018) An investigation into the validity of asynchronous web-based video employment-interview ratings. C*onsulting psychology journal*. 70 (2), 129–146.

Hilliard, A. et al. (2022) Robots are judging me: perceived fairness of algorithmic recruitment tools. *Frontiers in psychology*. 13, 940456.

Hinkin, T. R. (1998) A brief tutorial on the development of measures for use in survey questionnaires. *Organizational research methods*. 1 (1), 104-121.

Hoff, K. A. and Bashir, M. (2015) Trust in automation: integrating empirical evidence on factors that influence trust. *Human factors*. 57 (3), 407–434.

Hooker, B. (2005) Fairness. *Ethical theory and moral practice*. 8 (4), 329–352.

Hunkenschroer, A. L. and Kriebitz, A. (2022) Is AI recruiting (un) ethical? A human rights perspective on the use of AI for hiring. *AI and Ethics*. 3 (1), 1–15.

Hunkenschroer, A. L. and Luetge, C. (2022) Ethics of AI-Enabled recruiting and selection: a review and research agenda. *Journal of business ethics*. 178 (4), 977–1007.

Imbens, G. W. and Rubin, D. B. (2015) *Causal inference in statistics, social, and biomedical sciences*. Cambridge University Press.

Jaser, Z. et al. (2022) Where automated job interviews fall short. *Harvard Business Review*.

Jones, D. A. and Martens, M. L. (2009) The mediating role of overall fairness and the

moderating role of trust certainty in justice-criteria relationships: the formation and use of fairness heuristics in the workplace. *Journal of organizational behavior*. 30 (8), 1025–1051.

Kaibel, C. et al. (2019) 'Applicant perceptions of hiring algorithms - uniqueness and discrimination experiences as moderators', *Proceedings of 2019 Academy of Management. Boston, USA, 9–13 August,* 1181–1186.

Kaplan, A. and Haenlein, M. (2019) Siri, Siri, in my hand: who's the fairest in the land? On the interpretations, illustrations, and implications of artificial intelligence. *Business horizons*. 62 (1), 15–25.

Kazim, E. et al. (2021) Systematizing audit in algorithmic recruitment. *Journal of intelligence*. 9 (3), 46.

Kesharwani, A. (2020) Do (how) digital natives adopt a new technology differently than digital immigrants? A longitudinal study. *Information & management*. 57 (2), 103170.

Kim, P. T. (2016) Data-driven discrimination at work. *William & Mary law review*. 58, 857–936.

Kizilcec, R. F. (2016) 'How much information? Effects of transparency on trust in an algorithmic interface', *Proceedings of the 2016 CHI conference on human factors in computing systems*, , San Jose, USA, 7–12 May, 2390–2395.

Kleinberg, J. and Mullainathan, S. (2019) 'Simplicity creates inequity: implications for fairness, stereotypes and interpretability', *Proceedings of the 2019 ACM conference on economics and computation,* Phoenix, USA, 24–28 June, 807–808.

Köchling, A. and 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 (Göttingen)*. 13 (3), 795–848.

Köchling, A. et al. (2022) Can I show my skills? Affective responses to artificial intelligence in the recruitment process. *Review of managerial science*. 1–30.

Lahuis, D. M. et al. (2003) The effect of legitimizing explanations on applicants' perceptions of selection assessment fairness. *Journal of applied social psychology*. 33 (10), 2198–2215.

Langer, M. et al. (2017) Examining digital interviews for personnel selection: Applicant reactions and interviewer ratings. *International journal of selection and assessment*. 25 (4), 371–382.

Langer, M. et al. (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.

Langer, M. et al. (2019) Highly automated job interviews: acceptance under the influence of stakes. *International journal of selection and assessment*. 27 (3), 217–234.

Langer, M. et al. (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.

Lavanchy, M. et al. (2023) Applicants' fairness perceptions of algorithm-driven hiring procedures. *Journal of business ethics*. 1–26.

Leary, M. R. and Kowalski, R. M. (1990) Impression management: a literature review and two-component model. *Psychological bulletin*. 107 (1), 34–47.

Lee, C. and Cha, K. (2023) FAT-CAT—explainability and augmentation for an AI system: a case study on AI recruitment-system adoption. *International journal of human-computer studies*. 171, 102976.

Lee, I. and Shin, Y. J. (2020) Machine learning for enterprises: applications, algorithm selection, and challenges. *Business horizons*. 63 (2), 157–170.

Lee, M. K. (2018) Understanding perception of algorithmic decisions: fairness, trust, and emotion in response to algorithmic management. *Big data & society*. 5 (1), 1–16.

Lee, M. K. and Baykal, S. (2017) 'Algorithmic mediation in group decisions: fairness perceptions of algorithmically mediated vs. discussion-based social division', *Proceedings of the 2017 acm conference on computer supported cooperative work and social computing,* Portland, USA, 25 February 25–01 March, 1035–1048.

Leicht-Deobald, U. et al. (2019) The challenges of algorithm-based HR decision-making for personal integrity. *Journal of business ethics*. 160 (2), 377–392.

Levashina, J. et al. (2014) The structured employment interview: narrative and quantitative review of the research literature. *Personnel psychology*. 67 (1), 241–293.

Leventhal, G. S. (1980) 'What should be done with equity theory', in Gergen, K. J.,

Greenberg, M. S. and Willis, R. H. (ed.) *Social exchange: advances in theory and research*. Springer, 27–55.

Li, L. et al. (2021) 'Algorithmic hiring in practice: recruiter and HR professional's perspectives on AI use in hiring', *Proceedings of the 2021 AAAI/ACM conference on AI, ethics, and society,* Virtual Event, USA, 19–21 May, 166–176.

Lind, E. A. (2001) Fairness heuristic theory: Justice judgments as pivotal cognitions in organizational relations. *Advances in organizational justice*. 56 (8), 56–88.

Lind, E. A. et al. (2001) Primacy effects in justice judgments: testing predictions from fairness heuristic theory. *Organizational behavior and human decision processes*. 85 (2), 189–210.

Lipton, Z. (2018) The mythos of model interpretability. *Communications of the ACM*. 61 (10), 36–43.

Liu, X. and London, K. (2016) 'T.A.I: a tangible AI interface to enhance human-artificial intelligence (AI) communication beyond the screen', *Proceedings of the 2016 ACM conference on designing interactive systems,* Brisbane, Australia, 4–8 June, 281–285.

Longoni, C. et al. (2019) Resistance to medical artificial intelligence. *The Journal of consumer research*. 46 (4), 629–650.

Lukacik, E.-R. et al. (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.

Magnusson K. (no date) Understanding statistical power and significance testing. Available at: https://rpsychologist.com/d3/nhst (Accessed: 06 April 2023).

Marcinkowski, F. et al. (2020) 'Implications of AI (un-) fairness in higher education admissions: the effects of perceived AI (un-) fairness on exit, voice and organizational reputation', *Proceedings of the 2020 conference on fairness, accountability, and transparency*. Barcelona, Spain, 27–30 January, 122-130.

McAllister, D. J. (1995) Affect-and cognition-based trust as foundations for interpersonal cooperation in organizations. *Academy of management journal*. 38 (1), 24–59.

McCarthy, J. M. et al. (2017a) Applicant perspectives during selection: a review addressing 'so what?,' 'what's new?,' and 'where to next?' *Journal of management*. 43 (6), 1693–1725.

McCarthy, J. M. et al. (2017b) Using pre-test explanations to improve test-taker reactions: Testing a set of 'wise' interventions. *Organizational behavior and human decision processes*. 141, 43–56.

Mejia, C. and Torres, E. N. (2018) Implementation and normalization process of asynchronous video interviewing practices in the hospitality industry. *International journal of contemporary hospitality management*. 30 (2), 685–701.

Mirowska, A. and Mesnet, L. (2022) Preferring the devil you know: potential applicant reactions to artificial intelligence evaluation of interviews. *Human resource management journal*. 32 (2), 364–383.

Morse, L. et al. (2022) Do the ends justify the means? Variation in the distributive and procedural fairness of machine learning algorithms. *Journal of business ethics*. 181 (4), 1083–1095.

Mujtaba, D. F. and Mahapatra, N. R. (2019) 'Ethical considerations in ai-based recruitment', In 2019 *IEEE international symposium on technology and society (ISTAS),* Medford, USA, 15-16 November, 1–7.

Newman, D. T. et al. (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.

Ötting, S. K. and 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.

Pawlenka, C. (2005) The idea of fairness: a general ethical concept or one particular to sports ethics? *Journal of the philosophy of sport*. 32 (1), 49–64.

Petrinovich, L. et al. (1993) An empirical study of moral intuitions: toward an evolutionary ethics. *Journal of personality and social psychology*. 64 (3), 467–478.

Ployhart, R. E. et al. (2017) Solving the supreme problem: 100 years of selection and recruitment at the journal of applied psychology. *Journal of applied psychology*. 102 (3), 291–304.

Posthuma, R. A. et al. (2002) Beyond employment interview validity: a comprehensive narrative review of recent research and trends over time. *Personnel psychology*. 55 (1), 1–81.

Rai, A. et al. (2019) Next-generation digital platforms: towards human-AI hybrids.

MIS quarterly. 43, 3–8.

Raghavan, M. et al. (2020) 'Mitigating bias in algorithmic hiring: evaluating claims and practices', *Proceedings of the 2020 conference on fairness, accountability, and transparency,* Barcelona, Spain, 27–30 January, 469-481.

Rasipuram, S. and Jayagopi, D. B. (2018) Automatic assessment of communication skill in interview-based interactions. *Multimedia tools and applications*. 77 (14), 18709–18739.

Ribeiro, M. T. et al. (2016) '"Why should I trust you?" Explaining the predictions of any classifier', *Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining,* San Francisco, USA, 13–17 August, 1135–1144.

Rosenthal, R. (1976) *Experimenter effects in behavioral research*. Irvington

Roulin, N. et al. (2014) Interviewers' perceptions of impression management in employment interviews. *Journal of managerial psychology*. 29 (2), 141–163.

Rupp, D. E. et al. (2017) A critical analysis of the conceptualization and measurement of organizational justice: Is it time for reassessment? *Academy of management annals*, 11(2), 919–959.

Rynes, S. L. et al. (1991) The importance of recruitment in job choice: a different way of looking. Personnel psychology. 44 (3), 487–521.

Salley, H. A. (2022) *New interviewing technologies: what do job applicants think?* PhD thesis. Middle Tennessee State University.

Sánchez-Monedero, J. et al. (2020) 'What does it mean to 'solve' the problem of discrimination in hiring? Social, technical and legal perspectives from the UK on automated hiring systems', *Proceedings of the 2020 conference on fairness, accountability, and transparency,* Barcelona, Spain, 27–30 January, 458–468.

Saxena, N. A. et al. (2020) How do fairness definitions fare? Testing public attitudes towards three algorithmic definitions of fairness in loan allocations. *Artificial intelligence*. 283, 103238.

Scherer, M. U. (2015) Regulating artificial intelligence systems: risks, challenges, competencies, and strategies. *Harvard journal of law & technology*. 29 (2), 354–398.

Schneider, B. (2017) The evolution of airline safety videos. Available at: https://www.bloomberg.com/news/articles/2017-12-22/the-evolution-of-airline-safety-videos (Accessed: 07 May 2023).

Schutte, N. et al. (2000) The factorial validity of the Maslach Burnout Inventory-General Survey (MBI-GS) across occupational groups and nations. *Journal of occupational and organizational psychology*. 73 (1), 53–66.

Selbst, A. D. et al. (2019) 'Fairness and abstraction in sociotechnical systems', *Proceedings of the 2019 conference on fairness, accountability, and transparency,* Atlanta, USA, 29–31 January, 59–68.

Sellers, R. (2014) Video interviewing and its impact on recruitment. *Strategic HR review*. 13 (3), 137.

Shaw, J. C. et al. (2003) To justify or excuse?: A meta-analytic review of the effects of explanations. *Journal of applied psychology*. 88 (3), 444–458.

Schinkel, S. et al. (2016) Applicant reactions to selection events: four studies into the role of attributional style and fairness perceptions. *International journal of selection and assessment*. 24 (2), 107–118.

Schumann, C. et al. (2020) 'We need fairness and explainability in algorithmic hiring', *Proceedings of the 19th international conference on autonomous agents and multiagent systems (AAMAS)*, Auckland, New Zealand, 9–13 May.

Seizov, O. and Wulf, A. J. (2020) Artificial intelligence and transparency: a blueprint for improving the regulation of AI applications in the EU. *European business law review*. 31 (4), 611–640.

Shin, D. (2021) The effects of explainability and causability on perception, trust, and acceptance: implications for explainable AI. *International journal of human-computer studies*. 146, 102551.

Shrestha, Y. et al. (2019) Organizational decision-making structures in the age of artificial intelligence. *California management review*. 61 (4), 66–83.

Shulner-Tal, A. et al. (2023) Enhancing fairness perception - towards human-centred AI and personalized explanations understanding the factors influencing laypeople's fairness perceptions of algorithmic decisions. *International journal of human-computer interaction*. 39 (7), 1455–1482.

Simbeck, K. (2019) HR analytics and ethics. *IBM journal of research and development*. 63(4/5), 1–12.

Skarlicki, D. P. and Folger, R. (1997) Retaliation in the workplace: the roles of distributive, procedural, and interactional justice. *Journal of applied psychology*. 82 (3), 434–443.

Spence, M. (1973) Job Market Signaling. *The quarterly journal of economics*. 87 (3), 355–374.

Steenkamp, J. B. E. et al. (2010) Socially desirable response tendencies in survey research. *Journal of marketing research*. 47 (2), 199–214.

Suen, H. Y. et al. (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.

Suen, H. Y. and Hung, K. E. (2023) Building trust in automatic video interviews using various AI interfaces: tangibility, immediacy, and transparency. *Computers in human behavior*. 143, 107713.

Swider, B. W. et al. (2015) Searching for the right fit: development of applicant person-organization fit perceptions during the recruitment process. *Journal of applied psychology*. 100 (3), 880–893.

Tambe, P. et al. (2019) Artificial intelligence in human resources management: challenges and a path forward. *California management review*. 61 (4), 15–42.

Taylor, B. J. (2006) Factorial surveys: using vignettes to study professional judgment. *The British journal of social work*. 36 (7), 1187–1207.

Tyler, T. R. et al. (1985). The influence of perceived injustice on the endorsement of political leaders. *Journal of applied social psychology*. 15 (8), 700–725.

Tene, O. and Polonetsky, J. (2013). A theory of creepy: technology, privacy and shifting social norms. *Yale journal of law & technology*. 16, 59.

Torres, E. N. and Gregory, A. (2018) Hiring manager's evaluations of asynchronous video interviews: the role of candidate competencies, aesthetics, and resume placement. *International journal of hospitality management*. 75, 86–93.

Torres, E. N. and Mejia, C. (2017) Asynchronous video interviews in the hospitality industry: considerations for virtual employee selection. *International journal of hospitality management*. 61, 4–613.

Truxillo, D. M. et al. (2002) Selection fairness information and applicant reactions: a longitudinal field study. *Journal of applied psychology*. 87 (6), 1020–1031.

Truxillo, D. M. et al. (2009) Effects of explanations on applicant reactions: a meta-analytic review. *International journal of selection and assessment*. 17 (4), 346–361.

van Berkel, N. et al. (2021) 'Effect of information presentation on fairness perceptions of machine learning predictors', *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems, Yokohama, Japan, 8–13 May,* 1–13.

van den Bos, K. et al. (1998) When do we need procedural fairness?: the role of trust in authority. *Journal of personality and social psychology*. 75 (6), 1449–1458.

van den Bos, K. et al. (2001) 'The psychology of procedural and distributive justice viewed from the perspective of fairness heuristic theory', in Cropanzano, R. (ed.) *Justice in the workplace: from theory to practice*. Lawrence Erlbaum Associates Publishers, 49–66.

van Esch, P. et al. (2019) Marketing AI recruitment: the next phase in job application and selection. *Computers in human behavior*. 90, 215–222.

van Esch, P. and Black, J.S. (2019) Factors that influence new generation candidates to engage with and complete digital, AI-enabled recruiting. *Business horizons*. 62 (6), 729–739.

van Esch, P. et al. (2021) Job candidates' reactions to AI-enabled job application processes. *AI and Ethics*. 1, 119–130.

Vasconcelos, M. et al. (2018) 'Modeling epistemological principles for bias mitigation in AI systems: an illustration in hiring decisions', *Proceedings of the 2018 AAAI/ACM Conference on AI, ethics, and society,* New Orleans, USA, 2–3 February, 323–329.

von Krogh, G. (2018) Artificial intelligence in organizations: new opportunities for phenomenon-based theorizing. *Academy of management discoveries*. 4 (4), 404–409.

Wang, R. et al. (2020) 'Factors influencing perceived fairness in algorithmic decision-making: algorithm outcomes, development procedures, and individual differences', *Proceedings of the 2020 CHI conference on human factors in computing systems*, Honolulu, USA, 25–30 April, 1–14.

Walker, H. J. et al. (2013) Is this how I will be treated? Reducing uncertainty through recruitment interactions. *Academy of management journal*. 56 (5), 1325–1347.

Warszta, T. (2012) *Application of Gilliland's model of applicants' reactions to the field of web-based selection*. PhD thesis. Christian-Albrechts Universität Kiel.

White, R. A. (1977) The influence of experimenter motivation, attitudes, and methods of handling subjects on psi test results. *Handbook of parapsychology*. 273–301.

Woods, S. A. et al. (2020) Personnel selection in the digital age: a review of validity and applicant reactions, and future research challenges. *European journal of work and organizational psychology*. 29 (1), 64–77.

Yarger, L. et al. (2020) Algorithmic equity in the hiring of underrepresented IT job candidates. *Online information review*. 44 (2), 383–395.

Yu, L. and Li, Y. (2022) Artificial intelligence decision-making transparency and employees' trust: the parallel multiple mediating effect of effectiveness and discomfort. *Behavioral sciences*. 12 (5), 127.

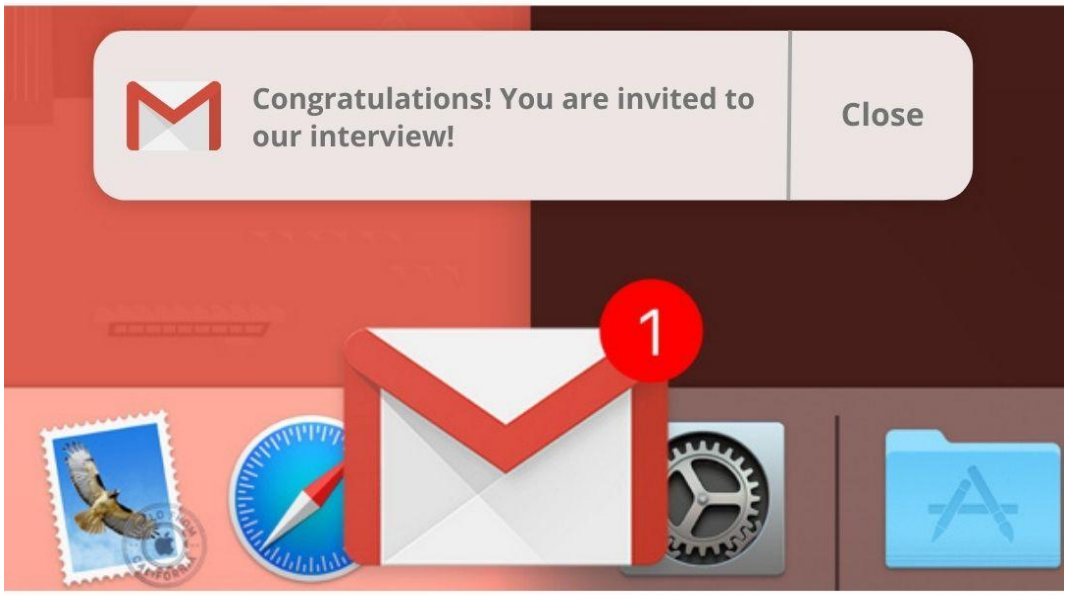
Zhang, P. (2013) The affective response model: a theoretical framework of affective concepts and their relationships in the ict context. *MIS quarterly*. 37 (1), 247–274.

Zhang, X. et al. (2020) 'Developing fairness rules for talent intelligence management system', *Proceedings of the 53rd Hawaii international conference on system sciences*. Maui, USA, 7–10 January, 5882–5891

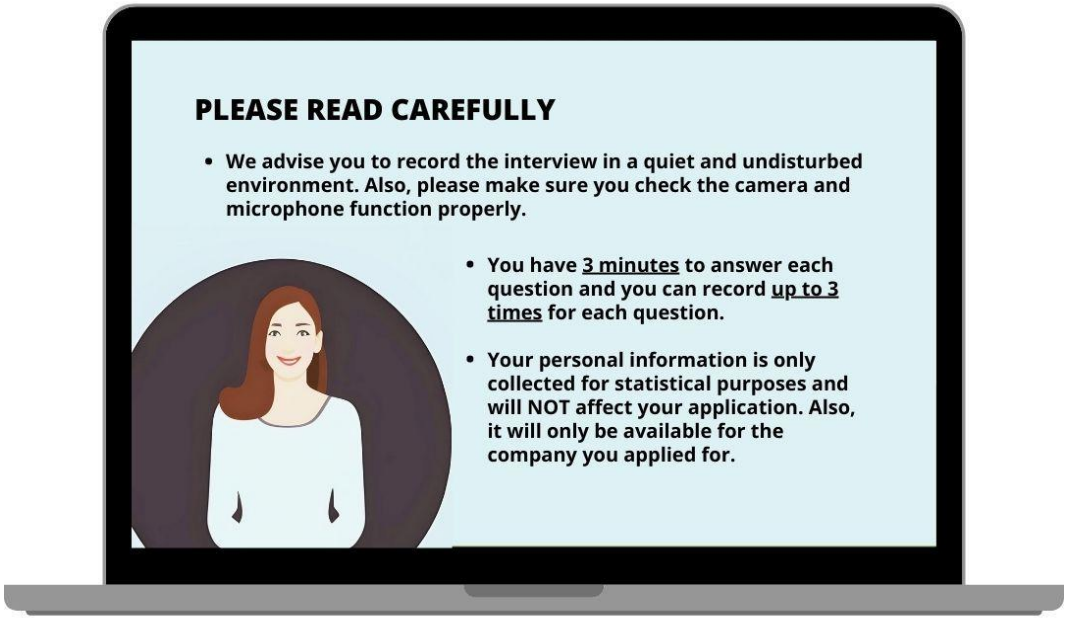
Zielinski, D. (2020) Addressing artificial intelligence-based hiring concerns. *HR magazine*. 1–1.

8. Appendices

Appendix A - Manipulation material

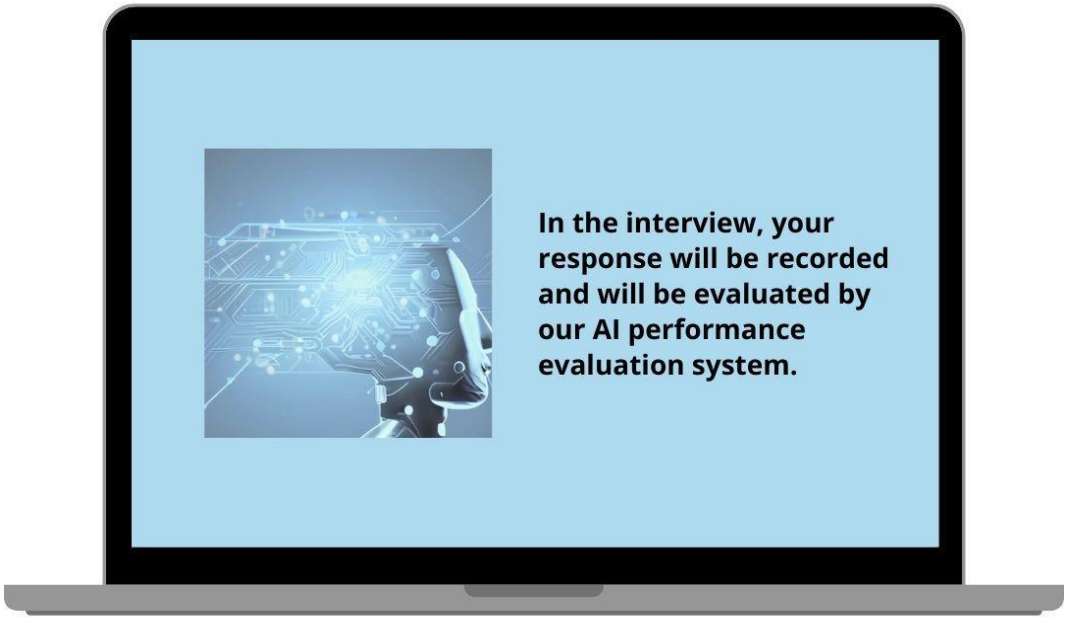


Appendix A.1 Interview notification

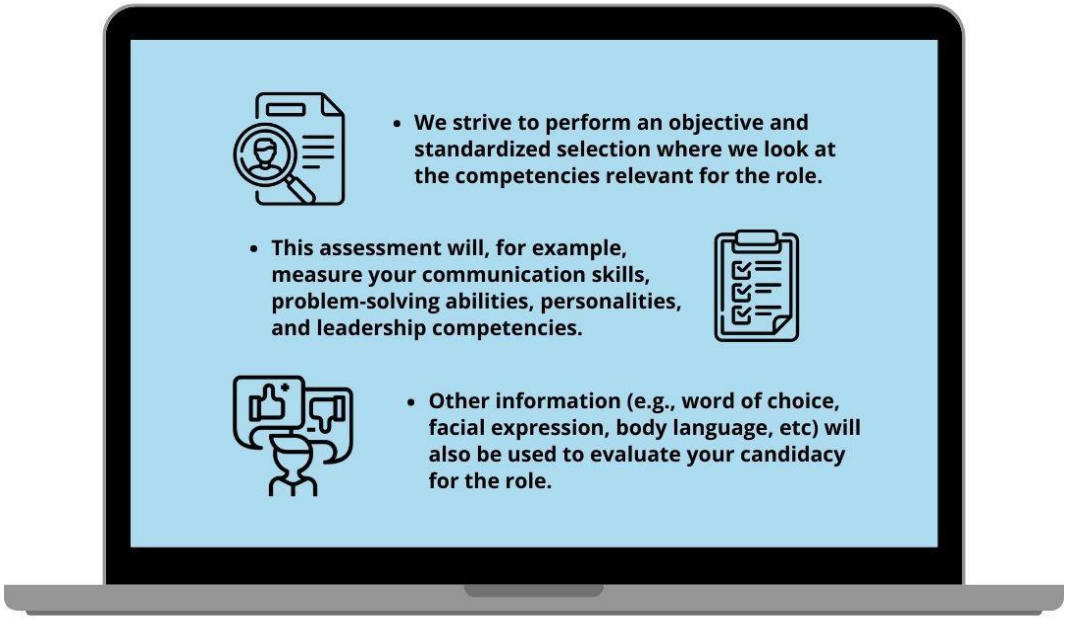


Appendix A.2 Guidance and privacy rules

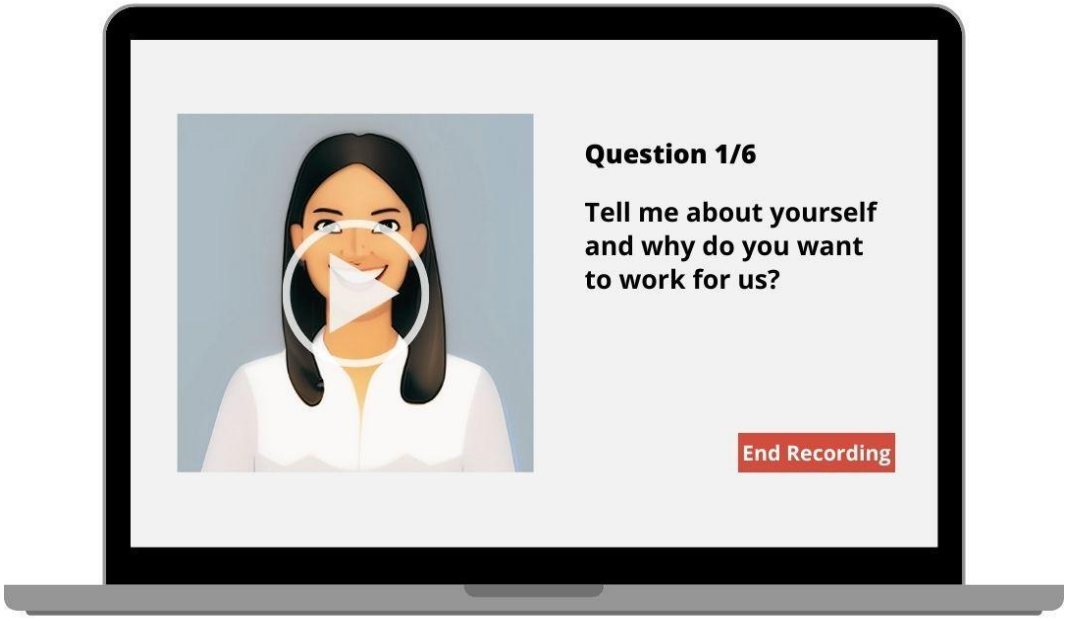
Appendix A.3 Disclosure of human as the decision maker



Appendix A.4 Disclosure of AI as the decision maker



Appendix A.5 Explanations



Appendix A.6 Interview simulation

Appendix B - Questionnaire

Fairness perception items	Source	Scale
I think the shown procedure was fair.	Warszta (2012)	7-point Likert scale (1="Strongly disagree", 7= "Strongly agree")
A person who scored well in this interview will perform well in his/her role.	Bauer et al. (2001)	
This interview gives applicants the opportunity to show what they can really do.	Bauer et al. (2001)	
Applicants would be able to have their interview results reviewed if they wanted.	Bauer et al. (2001)	
Please select "agree" below.*	Self-developed	
The interview was administered to all applicants in the same way.	Bauer et al. (2001)	
I knew what to expect in the interview.	Bauer et al. (2001)	
I was treated honestly and openly during the interview process.	Bauer et al. (2001)	
The selection process was objective and without bias.	Self-developed	

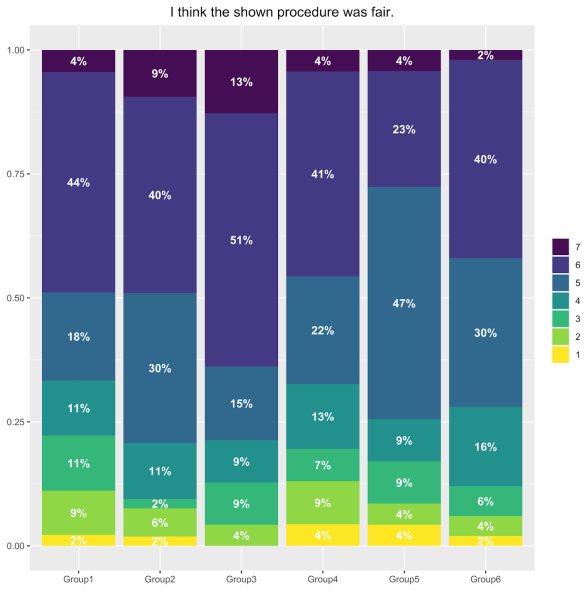
* Attention check item

Appendix B.1 Fairness perception items

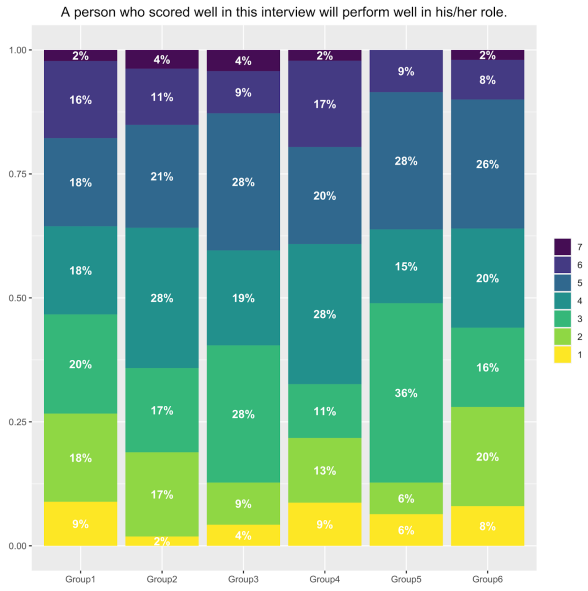
Demographic information	Options
Gender	Male, Female, Non-binary, Prefer not to say
Age	[*Number entry*]
Education level	High school, Bachelor's degree, Master's degree, Other, Prefer not to say
Please indicate your knowledge level of AI	5-point Likert scale (1="Not knowledgeable at all", 5="Extremely knowledgeable")

Appendix B.2 Demographic items

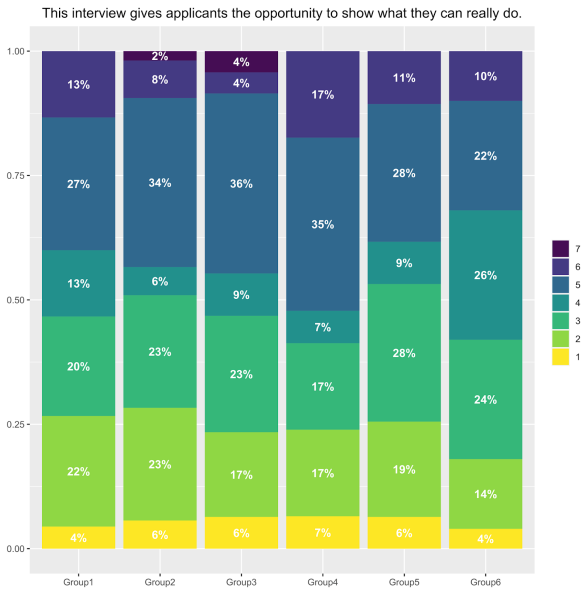
Appendix C - Stacked bar charts based on questionnaire items



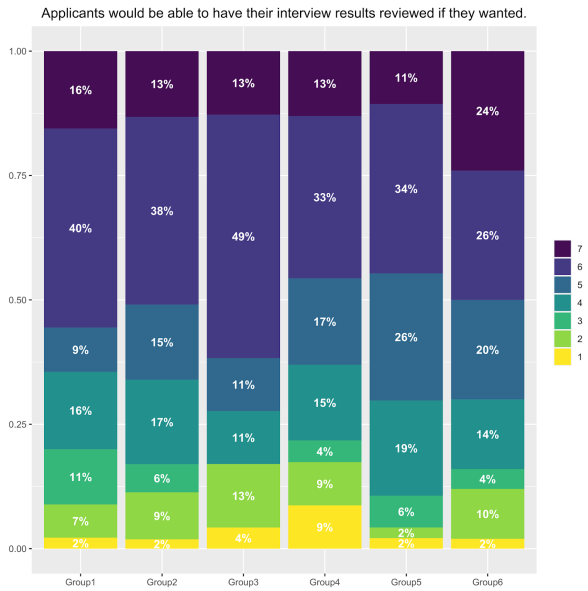
C.1 Overall fairness



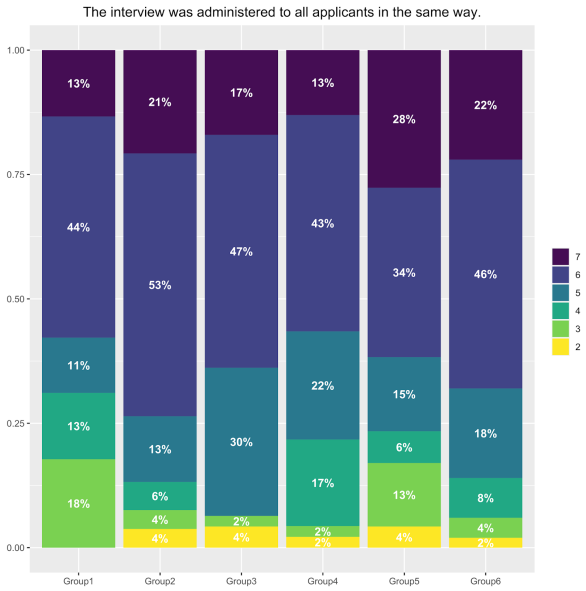
C.2 Job relatedness



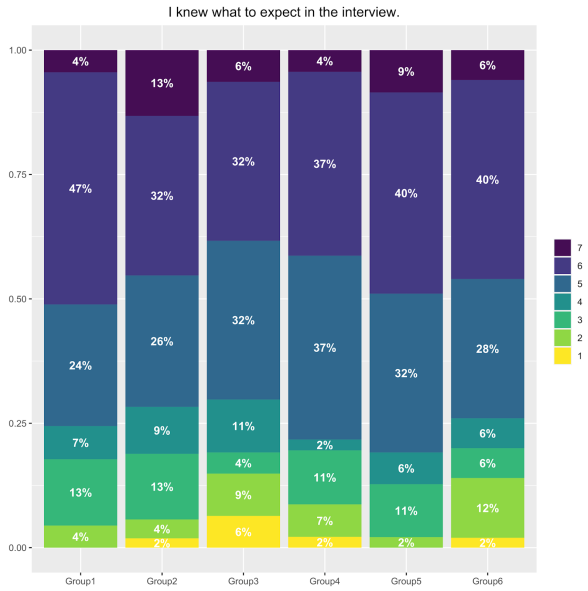
C.3 Opportunity to perform



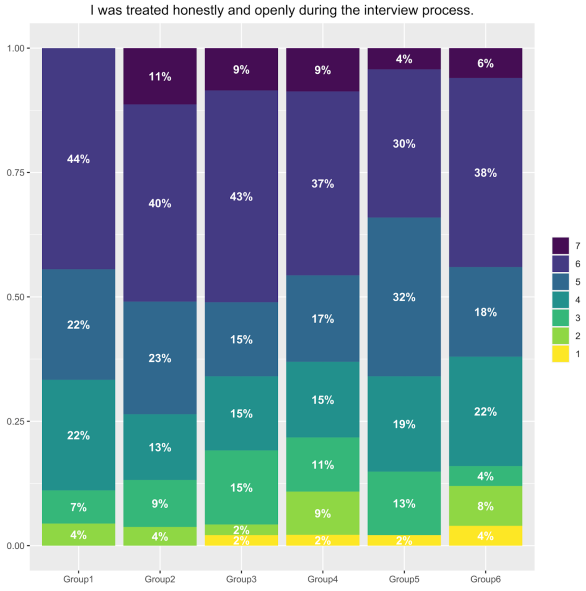
C.4 Reconsideration opportunity



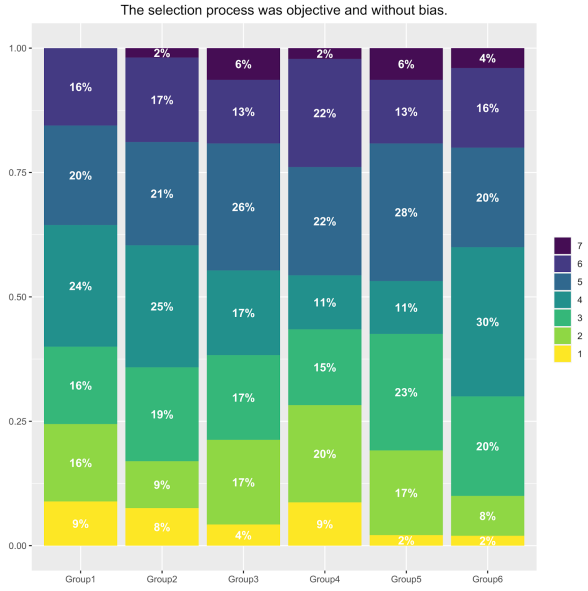
C.5 Consistency



C.6 Selection information



C.7 Honesty (openness)



C.8 Bias suppression

Appendix D - R scripts and survey data

The R scripts used for the whole thesis are provided in the next pages. Output of the R markdown file is directly bound to this appendix. To avoid being lengthy, the scripts in this appendix do not include the output. The data of our survey and the file including both codes and the output can be assessed via this URL to the author's Google Drive folder.

1351 - Master Thesis
Data analysis: A step-by-step explanation

Shao-Fu Lu (42056@student.hhs.se)

In this file, I demonstrate how I process the data from the survey for examination and for replication purposes. However, because it will be extremely lengthy if all the outputs are also shown, I hide the outputs and only show the codes. Still, replication is possible if the codes are run exactly as follows.

Data preparation

Load data

Firstly, I load the raw data in csv that is directly downloaded from Qualtrics. Following the steps in this step-by-step demonstration, the results and analyses can be replicated from this "20230407.csv" file; particularly, 20230407 refers to the date we stopped collecting responses for the survey (see Section 3.3.1).

```
df <- read.csv("20230407.csv")
```

Note: Qualtrics automatically collected the information of respondents' IP addresses and locations (i.e., latitude and longitude). To keep our respondents anonymous, I have removed these information from the file.

Group 1 Unknown - Explanations not provided

A between-subject experiment is conducted in this research and there are six treatment groups (see Table 3.1). Therefore, the following steps will be repeated six times to organize the data and allocate the participants to their corresponding groups.

All questions in the questionnaire are indexed with a unique number. For example, G1Q6.1 refers to the first question for participants in treatment group 1. The number after Q means that it is in the sixth block in the survey. But this number will be different across different groups because of different manipulations (i.e., the information given to different treatment groups before the participants answer the questionnaire).

```
group1 <- df[df$G1Q6.1 != "", ] # extract subjects from group 1
df[1:5, 1:2] # first 5 rows and first 2 columns
```

The first two rows are default information directly exported from Qualtrics.
The next step is to filter out inattentive respondents. The fifth question is an attention check (see Appendix B.1) and the correct answer is the sixth option.

```
nrow(group1[group1$G1Q6.5 != 6, ]) # 7 failed (9 - 2 default rows)
group1 <- group1[group1$G1Q6.5 == 6, ] # remove who failed
group1[, "G1Q6.5"] <- NULL # remove attention check column
```

As mentioned, all the questions are indexed with a unique number. But here, I rename the columns to what they actually stand for (see Appendix B) for the sake of understandability and consistency hereafter.

```
head(group1$G1Q6.1) # The first 6 observations to the question G1Q6.1
library(dplyr)
# Rename the numbers of questions in the questionnaire
group1 <- group1 %>%
  rename("overall" = "G1Q6.1", "relatedness" = "G1Q6.2",
         "perform" = "G1Q6.3", "reconsider" = "G1Q6.4",
         "consistency" = "G1Q6.6", "information" = "G1Q6.7",
         "openness" = "G1Q6.8", "unbiased" = "G1Q6.9",
         "gender" = "G1Q7.1", "age" = "G1Q7.2", "education" = "G1Q7.3")

head(group1$overall) # to show it is the same column as G1Q6.1
```

Also, because all the demographic items are given numeric values, I rename the numeric values of demographics to what they actually mean.

```
head(group1$gender) # The first 6 observations of the variable gender

group1$gender <- ifelse(group1$gender == 1, "Male",
                 ifelse(group1$gender == 2, "Female",
                   ifelse(group1$gender == 3, "Non-binary",
                     "Prefer not to say")))

head(group1$gender) # The first 6 observations of the variable gender
head(group1$education) # The first 6 observations of the variable education

group1$education <- ifelse(group1$education == 1, "High School",
                    ifelse(group1$education == 2, "Bachelor's",
                      ifelse(group1$education == 3, "Master's",
                        ifelse(group1$education == 4, "Other",
                          "Prefer not to say"))))

head(group1$education) # The first 6 observations of the variable education
```

Because the raw data downloaded from Qualtrics separates all treatment groups into different columns, I need to manually remove the empty columns and keep only the ones for group 1.

```
emptycols1 <- colSums(group1 == "") == nrow(group1) # mark the empty columns
group1 <- group1[!emptycols1] # remove the empty columns
```

Remove the column for free text entry because it is coded as “Other” in the previous step. Lastly, I add a column of group number to this dataframe.

```
group1[, "G1Q7.3_4_TEXT"] <- NULL # remove column
group1 <- cbind(group = "Group1", group1, AI_knowledge = NA) # add group number
```

Group 2 Unknown - Explanations provided

The following codes are almost identical as above, except for the question index.

```
group2 <- df[df$G2Q7.1 != "", ] # extract subjects from group 2

# Attention check
nrow(group2[group2$G2Q7.5 != 6, ]) # 1 failed (3 - 2 default rows)
group2 <- group2[group2$G2Q7.5 == 6, ] # remove who failed
group2[, "G2Q7.5"] <- NULL # remove attention check column

# Rename the columns
group2 <- group2 %>%
  rename("overall" = "G2Q7.1", "relatedness" = "G2Q7.2",
         "perform" = "G2Q7.3", "reconsider" = "G2Q7.4",
         "consistency" = "G2Q7.6", "information" = "G2Q7.7",
         "openness" = "G2Q7.8", "unbiased" = "G2Q7.9",
         "gender" = "G2Q8.1", "age" = "G2Q8.2", "education" = "G2Q8.3")

# Rename the values
group2$gender <- ifelse(group2$gender == 1, "Male",
                 ifelse(group2$gender == 2, "Female",
                   ifelse(group2$gender == 3, "Non-binary",
                     "Prefer not to say")))

group2$education <- ifelse(group2$education == 1, "High School",
                    ifelse(group2$education == 2, "Bachelor's",
                      ifelse(group2$education == 3, "Master's",
                        ifelse(group2$education == 4, "Other",
                          "Prefer not to say"))))

emptycols2 <- colSums(group2 == "") == nrow(group2) # mark the empty columns
group2 <- group2[!emptycols2] # remove the empty columns
group2[, "G2Q8.3_4_TEXT"] <- NULL # remove column
group2 <- cbind(group = "Group2", group2, AI_knowledge = NA) # add group number
```

Group 3 Human - Explanations not provided

```
group3 <- df[df$G3Q7.1 != "", ] # extract subjects from group 3

# Attention check
nrow(group3[group3$G3Q7.5 != 6, ]) # 9 failed (11 - 2 default rows)
group3 <- group3[group3$G3Q7.5 == 6, ] # remove who failed
group3[, "G3Q7.5"] <- NULL # remove attention check column

# Rename the columns
group3 <- group3 %>%
  rename("overall" = "G3Q7.1", "relatedness" = "G3Q7.2",
         "perform" = "G3Q7.3", "reconsider" = "G3Q7.4",
         "consistency" = "G3Q7.6", "information" = "G3Q7.7",
         "openness" = "G3Q7.8", "unbiased" = "G3Q7.9",
         "gender" = "G3Q8.1", "age" = "G3Q8.2", "education" = "G3Q8.3")
```

```
# Rename the values
group3$gender <- ifelse(group3$gender == 1, "Male",
                 ifelse(group3$gender == 2, "Female",
                     ifelse(group3$gender == 3, "Non-binary",
                         "Prefer not to say")))

group3$education <- ifelse(group3$education == 1, "High School",
                    ifelse(group3$education == 2, "Bachelor's",
                        ifelse(group3$education == 3, "Master's",
                            ifelse(group3$education == 4, "Other",
                                "Prefer not to say"))))

emptycols3 <- colSums(group3 == "") == nrow(group3) # mark the empty columns
group3 <- group3[!emptycols3] # remove the empty columns
group3[, "G3Q8.3_4_TEXT"] <- NULL # remove column
group3 <- cbind(group = "Group3", group3, AI_knowledge = NA) # add group number
```

Group 4 Human - Explanations provided

```
group4 <- df[df$G4Q8.1 != "", ] # extract subjects from group 4

# Attention check
nrow(group4[group4$G4Q8.5 != 6, ]) # 5 failed (7 - 2 default rows)
group4 <- group4[group4$G4Q8.5 == 6, ] # remove who failed
group4[, "G4Q8.5"] <- NULL # remove attention check column

# Rename the columns
group4 <- group4 %>%
  rename("overall" = "G4Q8.1", "relatedness" = "G4Q8.2",
         "perform" = "G4Q8.3", "reconsider" = "G4Q8.4",
         "consistency" = "G4Q8.6", "information" = "G4Q8.7",
         "openness" = "G4Q8.8", "unbiased" = "G4Q8.9",
         "gender" = "G4Q9.1", "age" = "G4Q9.2", "education" = "G4Q9.3")

# Rename the values
group4$gender <- ifelse(group4$gender == 1, "Male",
                 ifelse(group4$gender == 2, "Female",
                     ifelse(group4$gender == 3, "Non-binary",
                         "Prefer not to say")))

group4$education <- ifelse(group4$education == 1, "High School",
                    ifelse(group4$education == 2, "Bachelor's",
                        ifelse(group4$education == 3, "Master's",
                            ifelse(group4$education == 4, "Other",
                                "Prefer not to say"))))

emptycols4 <- colSums(group4 == "") == nrow(group4) # mark the empty columns
group4 <- group4[!emptycols4] # remove the empty columns
group4[, "G4Q9.3_4_TEXT"] <- NULL # remove column
group4 <- cbind(group = "Group4", group4, AI_knowledge = NA) # add group number
```

Group 5 AI - Explanations not provided

```
group5 <- df[df$G5Q8.1 != "", ] # extract subjects from group 5

# Attention check
nrow(group5[group5$G5Q8.5 != 6, ]) # 3 failed (5 - 2 defaults rows)
group5 <- group5[group5$G5Q8.5 == 6, ] # remove who failed
group5[, "G5Q8.5"] <- NULL # remove attention check column

# Rename the columns
group5 <- group5 %>%
  rename("overall" = "G5Q8.1", "relatedness" = "G5Q8.2",
         "perform" = "G5Q8.3", "reconsider" = "G5Q8.4",
         "consistency" = "G5Q8.6", "information" = "G5Q8.7",
         "openness" = "G5Q8.8", "unbiased" = "G5Q8.9",
         "gender" = "G5Q9.1", "age" = "G5Q9.2",
         "education" = "G5Q9.3", "AI_knowledge" = "G5Q9.4")
```

Note: The item "AI knowledge" is only asked in AI-based conditions (treatment 5 and 6), as explained in Section 3.2.2.2.

```
# Rename the values
group5$gender <- ifelse(group5$gender == 1, "Male",
                 ifelse(group5$gender == 2, "Female",
                   ifelse(group5$gender == 3, "Non-binary",
                     "Prefer not to say")))

group5$education <- ifelse(group5$education == 1, "High School",
                    ifelse(group5$education == 2, "Bachelor's",
                      ifelse(group5$education == 3, "Master's",
                        ifelse(group5$education == 4, "Other",
                          "Prefer not to say"))))

emptycols5 <- colSums(group5 == "") == nrow(group5) # mark empty columns
group5 <- group5[!emptycols5] # remove empty columns
group5[, "G5Q9.3_4_TEXT"] <- NULL # remove column
group5 <- cbind(group = "Group5", group5) # add a group variable
```

Group 6 AI - Explanations provided

```
group6 <- df[df$G6Q9.1 != "", ] # extract subjects from group 6

# Attention check
nrow(group6[group6$G6Q9.5 != 6, ]) # 5 failed (7 - 2 default rows)
group6 <- group6[group6$G6Q9.5 == 6, ] # remove who failed
group6[, "G6Q9.5"] <- NULL # remove attention check column

# Rename the columns
group6 <- group6 %>%
```

```
 rename("overall" = "G6Q9.1", "relatedness" = "G6Q9.2",
        "perform" = "G6Q9.3", "reconsider" = "G6Q9.4",
        "consistency" = "G6Q9.6", "information" = "G6Q9.7",
        "openness" = "G6Q9.8", "unbiased" = "G6Q9.9",
        "gender" = "G6Q10.1", "age" = "G6Q10.2",
        "education" = "G6Q10.3", "AI_knowledge" = "G6Q10.4")

# Rename the values
group6$gender <- ifelse(group6$gender == 1, "Male",
                 ifelse(group6$gender == 2, "Female",
                    ifelse(group6$gender == 3, "Non-binary",
                       "Prefer not to say")))

group6$education <- ifelse(group6$education == 1, "High School",
                    ifelse(group6$education == 2, "Bachelor's",
                       ifelse(group6$education == 3, "Master's",
                          ifelse(group6$education == 4, "Other",
                             "Prefer not to say"))))

emptycols6 <- colSums(group6 == "") == nrow(group6) # mark empty columns
group6 <- group6[!emptycols6] # remove empty columns
group6[, "G6Q10.3_4_TEXT"] <- NULL # remove column
group6 <- cbind(group = "Group6", group6) # add a group variable
```

Merge six treatment groups

After I have created six tables for all six groups, I merge them into a concise dataframe (i.e., "all") that only includes the information I need for the analyses.

```
df_merge <- rbind(group1, group2, group3, group4, group5, group6)

# convert the values to numeric
df_merge$overall <- as.numeric(df_merge$overall)
df_merge$relatedness <- as.numeric(df_merge$relatedness)
df_merge$perform <- as.numeric(df_merge$perform)
df_merge$reconsider <- as.numeric(df_merge$reconsider)
df_merge$consistency <- as.numeric(df_merge$consistency)
df_merge$information <- as.numeric(df_merge$information)
df_merge$openness <- as.numeric(df_merge$openness)
df_merge$unbiased <- as.numeric(df_merge$unbiased)
df_merge$age <- as.numeric(df_merge$age, na.rm = T)
df_merge$AI_knowledge <- as.numeric(df_merge$AI_knowledge, na.rm = T)

# Only keep useful variables
all <- df_merge[, c("group", "overall", "relatedness", "perform", "reconsider",
                    "consistency", "information", "openness", "unbiased",
                    "gender", "age", "education", "AI_knowledge")]
```

Demographic and treatment group distribution

The results are presented in Section 3.3.2 Participants.

Age

```
round(mean(all$age, na.rm = TRUE), digits = 2) # mean
round(sd(all$age, na.rm = TRUE), digits = 2)  # sd
min(all$age, na.rm = TRUE) # minimum
max(all$age, na.rm = TRUE) # maximum
```

Gender distribution

```
table(all$gender)
```

Education distribution

```
table(all$education)
```

AI Knowledge

```
round(mean(all$AI_knowledge, na.rm = TRUE), digits = 2)
round(sd(all$AI_knowledge, na.rm = TRUE), digits = 2)
table(all$AI_knowledge)
```

Treatment group distribution (Table 3.2)

```
table(all$group)
```

Results and analyses

4.1 Measure testing

```
library(ltm)
alpha <- all[, c("relatedness", "perform", "reconsider", "consistency",
                 "information", "openness", "unbiased")]
cronbach.alpha(alpha)
```

Correlation matrix

The results are presented in Table 4.1.

```
cor <- all

# 0 = unknown, 1 = known
cor_known <- ifelse(cor$group == "Group1" | cor$group == "Group2", 0, 1)

# 0 = unknown, 1 = human, 2 = AI
cor_group <- ifelse(cor$group == "Group1" | cor$group == "Group2", 0,
                    ifelse(cor$group == "Group3" | cor$group == "Group4", 1, 2))
```

```
# 0 = without explanation, 1 = with explanation
cor$group <- ifelse(cor$group == "Group1" | cor$group == "Group3"
                    | cor$group == "Group5", 0, 1)

cor <- cor %>%
 rename("explanation" = "group")

cor <- cbind(known=cor_known, decision.maker=cor_group, cor)

cor$gender <- ifelse(cor$gender == "Male", 1,
                     ifelse(cor$gender == "Female", 2,
                          ifelse(cor$gender == "Non-binary", 3, 4)))

cor$education <- ifelse(cor$education == "High School", 1,
                        ifelse(cor$education == "Bachelor's", 2,
                             ifelse(cor$education == "Master's", 3,
                                  ifelse(cor$education == "Other", 4, 5))))
```

Correlations and p-values are available from running these code as follows.

```
cor_2 <- rcorr(as.matrix(cor))
r = data.frame(cor_2$r) # r value
round(cor_2$r, digits = 2)
p = data.frame(cor_2$P) # p value
round(cor_2$P, digits = 3)
```

Alternatively, r and p values can be saved in two separate tables.

```
write.csv(round(r, digits = 2), file = "Correlation_r.csv")
write.csv(round(p, digits = 4), file = "Correlation_p.csv")
```

4.2 Potential covariate analysis

The results are summarized in Table 4.2.

```
covariate <- all[, c("group", "gender", "age", "education", "AI_knowledge")]

covariate$gender <- ifelse(covariate$gender == "Male", 1,
                    ifelse(covariate$gender == "Female", 2,
                         ifelse(covariate$gender == "Non-binary", 3, 4)))

covariate$education <- ifelse(covariate$education == "High School", 1,
                       ifelse(covariate$education == "Bachelor's", 2,
                            ifelse(covariate$education == "Master's", 3,
                                 ifelse(covariate$education == "Other", 4, 5))))

# Fit an analysis of variance (ANOVA) model
summary(aov(gender ~ group, data = covariate))
summary(aov(age ~ group, data = covariate))
```

```
summary(aov(education ~ group, data = covariate))
summary(aov(AI_knowledge ~ group, data = covariate))

 # Pearson's Chi-squared Test
chisq.test(covariate$group, covariate$gender)
chisq.test(covariate$group, covariate$age)
chisq.test(covariate$group, covariate$education)
chisq.test(covariate[!is.na(covariate$AI_knowledge), ]$group,
        covariate[!is.na(covariate$AI_knowledge), ]$AI_knowledge)
```

4.3 Hypothesis testing

Hypothesis 1

Firstly, t-tests are used to analyze the differences between human and AI group in the conditions without explanations (Table 4.3).

```
library(lsr)

# t-tests between group 3 and group 5 and Cohen's d effect size
t.test(all[all$group=="Group3", ]$overall, all[all$group=="Group5", ]$overall)
round(sd(all[all$group=="Group3", ]$overall), digits = 2)
round(sd(all[all$group=="Group5", ]$overall), digits = 2)
cohensD(all[all$group=="Group3", ]$overall, all[all$group=="Group5", ]$overall)

t.test(all[all$group=="Group3", ]$relatedness, all[all$group=="Group5", ]$relatedness)
round(sd(all[all$group=="Group3", ]$relatedness), digits = 2)
round(sd(all[all$group=="Group5", ]$relatedness), digits = 2)
cohensD(all[all$group=="Group3", ]$relatedness, all[all$group=="Group5", ]$relatedness)

t.test(all[all$group=="Group3", ]$perform, all[all$group=="Group5", ]$perform)
round(sd(all[all$group=="Group3", ]$perform), digits = 2)
round(sd(all[all$group=="Group5", ]$perform), digits = 2)
cohensD(all[all$group=="Group3", ]$perform, all[all$group=="Group5", ]$perform)

t.test(all[all$group=="Group3", ]$reconsider, all[all$group=="Group5", ]$reconsider)
round(sd(all[all$group=="Group3", ]$reconsider), digits = 2)
round(sd(all[all$group=="Group5", ]$reconsider), digits = 2)
cohensD(all[all$group=="Group3", ]$reconsider, all[all$group=="Group5", ]$reconsider)

t.test(all[all$group=="Group3", ]$consistency, all[all$group=="Group5", ]$consistency)
round(sd(all[all$group=="Group3", ]$consistency), digits = 2)
round(sd(all[all$group=="Group5", ]$consistency), digits = 2)
cohensD(all[all$group=="Group3", ]$consistency, all[all$group=="Group5", ]$consistency)

t.test(all[all$group=="Group3", ]$information, all[all$group=="Group5", ]$information)
round(sd(all[all$group=="Group3", ]$information), digits = 2)
round(sd(all[all$group=="Group5", ]$information), digits = 2)
cohensD(all[all$group=="Group3", ]$information, all[all$group=="Group5", ]$information)
```

```
t.test(all[all$group=="Group3", ]$openness, all[all$group=="Group5", ]$openness)
round(sd(all[all$group=="Group3", ]$openness), digits = 2)
round(sd(all[all$group=="Group5", ]$openness), digits = 2)
cohensD(all[all$group=="Group3", ]$openness, all[all$group=="Group5", ]$openness)

t.test(all[all$group=="Group3", ]$unbiased, all[all$group=="Group5", ]$unbiased)
round(sd(all[all$group=="Group3", ]$unbiased), digits = 2)
round(sd(all[all$group=="Group5", ]$unbiased), digits = 2)
cohensD(all[all$group=="Group3", ]$unbiased, all[all$group=="Group5", ]$unbiased)
```

Hypothesis 2

Next, t-tests are used to analyze the differences between unknown and known decision makers in the conditions without explanations (Table 4.4).

```
# t-tests between group 1 and group 3+5 and Cohen's d effect size
group3and5 <- all[all$group == "Group3" | all$group == "Group5", ]

t.test(as.numeric(group1$overall), group3and5$overall)
round(sd(as.numeric(group1$overall)), digits = 2)
round(sd(group3and5$overall), digits = 2)
cohensD(as.numeric(group1$overall), group3and5$overall)

t.test(as.numeric(group1$relatedness), group3and5$relatedness)
round(sd(as.numeric(group1$relatedness)), digits = 2)
round(sd(group3and5$relatedness), digits = 2)
cohensD(as.numeric(group1$relatedness), group3and5$relatedness)

t.test(as.numeric(group1$perform), group3and5$perform)
round(sd(as.numeric(group1$perform)), digits = 2)
round(sd(group3and5$perform), digits = 2)
cohensD(as.numeric(group1$perform), group3and5$perform)

t.test(as.numeric(group1$reconsider), group3and5$reconsider)
round(sd(as.numeric(group1$reconsider)), digits = 2)
round(sd(group3and5$reconsider), digits = 2)
cohensD(as.numeric(group1$reconsider), group3and5$reconsider)

t.test(as.numeric(group1$consistency), group3and5$consistency)
round(sd(as.numeric(group1$consistency)), digits = 2)
round(sd(group3and5$consistency), digits = 2)
cohensD(as.numeric(group1$consistency), group3and5$consistency)

t.test(as.numeric(group1$information), group3and5$information)
round(sd(as.numeric(group1$information)), digits = 2)
round(sd(group3and5$information), digits = 2)
cohensD(as.numeric(group1$information), group3and5$information)

t.test(as.numeric(group1$openness), group3and5$openness)
round(sd(as.numeric(group1$openness)), digits = 2)
```

```
round(sd(group3and5$openness), digits = 2)
cohensD(as.numeric(group1$openness), group3and5$openness)

t.test(as.numeric(group1$unbiased), group3and5$unbiased)
round(sd(as.numeric(group1$unbiased)), digits = 2)
round(sd(group3and5$unbiased), digits = 2)
cohensD(as.numeric(group1$unbiased), group3and5$unbiased)
```

Hypothesis 3

Moderated hierarchical multiple regression analysis

To test hypothesis 3, a two-step regression analysis is applied to analyzing the moderating effects of explanations.

```
moderator <- all

# 0 = no explanation, 1 = with explanation
moderator$explain <- ifelse(moderator$group=="Group1" | moderator$group=="Group3" |
                      moderator$group=="Group5", 0, 1)

moderator$group <- ifelse(moderator$group=="Group1" |
                    moderator$group=="Group2", "Unknown",
                  ifelse(moderator$group=="Group3" |
                       moderator$group=="Group4", "Human", "AI"))

moderator$group <- as.factor(moderator$group)
moderator$group <- relevel(moderator$group, ref="Unknown") # baseline
mod.a1 <- lm(overall ~ group + explain, data = moderator)
mod.a2 <- lm(overall ~ group * explain, data = moderator) # interaction
mod.b1 <- lm(relatedness ~ group + explain, data = moderator)
mod.b2 <- lm(relatedness ~ group * explain, data = moderator) # interaction
mod.c1 <- lm(perform ~ group + explain, data = moderator)
mod.c2 <- lm(perform ~ group * explain, data = moderator) # interaction
mod.d1 <- lm(reconsider ~ group + explain, data = moderator)
mod.d2 <- lm(reconsider ~ group * explain, data = moderator) # interaction
mod.e1 <- lm(consistency ~ group + explain, data = moderator)
mod.e2 <- lm(consistency ~ group * explain, data = moderator) # interaction
mod.f1 <- lm(information ~ group + explain, data = moderator)
mod.f2 <- lm(information ~ group * explain, data = moderator) # interaction
mod.g1 <- lm(openness ~ group + explain, data = moderator)
mod.g2 <- lm(openness ~ group * explain, data = moderator) # interaction
mod.h1 <- lm(unbiased ~ group + explain, data = moderator)
mod.h2 <- lm(unbiased ~ group * explain, data = moderator) # interaction
```

The results shown in Table 4.5 are from this table named “moderation.html”.

```
library(texreg)
htmlreg(list(mod.a1, mod.a2, mod.b1, mod.b2, mod.c1, mod.c2, mod.d1, mod.d2,
        mod.e1, mod.e2, mod.f1, mod.f2, mod.g1, mod.g2, mod.h1, mod.h2),
```

```
    stars = c(0.1, 0.05, 0.01, 0.001),
    digits = 3, star.symbol = "*", symbol = "†",
    custom.model.names = c("overall", "overall.2",
                           "relatedness", "relatedness.2",
                           "perform", "perform.2",
                           "reconsider", "reconsider.2",
                           "consistency", "consistency.2",
                           "information", "information.2",
                           "openness", "openness.2",
                           "unbiased", "unbiased.2"),
    file = "moderation.html")
```

4.5 Subgroup analysis

The results are presented in Table 4.7.

Subgroup_gender

```
male <- subset(all, all$gender=="Male")
female <- subset(all, all$gender=="Female")

t.test(male$overall, female$overall)
round(sd(male$overall), digits = 2)
round(sd(female$overall), digits = 2)
round(cohensD(male$overall, female$overall), digits = 2)

t.test(male$relatedness, female$relatedness)
round(sd(male$relatedness), digits = 2)
round(sd(female$relatedness), digits = 2)
round(cohensD(male$relatedness, female$relatedness), digits = 2)

t.test(male$perform, female$perform)
round(sd(male$perform), digits = 2)
round(sd(female$perform), digits = 2)
round(cohensD(male$perform, female$perform), digits = 2)

t.test(male$reconsider, female$reconsider)
round(sd(male$reconsider), digits = 2)
round(sd(female$reconsider), digits = 2)
round(cohensD(male$reconsider, female$reconsider), digits = 2)

t.test(male$consistency, female$consistency)
round(sd(male$consistency), digits = 2)
round(sd(female$consistency), digits = 2)
round(cohensD(male$consistency, female$consistency), digits = 2)

t.test(male$information, female$information)
round(sd(male$information), digits = 2)
round(sd(female$information), digits = 2)
round(cohensD(male$information, female$information), digits = 2)
```

```
t.test(male$openness, female$openness)
round(sd(male$openness), digits = 2)
round(sd(female$openness), digits = 2)
round(cohensD(male$openness, female$openness), digits = 2)

t.test(male$unbiased, female$unbiased)
round(sd(male$unbiased), digits = 2)
round(sd(female$unbiased), digits = 2)
round(cohensD(male$unbiased, female$unbiased), digits = 2)
```

Subgroup_age

```
mean(all$age, na.rm = T)

old <- subset(all, all$age>24.5 & !is.na(all$age))
young <- subset(all, all$age<24.5 & !is.na(all$age))

t.test(old$overall, young$overall)
round(sd(old$overall), digits = 2)
round(sd(young$overall), digits = 2)
round(cohensD(old$overall, young$overall), digits = 2)

t.test(old$relatedness, young$relatedness)
round(sd(old$relatedness), digits = 2)
round(sd(young$relatedness), digits = 2)
round(cohensD(old$relatedness, young$relatedness), digits = 2)

t.test(old$perform, young$perform)
round(sd(old$perform), digits = 2)
round(sd(young$perform), digits = 2)
round(cohensD(old$perform, young$perform), digits = 2)

t.test(old$reconsider, young$reconsider)
round(sd(old$reconsider), digits = 2)
round(sd(young$reconsider), digits = 2)
round(cohensD(old$reconsider, young$reconsider), digits = 2)

t.test(old$consistency, young$consistency)
round(sd(old$consistency), digits = 2)
round(sd(young$consistency), digits = 2)
round(cohensD(old$consistency, young$consistency), digits = 2)

t.test(old$information, young$information)
round(sd(old$information), digits = 2)
round(sd(young$information), digits = 2)
round(cohensD(old$information, young$information), digits = 2)

t.test(old$openness, young$openness)
round(sd(old$openness), digits = 2)
```

```
round(sd(young$openness), digits = 2)
round(cohensD(old$openness, young$openness), digits = 2)

t.test(old$unbiased, young$unbiased)
round(sd(old$unbiased), digits = 2)
round(sd(young$unbiased), digits = 2)
round(cohensD(old$unbiased, young$unbiased), digits = 2)
```

Subgroup_education

```
bachelor <- subset(all, all$education=="Bachelor's")
master <- subset(all, all$education=="Master's")

t.test(bachelor$overall, master$overall)
round(sd(bachelor$overall), digits = 2)
round(sd(master$overall), digits = 2)
round(cohensD(bachelor$overall, master$overall), digits = 2)

t.test(bachelor$relatedness, master$relatedness)
round(sd(bachelor$relatedness), digits = 2)
round(sd(master$relatedness), digits = 2)
round(cohensD(bachelor$relatedness, master$relatedness), digits = 2)

t.test(bachelor$perform, master$perform)
round(sd(bachelor$perform), digits = 2)
round(sd(master$perform), digits = 2)
round(cohensD(bachelor$perform, master$perform), digits = 2)

t.test(bachelor$reconsider, master$reconsider)
round(sd(bachelor$reconsider), digits = 2)
round(sd(master$reconsider), digits = 2)
round(cohensD(bachelor$reconsider, master$reconsider), digits = 2)

t.test(bachelor$consistency, master$consistency)
round(sd(bachelor$consistency), digits = 2)
round(sd(master$consistency), digits = 2)
round(cohensD(bachelor$consistency, master$consistency), digits = 2)

t.test(bachelor$information, master$information)
round(sd(bachelor$information), digits = 2)
round(sd(master$information), digits = 2)
round(cohensD(bachelor$information, master$information), digits = 2)

t.test(bachelor$openness, master$openness)
round(sd(bachelor$openness), digits = 2)
round(sd(master$openness), digits = 2)
round(cohensD(bachelor$openness, master$openness), digits = 2)

t.test(bachelor$unbiased, master$unbiased)
round(sd(bachelor$unbiased), digits = 2)
```

```
round(sd(master$unbiased), digits = 2)
round(cohensD(bachelor$unbiased, master$unbiased), digits = 2)
```

Data Visualizations

Figure 4.1

Distribution of survey results (human versus AI).

```
library(tidyr)
library(ggplot2)
library(scales)
library(tidyverse)

h1 <- group3and5[c("group", "overall", "relatedness",
                   "perform", "reconsider", "consistency",
                   "information", "openness", "unbiased")]

h1 <- cbind(subject = 1:94, h1)
h1.gather <- gather(h1, "item", "value", 3:10)

h1.vasual <- h1.gather %>%
 group_by(group, item, value) %>%
 count(name = "count") %>%
 group_by(group, item) %>%
 mutate(percent = count/sum(count)) %>%
 ungroup() %>%
 mutate(percentage = percent(percent, accuracy = 1)) %>%
 mutate(value = fct_relevel(factor(value),
                    "1", "2", "3", "4", "5", "6", "7"),
      value = fct_rev(value))

h1.vasual$item <- factor(h1.vasual$item,
             levels = c("group", "overall", "relatedness",
                    "perform", "reconsider", "consistency",
                    "information", "openness", "unbiased"))

h1.vasual$group <- ifelse(h1.vasual$group == "Group3", "Human", "AI")

h1.diverging <- h1.vasual %>%
 mutate(percent = if_else(value %in% c("1", "2", "3"), -percent, percent)) %>%
 mutate(percentage = percent(percent, accuracy = 1))

h1.diverging.2 <- h1.diverging %>%
 mutate(percentage = abs(percent)) %>%
 mutate(percentage = percent(percentage, accuracy = 1)) %>%
 mutate(value = fct_relevel(factor(value),
                    "3", "2", "1", "4", "5", "6", "7"),
      value = fct_rev(value))
```

```
ggplot(h1.diverging.2) +
 aes(x = group, y = percent, fill = value) +
 geom_col() +
 geom_text(aes(label = percentage),
          position = position_stack(vjust = 0.5),
          color = "white",
          fontface = "bold") +
 coord_flip() +
 scale_fill_viridis_d(breaks=c("1", "2", "3", "4", "5", "6", "7")) +
 guides(fill = guide_legend(nrow = 1)) +
 labs(x = "", y = NULL, fill = NULL) +
 facet_grid(rows = vars(item)) +
 theme_minimal() +
 theme(axis.text.x = element_blank(),
      legend.position="top")
```

Alternatively, an image can be saved with the following codes.

```
png(filename = "H1_visual.png",
   unit  = "cm", width = 15, height = 18,
   res = 500)

ggplot(h1.diverging.2) +
 aes(x = group, y = percent, fill = value) +
 geom_col() +
 geom_text(aes(label = percentage),
          position = position_stack(vjust = 0.5),
          color = "white",
          fontface = "bold") +
 coord_flip() +
 scale_fill_viridis_d(breaks=c("1", "2", "3", "4", "5", "6", "7")) +
 guides(fill = guide_legend(nrow = 1)) +
 labs(x = "", y = NULL, fill = NULL) +
 facet_grid(rows = vars(item)) +
 theme_minimal() +
 theme(axis.text.x = element_blank(),
      legend.position="top")

dev.off()
```

Figure 4.2

Distribution of survey results (unknown versus known).

```
group135 <- all[all$group=="Group1"|all$group=="Group3"|all$group=="Group5",
            c("group", "overall", "relatedness", "perform", "reconsider",
              "consistency", "information", "openness", "unbiased")]

h2 <- cbind(subject = 1:nrow(group135), group135)
h2.gather <- gather(h2, "item", "value", 3:10)
```

```
h2.gather$group <- ifelse(h2.gather$group == "Group1", "Unknown", "Known")

h2.vasual <- h2.gather %>%
 group_by(group, item, value) %>%
 count(name = "count") %>%
 group_by(group, item) %>%
 mutate(percent = count/sum(count)) %>%
 ungroup() %>%
 mutate(percentage = percent(percent, accuracy = 1)) %>%
 mutate(value = fct_relevel(factor(value),
                    "1", "2", "3", "4", "5", "6", "7"),
      value = fct_rev(value))

h2.vasual$item <- factor(h2.vasual$item,
                  levels = c("group", "overall", "relatedness",
                         "perform", "reconsider", "consistency",
                         "information", "openness", "unbiased"))

h2.diverging <- h2.vasual %>%
 mutate(percent = if_else(value %in% c("1", "2", "3"), -percent, percent)) %>%
 mutate(percentage = percent(percent, accuracy = 1))

h2.diverging.2 <- h2.diverging %>%
 mutate(percentage = abs(percent)) %>%
 mutate(percentage = percent(percentage, accuracy = 1)) %>%
 mutate(value = fct_relevel(factor(value),
                    "3", "2", "1", "4", "5", "6", "7"),
      value = fct_rev(value))

ggplot(h2.diverging.2) +
 aes(x = group, y = percent, fill = value) +
 geom_col() +
 geom_text(aes(label = percentage),
         position = position_stack(vjust = 0.5),
         color = "white",
         fontface = "bold") +
 coord_flip() +
 scale_fill_viridis_d(breaks=c("1", "2", "3", "4", "5", "6", "7")) +
 guides(fill = guide_legend(nrow = 1)) +
 labs(x= NULL, y = NULL, fill = NULL) +
 facet_grid(rows = vars(item)) +
 theme_minimal() +
 theme(axis.text.x = element_blank(),
     legend.position="top")
```

Alternatively, an image can be saved with the following codes.

```
png(filename = "H2_visual.png",
   unit  = "cm", width = 15, height = 18,
   res = 500)
```

```
ggplot(h2.diverging.2) +
 aes(x = group, y = percent, fill = value) +
 geom_col() +
 geom_text(aes(label = percentage),
          position = position_stack(vjust = 0.5),
          color = "white",
          fontface = "bold") +
 coord_flip() +
 scale_fill_viridis_d(breaks=c("1", "2", "3", "4", "5", "6", "7")) +
 guides(fill = guide_legend(nrow = 1)) +
 labs(x= NULL, y = NULL, fill = NULL) +
 facet_grid(rows = vars(item)) +
 theme_minimal() +
 theme(axis.text.x = element_blank(),
      legend.position="top")

dev.off()
```

Figure 4.3

Moderating effects of explanations.

```
library(emmeans)
mod.overall <- emmip(mod.a2, explain~group, CIs = T,
     xlab="", ylab="", tlab="Explain",
     engine="ggplot") +
 ylim(3, 6) +
 ggtitle("Overall") +
 scale_color_manual(labels = c("No", "Yes"), values=c("#FE6DB6","#0078D7")) +
 ggthemes::theme_clean()

mod.relatedness <- emmip(mod.b2, explain~group, CIs = T,
          xlab="", ylab="", tlab="Explain",
          engine="ggplot") +
 ylim(3, 6) +
 ggtitle("Relatedness") +
 scale_color_manual(labels = c("No", "Yes"), values=c("#FE6DB6","#0078D7")) +
 ggthemes::theme_clean()

mod.perform <- emmip(mod.c2, explain~group, CIs=T,
          xlab="", ylab="", tlab="Explain",
          engine="ggplot") +
 ylim(3, 6) +
 ggtitle("Perform") +
 scale_color_manual(labels = c("No", "Yes"), values=c("#FE6DB6","#0078D7")) +
 ggthemes::theme_clean()

mod.reconsider <- emmip(mod.d2, explain~group, CIs=T,
          xlab="", ylab="", tlab="Explain",
```

```
        engine="ggplot") +
 ylim(3, 6) +
 ggtitle("Reconsider") +
 scale_color_manual(labels = c("No", "Yes"), values=c("#FE6DB6","#0078D7")) +
 ggthemes::theme_clean()

mod.consistency <- emmip(mod.e2, explain~group, CIs=T,
        xlab="", ylab="", tlab="Explain",
        engine="ggplot") +
 ylim(3, 6) +
 ggtitle("Consistency") +
 scale_color_manual(labels = c("No", "Yes"), values=c("#FE6DB6","#0078D7")) +
 ggthemes::theme_clean()

mod.information <- emmip(mod.f2, explain~group, CIs=T,
        xlab="", ylab="", tlab="Explain",
        engine="ggplot") +
 ylim(3, 6) +
 ggtitle("Information") +
 scale_color_manual(labels = c("No", "Yes"), values=c("#FE6DB6","#0078D7")) +
 ggthemes::theme_clean()

mod.openness <- emmip(mod.g2, explain~group, CIs=T,
        xlab="", ylab="", tlab="Explain",
        engine="ggplot") +
 ylim(3, 6) +
 ggtitle("Openness") +
 scale_color_manual(labels = c("No", "Yes"), values=c("#FE6DB6","#0078D7")) +
 ggthemes::theme_clean()

mod.unbiased <- emmip(mod.h2, explain~group, CIs=T,
        xlab="", ylab="", tlab="Explain",
        engine="ggplot") +
 ylim(3, 6) +
 ggtitle("Unbiased") +
 scale_color_manual(labels = c("No", "Yes"), values=c("#FE6DB6","#0078D7")) +
 ggthemes::theme_clean()

# merge 8 in 1
library(ggpubr)
ggarrange(mod.overall, mod.relatedness,
        mod.perform, mod.reconsider,
        mod.consistency, mod.information,
        mod.openness, mod.unbiased,
        ncol = 2, nrow = 4,
        common.legend = T, legend = "top")
```

Alternatively, an image can be saved with the following codes.

```
png(filename = "moderation_plot.png",
   unit  = "cm", width = 15, height = 21,
   res = 500)

ggarrange(mod.overall, mod.relatedness,
        mod.perform, mod.reconsider,
        mod.consistency, mod.information,
        mod.openness, mod.unbiased,
        ncol = 2, nrow = 4,
        common.legend = T, legend = "top")

dev.off() # save as a file
```

Figure 4.4

Violin plot for male and female.

```
gender <- subset(all, all$gender=="Male" | all$gender=="Female")
gender$gender <- factor(gender$gender,
                  levels = c("Male", "Female"))

gender.1 <- gender %>%
 ggplot(aes(x=gender, y=overall, color=gender)) +
 geom_violin() +
 labs(title="",x="Overall", y="") +
 geom_boxplot(width=0.1, color="black", outlier.size=1) +
 scale_color_manual(values=c("#0078D7", "#FE6DB6")) +
 ylim(0.5, 7.5) +
 ggthemes::theme_clean() +
 theme(axis.text.x = element_blank(),
     legend.position="right")

gender.2 <- gender %>%
 ggplot(aes(x=gender, y=relatedness, color=gender)) +
 geom_violin() +
 labs(title="",x="Relatedness", y="") +
 geom_boxplot(width=0.1, color="black") +
 scale_color_manual(values=c("#0078D7", "#FE6DB6")) +
 ylim(0.5, 7.5) +
 ggthemes::theme_clean() +
 theme(axis.text.x = element_blank(),
     legend.position="right")

gender.3 <- gender %>%
 ggplot(aes(x=gender, y=perform, color=gender)) +
 geom_violin() +
 labs(title="",x="Perform", y="") +
 geom_boxplot(width=0.1, color="black") +
 scale_color_manual(values=c("#0078D7", "#FE6DB6")) +
 ylim(0.5, 7.5) +
```

```
  ggthemes::theme_clean() +
  theme(axis.text.x = element_blank(),
      legend.position="right")

gender.4 <- gender %>%
  ggplot(aes(x=gender, y=reconsider, color=gender)) +
  geom_violin() +
  labs(title="",x="Reconsider", y="") +
  geom_boxplot(width=0.1, color="black") +
  scale_color_manual(values=c("#0078D7", "#FE6DB6")) +
  ylim(0.5, 7.5) +
  ggthemes::theme_clean() +
  theme(axis.text.x = element_blank(),
      legend.position="right")

gender.5 <- gender %>%
  ggplot(aes(x=gender, y=consistency, color=gender)) +
  geom_violin() +
  labs(title="",x="Consistency", y="") +
  geom_boxplot(width=0.1, color="black") +
  scale_color_manual(values=c("#0078D7", "#FE6DB6")) +
  ylim(0.5, 7.5) +
  ggthemes::theme_clean() +
  theme(axis.text.x = element_blank(),
      legend.position="right")

gender.6 <- gender %>%
  ggplot(aes(x=gender, y=information, color=gender)) +
  geom_violin() +
  labs(title="",x="Information", y="") +
  geom_boxplot(width=0.1, color="black") +
  scale_color_manual(values=c("#0078D7", "#FE6DB6")) +
  ylim(0.5, 7.5) +
  ggthemes::theme_clean() +
  theme(axis.text.x = element_blank(),
      legend.position="right")

gender.7 <- gender %>%
  ggplot(aes(x=gender, y=openness, color=gender)) +
  geom_violin() +
  labs(title="",x="Openness", y="") +
  geom_boxplot(width=0.1, color="black") +
  scale_color_manual(values=c("#0078D7", "#FE6DB6")) +
  ylim(0.5, 7.5) +
  ggthemes::theme_clean() +
  theme(axis.text.x = element_blank(),
      legend.position="right")

gender.8 <- gender %>%
  ggplot(aes(x=gender, y=unbiased, color=gender)) +
```

```
 geom_violin() +
 labs(title="",x="Unbiased", y="") +
 geom_boxplot(width=0.1, color="black") +
 scale_color_manual(values=c("#0078D7", "#FE6DB6")) +
 ylim(0.5, 7.5) +
 ggthemes::theme_clean() +
 theme(axis.text.x = element_blank(),
     legend.position="right")

# merge 8 in 1

png(filename = "subgroup_gender.png",
   unit  = "cm", width = 12, height = 20,
   res = 500)

ggarrange(gender.1, gender.2, gender.3, gender.4,
        gender.5, gender.6, gender.7, gender.8,
        ncol = 2, nrow = 4,
        common.legend = T, legend = "top")

dev.off()
```

Violin plot for older and younger.

```
age <- all[complete.cases(all$age), ]
age$age <- ifelse(age$age > 24.5, "Older", "Younger")
age$age <- factor(age$age,
              levels = c("Older", "Younger"))

age.1 <- age %>%
 ggplot(aes(x=age, y=overall, color=age)) +
 geom_violin() +
 labs(title="", x="Overall", y = "") +
 geom_boxplot(width=0.1, color="black", outlier.size=1) +
 scale_color_manual(values=c("darkgreen", "orange")) +
 ylim(0.5, 7.5) +
 ggthemes::theme_clean() +
 theme(axis.text.x = element_blank(),
     legend.position="right")

age.2 <- age %>%
 ggplot(aes(x=age, y=relatedness, color=age)) +
 geom_violin() +
 labs(title="", x="Relatedness", y = "") +
 geom_boxplot(width=0.1, color="black", outlier.size=1) +
 scale_color_manual(values=c("darkgreen", "orange")) +
 ylim(0.5, 7.5) +
 ggthemes::theme_clean() +
 theme(axis.text.x = element_blank(),
     legend.position="right")
```

```
age.3 <- age %>%
 ggplot(aes(x=age, y=perform, color=age)) +
 geom_violin() +
 labs(title="", x="Perform", y = "") +
 geom_boxplot(width=0.1, color="black", outlier.size=1) +
 scale_color_manual(values=c("darkgreen", "orange")) +
 ylim(0.5, 7.5) +
 ggthemes::theme_clean() +
 theme(axis.text.x = element_blank(),
     legend.position="right")

age.4 <- age %>%
 ggplot(aes(x=age, y=reconsider, color=age)) +
 geom_violin() +
 labs(title="", x="Reconsider", y = "") +
 geom_boxplot(width=0.1, color="black", outlier.size=1) +
 scale_color_manual(values=c("darkgreen", "orange")) +
 ylim(0.5, 7.5) +
 ggthemes::theme_clean() +
 theme(axis.text.x = element_blank(),
     legend.position="right")

age.5 <- age %>%
 ggplot(aes(x=age, y=consistency, color=age)) +
 geom_violin() +
 labs(title="", x="Consistency", y = "") +
 geom_boxplot(width=0.1, color="black", outlier.size=1) +
 scale_color_manual(values=c("darkgreen", "orange")) +
 ylim(0.5, 7.5) +
 ggthemes::theme_clean() +
 theme(axis.text.x = element_blank(),
     legend.position="right")

age.6 <- age %>%
 ggplot(aes(x=age, y=information, color=age)) +
 geom_violin() +
 labs(title="", x="Information", y = "") +
 geom_boxplot(width=0.1, color="black", outlier.size=1) +
 scale_color_manual(values=c("darkgreen", "orange")) +
 ylim(0.5, 7.5) +
 ggthemes::theme_clean() +
 theme(axis.text.x = element_blank(),
     legend.position="right")

age.7 <- age %>%
 ggplot(aes(x=age, y=openness, color=age)) +
 geom_violin() +
 labs(title="", x="Openness", y = "") +
 geom_boxplot(width=0.1, color="black", outlier.size=1) +
```

```
 scale_color_manual(values=c("darkgreen", "orange")) +
 ylim(0.5, 7.5) +
 ggthemes::theme_clean() +
 theme(axis.text.x = element_blank(),
     legend.position="right")

age.8 <- age %>%
 ggplot(aes(x=age, y=unbiased, color=age)) +
 geom_violin() +
 labs(title="", x="Unbiased", y = "") +
 geom_boxplot(width=0.1, color="black", outlier.size=1) +
 scale_color_manual(values=c("darkgreen", "orange")) +
 ylim(0.5, 7.5) +
 ggthemes::theme_clean() +
 theme(axis.text.x = element_blank(),
     legend.position="right")

# merge 8 in 1

png(filename = "subgroup_age.png",
   unit  = "cm", width = 12, height = 20,
   res = 500)

ggarrange(age.1, age.2, age.3, age.4, age.5, age.6, age.7, age.8,
        ncol = 2, nrow = 4,
        common.legend = T, legend = "top")

dev.off()
```

Violin plot for bachelor's and master's.

```
education <- subset(all, all$education=="Bachelor's" | all$education=="Master's")
education$education <- factor(education$education,
                  levels = c("Bachelor's", "Master's"))

education.1 <- education %>%
 ggplot(aes(x=education, y=overall, color=education)) +
 geom_violin() +
 labs(title="", x="Overall", y = "") +
 geom_boxplot(width=0.1, color="black", outlier.size=1) +
 scale_color_manual(values=c("#8931EF", "#666666")) +
 ylim(0.5, 7.5) +
 ggthemes::theme_clean() +
 theme(axis.text.x = element_blank(),
     legend.position="right")

education.2 <- education %>%
 ggplot(aes(x=education, y=relatedness, color=education)) +
 geom_violin() +
 labs(title="", x="Relatedness", y = "") +
```

```
  geom_boxplot(width=0.1, color="black", outlier.size=1) +
  scale_color_manual(values=c("#8931EF", "#666666")) +
  ylim(0.5, 7.5) +
  ggthemes::theme_clean() +
  theme(axis.text.x = element_blank(),
      legend.position="right")

education.3 <- education %>%
  ggplot(aes(x=education, y=perform, color=education)) +
  geom_violin() +
  labs(title="", x="Perform", y = "") +
  geom_boxplot(width=0.1, color="black", outlier.size=1) +
  scale_color_manual(values=c("#8931EF", "#666666")) +
  ylim(0.5, 7.5) +
  ggthemes::theme_clean() +
  theme(axis.text.x = element_blank(),
      legend.position="right")

education.4 <- education %>%
  ggplot(aes(x=education, y=reconsider, color=education)) +
  geom_violin() +
  labs(title="", x="Reconsider", y = "") +
  geom_boxplot(width=0.1, color="black", outlier.size=1) +
  scale_color_manual(values=c("#8931EF", "#666666")) +
  ylim(0.5, 7.5) +
  ggthemes::theme_clean() +
  theme(axis.text.x = element_blank(),
      legend.position="right")

education.5 <- education %>%
  ggplot(aes(x=education, y=consistency, color=education)) +
  geom_violin() +
  labs(title="", x="Consistency", y = "") +
  geom_boxplot(width=0.1, color="black", outlier.size=1) +
  scale_color_manual(values=c("#8931EF", "#666666")) +
  ylim(0.5, 7.5) +
  ggthemes::theme_clean() +
  theme(axis.text.x = element_blank(),
      legend.position="right")

education.6 <- education %>%
  ggplot(aes(x=education, y=information, color=education)) +
  geom_violin() +
  labs(title="", x="Information", y = "") +
  geom_boxplot(width=0.1, color="black", outlier.size=1) +
  scale_color_manual(values=c("#8931EF", "#666666")) +
  ylim(0.5, 7.5) +
  ggthemes::theme_clean() +
  theme(axis.text.x = element_blank(),
      legend.position="right")
```

```
education.7 <- education %>%
  ggplot(aes(x=education, y=openness, color=education)) +
  geom_violin() +
  labs(title="", x="Openness", y = "") +
  geom_boxplot(width=0.1, color="black", outlier.size=1) +
  scale_color_manual(values=c("#8931EF", "#666666")) +
  ylim(0.5, 7.5) +
  ggthemes::theme_clean() +
  theme(axis.text.x = element_blank(),
        legend.position="right")

education.8 <- education %>%
  ggplot(aes(x=education, y=unbiased, color=education)) +
  geom_violin() +
  labs(title="", x="Unbiased", y = "") +
  geom_boxplot(width=0.1, color="black", outlier.size=1) +
  scale_color_manual(values=c("#8931EF", "#666666")) +
  ylim(0.5, 7.5) +
  ggthemes::theme_clean() +
  theme(axis.text.x = element_blank(),
        legend.position="right")

# merge 8 in 1
png(filename = "subgroup_education.png",
    unit  = "cm", width = 12, height = 20,
    res = 500)

ggarrange(education.1, education.2, education.3, education.4,
          education.5, education.6, education.7, education.8,
          ncol = 2, nrow = 4,
          common.legend = T, legend = "top")

dev.off()
```