

This preprint version may not exactly replicate the final version published in the *International Journal of Selection and Assessment*. Copyright: Langer, M., König, C.J., & Papathanasiou, M. (in press). Highly-automated job interviews: Acceptance under the influence of stakes. *International Journal of Selection and Assessment*.

Highly-automated job interviews: Acceptance under the influence of stakes

^a Fachrichtung Psychologie, Universität des Saarlandes, Saarbrücken, Germany

This study was pre-registered on the Open Science Framework (osf.io/hgd5r) and on AsPredicted (<https://AsPredicted.org/i52c6.pdf>). This work was supported by the German Federal Ministry for Education and Research (BMBF; Project EmpaT, grant no.

16SV7229K). We thank the people from Charamel GmbH for their continuous support and for providing us with the virtual character Gloria.

Correspondence concerning this article should be addressed to Markus Langer, Universität des Saarlandes, Arbeits- & Organisationspsychologie, Campus A1 3, 66123 Saarbrücken, Germany. Phone: +49 681 302 4767. E-mail: Markus.Langer@uni-saarland.de

ABSTRACT

Technological advancements allow the automation of every part of job interviews (information acquisition, information analysis, action selection, action implementation) resulting in highly-automated interviews. Efficiency advantages exist, but it is unclear how people react to such interviews (and whether reactions depend on the stakes involved). Participants ($N = 123$) in a 2 (highly-automated, videoconference) \times 2 (high-stakes, low-stakes situation) experiment watched and assessed videos depicting a highly-automated interview for high-stakes (selection) and low-stakes (training) situations or an equivalent videoconference interview. Automated high-stakes interviews led to ambiguity and less perceived controllability. Additionally, highly-automated interviews diminished overall acceptance through lower social presence and fairness. To conclude, people seem to react negatively to highly-automated interviews and acceptance seems to vary based on the stakes.

Keywords: human resource management, highly-automated job interviews, videoconference interviews, applicant reactions, contextual influences.

Introduction

Technology has supported human resource management (HRM) for decades. Through technology, recruitment, personnel selection, and training can be done more easily and at reduced cost (Stone, Lukaszewski, Stone-Romero, & Johnson, 2013). Nowadays, organizations screen a considerable number of applicants through skype interviews or asynchronous digital interviews (Langer, König, & Krause, 2017), employees take part in e-learning sessions (Sitzmann, 2011), and some even train interpersonal skills with a virtual coach (Langer, König, Gebhard, & André, 2016). However, this might merely be the beginning as there is more technology on the horizon to support HRM (Stone, Deadrick, Lukaszewski, & Johnson, 2015).

Indeed, automated interview approaches have the potential to revolutionize job interview practice (Langer et al., 2016; Naim, Tanveer, Gildea, & Hoque, 2018). These automated approaches use sensor devices (i.e., devices capturing human behavior, e.g. nonverbal behavior through cameras), automatic extraction and evaluation of data (e.g., through tools trained to automatically applicant performance), and visualization (e.g., using virtual environments) to automate entire interview processes. Indeed, several of these approaches are already in place from low-stakes situations (e.g., for automated interview training; Langer et al., 2016) up to high-stakes interview-based personnel selection (e.g., automatic evaluation of job interviews; Guchait, Ruetzler, Taylor, & Toldi, 2014).

Researchers have called for studies investigating automated interview approaches because their effects on interviewees remain unknown (Blacksmith, Willford, & Behrend, 2016; Langer et al., 2017). In a first attempt to examine participant reactions to highly-automated interviews, we therefore mimicked the general idea behind these approaches (see Chamorro-Premuzic, Akhtar, Winsborough, & Sherman, 2017; HireVue, 2018; Langer, Schmid Mast, Meyer, Maass, & König, 2019; Precire, 2017). Specifically, we introduced

participants of the current study to a highly-automated interview tool within a virtual set-up (i.e., with a virtual agent as the interviewer). The tool acquires information (through sensors), analyzes information (automatic interview scoring), selects and decides about potential actions (decides about adequate follow-up questions), and implements these actions (controls the reactions of the virtual interviewer) (cf. Parasuraman, Sheridan, & Wickens, 2000). We then examined participants' reactions to such a tool in comparison to a classical technology-mediated interview (i.e., a videoconference interview). Furthermore, we wanted to advance prior research (M. K. Lee, 2018) that indicated that the context matters regarding people's reactions to automated tools. We argue that what is at stakes during a situation (i.e., the relevance of the situation for the future of the person; Peeters & Lievens, 2006) might be an explanation for why highly-automated processes are accepted within one situation compared to another.

Following, we first describe the process of high-level automation for job interviews. Then, we build upon Potosky's (2008) framework and work on decision agents for HRM procedures (Nolan, Carter, & Dalal, 2016; Ötting & Maier, 2018) and virtual agents (K. M. Lee & Nass, 2003; Mori, 1970; Mori, MacDorman, & Kageki, 2012) to develop ideas about potential differences between highly-automated and videoconference interviews. Finally, we explain the possible impact of automation within different levels of stakes.

Background and Hypotheses Development

Automation of HRM processes

Different levels of automation in HRM processes can be differentiated, as described by Parasuraman et al. (2000, p. 287): Their level 1 means that humans take all decisions and actions, level 5 implies that a system executes actions if a human approves these actions, whereas level 10 means that the automated system decides autonomously. More precisely, automated processes help to (a) acquire information, (b) analyze information, (c) select and

decide about potential actions, and (d) implement these actions (Hoff & Bashir, 2015; Parasuraman et al., 2000). For all of these aspects, various technology can help to achieve different degrees of automation (Parasuraman et al., 2000).

In the case of job interviews, acquiring information means to have an automated way to collect and extract data. At a low level of automation this could mean that video cameras and microphones automatically record video and audio data (Langer et al., 2019). On higher levels of automation, it might then be possible to use the recorded data to automatically transcribe the interviews, which saves time compared to taking notes during interviews (Middendorf & Macan, 2002). Furthermore, this kind of automatic data extraction can provide additional data about interviewees such as information about vocal features (e.g., speechrate) and nonverbal behavior (e.g., number of smiles) – information that can influence interview performance evaluations (DeGroot & Kluemper, 2007; Ruben, Hall, & Schmid Mast, 2015).

Automatically analyzing information means to evaluate the acquired information. At lower levels of automation, this might not be different from automatically building scores by averaging the evaluation of different interviewers (Bobko, Roth, & Buster, 2007; see also Nolan et al., 2016). At higher levels of automation, this could mean that machine learning algorithms were trained on past data of successful and unsuccessful applicants to learn what distinguishes them (Chamorro-Premuzic et al., 2017; Naim et al., 2018). This way, algorithms learn to automatically evaluate new interviewees regarding their interview performance and might even provide an overall score for the interviewees that could either serve as a recommendation for hiring managers (cf. Chamorro-Premuzic et al., 2017) or as an additional source of information within interview training (cf. Langer et al., 2016).

Automatically selecting and deciding about potential actions means to use the evaluated information to decide about further steps within the interview or in the following

hiring process. At a basic level of automation and without any input from a prior evaluation phase, implementing structured interviews as digital interviews could already be perceived as automatically selecting and deciding about actions as every applicant would automatically respond to the same set of interview questions automatically and in a standardized manner (see Brenner, Ortner, & Fay, 2016). At a higher level of automation, prior evaluations could be used to choose appropriate follow-up questions, select appropriate feedback for trainees (Gebhard et al., 2018), or to decide which interviewees should proceed to the next selection stage.

Automatically implementing actions means to actually implement the action decisions. At lower levels of automation, this could mean to automatically rank interviewees regarding their performance and presenting this ranking to hiring managers (Campion, Campion, & Reider, 2016). At higher levels of automation this might mean to automatically choose appropriate follow-up questions from a set of questions for a given participant or to automatically screen out certain applicants (cf. Faliagka, Tsakalidis, & Tzimas, 2012).

An additional aspect of highly-automated interviews that goes beyond the model of automation by Parasuraman et al. (2000) is the actual set-up of the interview. For instance, highly-automated interviews can be implemented as telephone interviews (Precire, 2017), digital interviews (Brenner et al., 2016) or virtual interviews (i.e., in a virtual environment with a virtual agent as interviewer; Langer, König, & Fitali, 2018).

In sum, decisions about the level of automation affect every aforementioned aspect and therefore also the design of the interview tool. These design choices consequently may alter how people react to such highly-automated interviews. For the current study, we decided to introduce participants to a highly-automated interview tool (i.e., that includes high levels of automation on all aforementioned aspects). Specifically, the highly-automated tool collects

audio and video information about the trainee/applicant and extracts vocal, nonverbal, and verbal information, analyzes this information (e.g., recognizes that the participant is nervous; analyzes the quality of the interview responses), decides about appropriate actions (e.g., decides how to react if an interviewee is nervous), and implements these actions (e.g., asks adequate follow-up questions or calms the interviewee) within a virtual environment and with a virtual agent as interviewer. In practice, similar tools are already used. For instance, similar virtual training tools are already used to support human trainers in negotiation training (Langer et al., 2019), and in the field of personnel selection, Schmid Mast and colleagues (2017) used highly-automated interviews to predict the job performance of student assistants, similar to the ones already offered by providers of automated interview solutions (e.g., HireVue, 2018; Precire, 2017).

Acceptance of Technology in HRM

Until now, research is lacking on how people react when they are confronted with highly-automated interviews. However, asserting positive reactions are often crucial for the effectiveness of such approaches (Stone et al., 2013). Since acceptance includes a positive overall evaluation of a procedure (Davis, Bagozzi, & Warshaw, 1989), within this study we define overall acceptance as a positive evaluation of an automated procedure within a given context. Following Langer and colleagues (2017), we argue that to raise overall acceptance of automated interviews the following variables need to be evaluated: (a) the social quality of the entire procedure (e.g., interpersonal treatment; Gilliland, 1993), (b) participants' perceptions of behavioral control regarding the influence on a given situation and the confidence of being able to execute a specific behavior during the situation (Ajzen, 2002b), (c) affective reactions (Langer et al., 2018; Tene & Polonetsky, 2015), (d) uprising privacy concerns during the situation (Bauer et al., 2006), and (e) general fairness perceptions (Colquitt & Zipay, 2015).

Factors Differentiating Videoconference and Highly-Automated Interviews

This study compares videoconference to highly-automated interviews, using a highly-automated interview tool that can be used for various contexts. Comparing videoconference and highly-automated interview approaches regarding acceptance leads to a variety of features that might differ between those two interview formats. Potosky's (2008) framework of media attributes provides a theoretical background to generate ideas about potential differences between the interview formats (see also Langer et al., 2017 who used Potosky's framework to compare videoconference interviews and asynchronous digital interviews). In the next paragraphs, we will briefly introduce Potosky's framework that distinguish between four attributes (i.e., social bandwidth, interactivity, transparency, and surveillance).

Social bandwidth relates to the possibility of exchanging communication information (Potosky, 2008). Specifically, media with a large social bandwidth allow exchanging a variety of social signals. In an interview with another human through videoconference, social bandwidth should be relatively high as humans send and receive many different verbal and nonverbal information. In highly-automated interviews, participants and the virtual character can also send and receive social signals, but social bandwidth should be relatively lower because automated technologies are still not as good as humans in recognizing and sending communicational content. For instance, cameras and associated social signal analysis software are still not free of errors in recognizing social signals (An, Yang, & Bhanu, 2015; Guo, Polanía, & Barner, 2018) and smiles by virtual characters might still be not as realistic as smiles by human beings (cf., Kätsyri, Förger, Mäkäraänen, & Takala, 2015).

The second aspect of Potosky's (2008) framework is interactivity which focuses on the opportunity to interact with a communication partner. Highly interactive media allow direct responses to communication partners. Thus, highly-automated interviews offer lower interactivity than videoconference interviews because participants interact with a virtual

communication partner and because the adaption to interviewees is also automatized (cf., Gebhard et al., 2018). Even the best highly-automated tools nowadays will not offer the same interactivity (e.g., possibility to ask open-ended questions; Frauendorfer, Schmid Mast, Nguyen, & Gatica-Perez, 2014) as videoconference interviews.

Transparency of a medium is high if participants do not realize that they are communicating through a medium and if there are no obstacles during the communication (Potosky, 2008). For instance, transparency would be low in cases in which there are video or audio interferences during videoconference interviews (Potosky, 2008). However, if these interviews occur without such interferences, people likely forget that they are communicating through technology, making videoconference interviews more transparent than highly-automated interviews (Langer et al., 2017). Furthermore, communicating through videoconference might be much more familiar than interacting with a highly-automated interview tool. This might also reduce transparency in highly-automated interviews because participants do not really understand what is happening during such an interview (Potosky, 2008). For instance, naïve participants will likely have no idea about the capabilities of a specific highly-automated tool (Langer et al., 2018). This lack of understanding might make it hard for humans to express themselves in a natural way, thus decreasing transparency.

Surveillance, the last aspect of Potosky's (2008) framework, relates to the extent to which an interaction appears to be observable by a third party. High surveillance could mean that the parties are aware that an interaction is recorded and later accessed by others (cf., Langer et al., 2017). Since many people use videoconference technologies, it might occur that they are concerned about their conversation being monitored. In the case of an highly-automated interview, however, it is much less clear if the video recording is just recorded, if there are other people watching the video in a live stream or if the video recordings are later watched by unauthorized people (see also Langer et al., 2017). Furthermore, even if the

objective surveillance in both versions of the interview is equal (e.g., recording of the interview), people might perceive more surveillance in the case of highly-automated interviews (cf. Barry & Fulmer, 2004) because it might be less certain what happens with one's data during such interviews (which, again, is also an issue of low transparency). Thus, perceived surveillance of a highly-automated interaction could be more severe (cf. McCole, Ramsey, & Williams, 2010).

In addition to the differences between highly-automated and videoconference interviews that can be deduced from Potosky's (2008) model, there are two aspects that differ between highly-automated and videoconference interviews: the decision agent and the set-up of the interview. Decisions in modern HRM procedures might not necessarily be made by humans (Ötting & Maier, 2018). In the case of highly-automated interviews, the interview tool might independently decide how to react to a given interviewee. Furthermore, the tool evaluates the interview responses and recommends only the top candidates for the next selection stages. In both cases, a highly-automated tool is the decision agent. In videoconference interviews, however, the decision agent is a human interviewer. Prior research has investigated the use of automatic-decision support systems and investigated the effects on people's reactions. For instance, Arkes, Shaffer, and Medow (in a medical setting; 2007) as well as Nolan et al. (in a personnel selection setting; 2016) showed that decision-support systems affect people's reactions to a given decision process. Both studies suggest that people get less credit for their influence on the decision-making process as soon as there are decision-support systems involved, which also implies that people believe that human influence on decisions reduces with automatic decision-support. In a similar vein, M. K. Lee (2018) found that people react negatively to automatic decisions in personnel selection situations compared to automatic decisions in work scheduling situations and argued that this difference is due to the reduction of human influence on the decision. On the positive side,

reducing the human influence on decisions might also be perceived to reduce possible biases and thus enhance consistency in the decision-making process (which people seem to tend to believe when thinking about automatic decisions; F. A. Miller, Katz, & Gans, 2018).

The last feature that differs between the interview formats is the interview set-up. We decided to implement the highly-automated interview within a virtual set-up with a virtual interviewer because this seems to be the closest automated-adaptation of a videoconference interview (cf. Langer et al., 2018). As the previous paragraphs already imply, the use of a virtual agent as interviewer might affect participants' reactions. At the positive side, virtual characters could positively affect perceived social bandwidth and interactivity during a situation (K. M. Lee & Nass, 2003) compared to a situation where there is only a recording screen (e.g., in a digital interview). However, compared to videoconference interviews, social presence and interpersonal treatment should still be relatively lower in highly-automated interviews, as a human interviewer should enhance social bandwidth and allow more interactivity. Furthermore, another potentially negative influence of using virtual characters comes from the uncanny valley hypothesis (Mori, 1970; Mori et al., 2012). This hypothesis suggests that acceptance of virtual character rises as they become more human-like up to a certain point where acceptance rapidly drops – this drop in acceptance is associated with negative feelings that constitute the uncanny valley (Langer & König, 2018). Consequently, having a virtual interviewer might lead to negative feelings elicited by an uncanny valley effect.

In order to measure the impact of differences between highly-automated interviews with virtual interviewers and videoconference interviews, the current study follows the example of Langer et al. (2017) and examines a variety of different acceptance variables relating to procedural justice research by Gilliland (1993). Specifically, we examine social presence (the feeling of present interpersonal warmth and empathy during an interaction;

Walter, Ortbach, & Niehaves, 2015) interpersonal treatment (feelings of being treated with respect and dignity; Bauer et al., 2001), perceived behavioral control (the believe that one can influence a specific situation [perceived controllability], the confidence of being able to execute a specific behavior and to control a situation [perceived self-efficacy]; Ajzen, 2002b), consistency (in the sense of equal treatment; Bauer et al., 2001), and fairness. Furthermore we examine acceptance variables stemming from research on novel technologies and virtual characters: privacy concerns (concerns about what happens to ones' private data; Malhotra, Kim, & Agarwal, 2004), and creepiness (an ambiguous negative affective reaction towards a specific situation, paired with feelings of not knowing how to judge and handle this situation; Langer & König, 2018). Higher social bandwidth and interactivity in videoconference interviews should condense in the acceptance variables social presence, interpersonal treatment, perceived behavioral control, and consistency. Higher transparency within videoconference interviews should be reflected by perceived behavioral control and creepiness. Surveillance should relate to privacy concerns. The differing decision agency could affect social presence, behavioral control, and consistency. A virtual character as interviewer could lead to less social presence and interpersonal treatment and could also enhance creepiness. Additionally, Potosky's attributes and the fact that there are different decision agents and a virtual interviewer could affect general fairness evaluations. All things considered, the aforementioned arguments predominantly assume favorable effects for videoconference interviews except for consistency, thus we state:

Hypothesis 1: Videoconference interviews will be evaluated higher on social presence, perceived interpersonal treatment, perceived behavioral control, fairness, privacy

concerns and lower on creepiness as well as consistency compared to highly-automated interviews.¹

In the current study, acceptance is the main dependent variable as it is of crucial importance for the effectiveness and ongoing use of various HRM procedures (e.g., personnel selection or training; Highhouse, Lievens, & Sinar, 2003; S. M. Lee, Kim, & Lee, 1995). To evaluate overall acceptance in different contexts, we draw on a concept from user experience studies, where overall acceptance is measured through perceived attractiveness of a procedure reflecting a simple “good versus bad” assessment of the respective procedure (Laugwitz, Held, & Schrepp, 2008).

Other than consistency, all of the aforementioned effects seem to be advantageous for videoconference interviews. As such, we expect the aforementioned acceptance variables to mediate a positive effect of videoconference interviews on the general attractiveness of the procedure (i.e., how positively participants perceive the procedure).

Hypothesis 2. The positive effect of videoconference compared to highly-automated interviews on the attractiveness of the procedure will be mediated by social presence, interpersonal treatment, perceived behavioral control, fairness, creepiness, and privacy concerns.

How the Stakes might Affect Acceptance

The highly-automated tool that we introduce can be used for various purposes, for high-stake personnel selection contexts and for low-stakes applications such as interview or

¹ To support the call for open science, this study was pre-registered (i.e., before we gathered any data, we clearly stated our hypotheses and methods online; Open Science Collaboration, 2015). In the pre-registration, we included efficiency as an additional dependent variable and an exploratory moderated-mediation model. Our measure for efficiency was not adequately reliable and was thus excluded. The results of the moderated-mediation did not change the conclusions of this study qualitatively, so this analysis was excluded for the paper.

negotiation training (Langer et al., 2019). High-stake situations can be defined as situations where people are under relatively high pressure because the situation has potentially far-reaching effects for the future life of a person, whereas low-stakes contexts induce less stress, can be completed more casually, and potentially require less preparation than high-stake contexts where people usually want to put their best foot forward (Jansen, König, Stadelmann, & Kleinmann, 2012).

Applying highly-automated tools within different contexts could affect people's reactions to these tools. M. K. Lee (2018) found that people react more favorably to highly-automated decisions for mechanical tasks (e.g., work scheduling) than for human tasks (e.g., hiring), which might be explained by what was at stakes. Decisions about work scheduling might affect if a person can go to the cinema or meet with friends. In contrast, hiring decisions might impact the future of the applicant (McCarthy & Goffin, 2004). Favorable hiring decisions possibly lead to a higher salary, moving to another city, integrating in a new team whereas unfavorable hiring decisions might lead to uncertainty about the next steps and to pressure seeking other options (Ryan & Ployhart, 2000).

We propose that the stakes of a context might affect reactions to highly-automated tools because people might perceive that a human decision agent should be in charge of high-stake processes and decisions instead of leaving those to highly-automated tools (Binns et al., 2018). Additionally, setting up a high-stake context in a virtual environment might also lead to negative feelings about the respective tool as people conclude that this kind of situation should rather be handled through an interaction with another human-being maybe because of the implicit sign that there is more social recognition when another person handles an important decision (M. K. Lee, 2018). To examine whether stakes might influence people's reactions to highly-automated tools, we manipulated the stakes through introducing participants to either a personnel selection or a training situation. Personnel selection contexts

should be high-stake situations as they are important for ones' future and often induce stress (McCarthy & Goffin, 2004). In fact, applicants invest a lot of effort preparing job interviews (Maurer, Solamon, Andrews, & Troxtel, 2001), are afraid before the job interview (McCarthy & Goffin, 2004), and even lie to get a job (Buehl, Melchers, Macan, & Kühnel, 2018). Although training contexts can also be high-stakes situations, they might be less stressful and potentially life-changing than personnel selection contexts (McCarthy & Goffin, 2004).

Considering the aforementioned arguments, it is arguable that there will be an interaction between the context and the interview type as automated high-stake interviews might evoke more negative reactions than automated low-stake interviews. However, these assumptions are preliminary which is why we ask:

Research Question 1: Are highly-automated interviews less accepted within high-stake situations?

Method

Sample

To generate ideas about the possible effect size between the experimental conditions, we consulted the meta-analysis of Blacksmith and colleagues (2016) on the effect of technology on applicant reactions. They found small to medium effect sizes for applicant reactions in favor of less automated interview methods (i.e., face-to-face interviews). Accordingly, we expected to find a moderate effect in the context of the current study. Sample size calculation with G*Power (Faul, Erdfelder, Buchner, & Lang, 2009) revealed that under the assumption of a moderate effect of Wilk's $\lambda = 0.92$, a sample of $N = 124$ was necessary for a power of $1 - \beta = 0.80$. Because of formerly experienced problems with online experiments (e.g., technical problems and long pauses taken by participants during the experiment), we oversampled until our sample consisted of $N = 132$ participants. We excluded three participants who stated we should not use their data as they did not execute

the experiment seriously. Furthermore, we excluded two participants because of technical problems, one participant because he/she skipped one part of the manipulation, and four other participants because they interrupted the experiment for more than one hour. The final sample consisted of $N = 123$ German participants (57% female), 76% of which were students (53% of them studied psychology). The mean age was 25.12 years ($SD = 7.44$). Most participants (46%) had experienced one to three job interviews before, 29% had experienced four to five job interviews, 9% had experienced six to nine interviews, 10% of participants had experienced more than ten interviews and the rest had not experienced any job interviews before. Some participants (13%) had experienced at least one technology-mediated job interview before. Participants were rewarded with course credit and the possibility of winning a small price.

Design, Procedure, and Manipulation

The entire study was conducted via an online survey platform. Participants were randomly assigned to one of the groups of the 2×2 between subject design (videoconference vs. highly-automated interview; low-stakes vs. high-stakes context). Participants first read the description of a situation where they were asked to imagine that a friend was invited to a job interview. This was done to introduce participants to an observer role and a word-of-mouth situation where participants watch a video of the situation their friend had experienced.

In the low-stake groups, participants received information describing that their friend wants to prepare for the job interview and finds a human trainer or a virtual training tool respectively. They were informed that the trainer/training tool asks interview questions and provides feedback for nonverbal behavior and speech. Furthermore, they were told that the trainer/training tool would provide feedback to the performance after the training session.

In the high-stake groups, participants received information indicating that their friend's interview will be conducted by a human interviewer or a virtual interview tool,

respectively. They were informed that the human interviewer/the interview tool asks interview questions and pays attention to interviewees' nonverbal behavior and speech. In addition, they were told that after the job interview, the human interviewer/the interview tool would decide if an applicant will be considered for a follow-up face-to-face interview.

Descriptions were equal in length and information, except for the experimental manipulation. Afterwards, participants watched videos respective to the videoconference and highly-automated conditions. After the videos, the low-stake groups were informed that their friend waits for feedback regarding interview performance and that the trainer/training tool will provide feedback for nonverbal behavior, speech, and content to improve future interview performance. The high-stake groups received the information that the human interviewer/the interview tool will analyze nonverbal behavior, speech, and content. Furthermore, they were told that the human interviewer/the interview tool will decide if the interviewee will advance to the next selection stage. We did not provide participants with any information about the outcomes of the conditions because feedback on the outcomes would have added another potential source of variance to the results (e.g., variance caused by the fact that people receive performance feedback by a human vs. a virtual character, see Hausknecht, Day, & Thomas, 2004, for the effects of outcomes on applicant reactions). In the end, participants answered to a questionnaire containing all dependent measures.

The Videos

First, we had to determine how to design the video for the highly-automated groups. In particular, the tool should include all aspects of highly-automated interviews: acquire information, analyze information, select and decide about potential actions, and implement these actions (Hoff & Bashir, 2015; Parasuraman et al., 2000). Furthermore, we followed the general idea underlying highly-automated approaches as we decided to highlight the importance of verbal, paraverbal, and nonverbal behavior information (see Naim et al., 2018;

Schmid Mast et al., 2017). We used the approach of Langer and colleagues (2016) who used a highly-automated interview tool to create the video for the highly-automated interview conditions. This tool acquires information through a depth camera (e.g., Intel RealSense®; Intel, 2018) and a microphone. For the analysis of information, the Social Signal Interpretation Framework (SSI; Wagner et al., 2013) uses input from the camera and microphone and integrates plugins such as OpenFace (Baltrušaitis, Robinson, & Morency, 2016) and PRAAT (Boersma & Van Heuven, 2001) to extract and interpret the raw data from the camera and microphone. For instance, OpenFace includes algorithms trained to recognize nodding or smiling as well as facial expressions (Baltrušaitis et al., 2016). Using the information from the SSI, the Visual Scene Maker (VSM; an authoring tool for virtual environments; Gebhard, Mehlmann, & Kipp, 2011) can then be used to further interpret interviewees' nonverbal behavior (i.e., select and decide about potential actions) and to manipulate the virtual environment and the virtual character (i.e., implement these actions). For example, if the SSI recognizes an interviewee's smile, it can send that data to the VSM. In the VSM, designers can integrate contextually different ways how to handle a smile (see Gebhard et al., 2018 for a detailed description of how the highly-automated interview tool works). For instance, the virtual character reacts with a smile to a user smile, but only in the beginning of the interview, whereas later in the interview, the virtual character may not smile back as adequacy of smiling differs for the stages of the interview (Ruben et al., 2015). Finally, the interview was set-up in a virtual environment (provided by the company Charamel) with a female virtual character as the interviewer. This environment shows an office and the virtual character behind a desk. The behavior of the interview character is controlled using the VSM and the input of the SSI.

Second, the study design required participants became aware of the capabilities of the highly-automated interview tool. We therefore used a similar video like Langer et al. (2018)

who introduced their participants to a highly-automated interview. The video of the videoconference/highly-automated interview showed a female human/virtual interviewer interacting with the supposed (female) friend of the participants; only the interviewer was visible. We tried to closely match the interviewers (both interviewers were female, had long black hair, blue eyes, and wore a professional outfit). To make sure that the observing participants were aware of the automated parts of the interview (i.e., acquiring and analysis of information, decision about actions, action implementation; the interviewer recognizes and adapts to interviewees' nonverbal behavior and emotions), the interviewee in the video becomes nervous after being asked the second question and hesitates to answer the question. As a result, the interviewer says that it was noticed that the interviewee seems to be nervous. The interviewer emphasizes that being nervous is completely comprehensible and try to calm the interviewee by acting very friendly. Afterwards, the interviewee recovers from her nervousness and answers to the question; then the video fades out.

The videos were similar in length; one identical video was used for both highly-automated contexts and one identical video for both videoconference interview contexts. In addition, the same audio track of the answers provided by the supposed friend of the participants was used for both videos.

Measures

Social presence, perceived behavior control, creepiness, and privacy concerns items ranged from 1 (*strongly disagree*) to 7 (*strongly agree*). For consistency, interpersonal treatment, and fairness items ranged from 1 (*strongly disagree*) to 5 (*strongly agree*). All items were adapted to participants' observer role.

Social presence was measured with five items adopted from Walter and colleagues (2015). A sample item is "The interviewer acted empathically."

Perceived behavioral control was measured with eight items taken from Langer and colleagues (2017) who followed the suggestions by Ajzen (2002) that perceived behavioral control scales should consist of perceived self-efficacy and perceived controllability items. A sample item for perceived self-efficacy is “I am sure that I could control the shown procedure through my behavior.” A sample item for perceived controllability is “It is possible to manage such a procedure.”

Consistency and interpersonal treatment were measured with three and four items from a German version of the Selection Procedural Justice Scale (Bauer et al., 2001; Warszta, 2012). A sample item for consistency is “This procedure is administered to all applicants in the same way.” A sample item for interpersonal treatment is “During the interview, the participant was treated politely.”

Fairness was measured with two items from Warszta (2012). A sample item is “I think the shown procedure is fair.”

Creepiness was measured with ten items from Langer and König (2018). The creepiness scale consists of two facets, namely emotional creepiness and creepy ambiguity, both measured with five items. A sample item for emotional creepiness is “During the shown situation, I had a queasy feeling.” A sample item for creepy ambiguity is “I did not know how to judge this situation.”

Privacy concerns were measured with six items. One item was taken from Smith, Milberg, and Burke (1996), two items from Malhotra, Kim, and Agarwal (2004), two items from Langer and colleagues (2018), and one item from Langer and colleagues (2017). A sample item is “Situations like the one shown threaten participants’ privacy.”

Attractiveness of the procedure was measured with six items from the User Experience Questionnaire (UEQ; Laugwitz, Held, & Schrepp, 2008). Items of the UEQ are pairs of opposites where people answer on a 7-point scale from -3 to +3 between the pairs of

opposites. For instance, if a participant answered +3 for the sample item combination “bad-good” the procedure is evaluated as very good.

Manipulation check measures. Two manipulation check measures were provided that ranged from 1 (*strongly disagree*) to 7 (*strongly agree*). To examine if people perceived selection situations as more high-stake than training situations, we measured high-stake evaluations with six items. These items were developed by the authors. A sample item is “Such situations are crucial for participants’ future.” To cover the interaction type manipulation the item provided was “The interviewer in the video was a human being.”

Results

Manipulation Checks

Participants were more likely to perceive the personnel selection context as high-stake $F(1, 119) = 10.36, p < .01, \eta^2_p = .08$. Furthermore, participants in the videoconference interview conditions were more likely to perceive the person shown as a human than in the highly-automated interviews, $t(80.55) = 22.30, p < .01, d = 4.03$.

Testing the Hypotheses

Table 1 provides an overview of descriptive statistics and intercorrelations. We used MANOVA and single ANOVAs to assess the differences between the interview types. Overall, there was a difference between videoconference and highly-automated interviews, $F(11, 109) = 3.06, p < .01, \text{Wilk's } \lambda = .76$. Hypothesis 1 examined the difference between videoconference and highly-automated interviews for social presence, interpersonal treatment, perceived behavioral control, consistency, fairness, creepiness, and privacy concerns. Results of the single ANOVAs are shown in Table 2 and Figure 1. Opposing Hypothesis 1, there were no differences between the interview approaches for social presence, interpersonal treatment, consistency, and fairness. However, the findings support Hypothesis 1 regarding the effect of the interview approach on perceived behavioral control

(moderately lower for highly-automated interactions), and privacy concerns (slightly to moderately higher for highly-automated interactions). Additionally, we found that highly-automated interviews evoked moderately more creepy ambiguity. Since Hypothesis 1 investigates the outcomes for seven variables which might lead to an accumulation of the alpha-error, we additionally examined the results using a Bonferroni-Holm corrected alpha threshold (Holm, 1979). This means we divided the alpha significance threshold through the amount of follow-up ANOVAs for the Hypothesis (i.e. $0.05/7 = 0.007$) and compared the significance values of the ANOVAs starting with the “most significant” result (i.e., for the effect of perceived controllability between highly-automated and videoconference interviews). For the next comparisons, the alpha threshold is divided through the amount of follow-up ANOVAs minus one and compared to the next significance value resulting from the ANOVAs (i.e., perceived self-efficacy). When comparing the significance values of the variables to the Bonferroni-Holm corrected alpha threshold, perceived controllability, self-efficacy, and creepy ambiguity showed a significant difference between the groups, whereas privacy concerns was not significant any more ($p = .032$ compared to a corrected alpha threshold of .028).

--- Insert Table 1, Table 2 and Figure 1 about here ---

In Hypothesis 2, we wanted to test the mediating effects of the dependent variables between the interview approach and the attractiveness of the procedure. Mediation significance tests were conducted with PROCESS (Hayes, 2013). Table 3 and Table 4 present the results. The findings indicate that overall, there was a positive indirect effect between the interview approach and the attractiveness of the procedure mediated by social presence and

fairness. This partially supports Hypothesis 2 as the other dependent variables were not significant mediators. The resulting model is presented in Figure 2.²

--- Insert Table 3, Table 4, and Figure 2 about here ---

Regarding Research Question 1, Figure 1h) indicates that there was a small to medium interaction effect for creepy ambiguity and Figures 1c) and 1f) imply that there were interaction effects for perceived controllability (small to medium effect) and fairness (moderate effect). When using a Bonferroni-Holm correction (i.e., $0.05/10 = 0.005$; as we had ten outcome variables) none of the interactions remained significant (fairness $p = .0067$ compared with 0.005, controllability $p = .023$ compared with 0.0055, creepy ambiguity $p = .028$ compared with 0.006).³

² In the pre-registration, we proposed to include information known and computer self-efficacy as covariates.

These analyses were conducted, however they did not change the results qualitatively so they were not included in the paper (they can be made available on request).

³ Participants in the current study did not experience a real interview so the results should be interpreted cautiously. Yet, as we were especially interested in the results regarding the high-stake interview, we conducted a follow-up study. In an attempt to check if the results for the high-stake interview also hold within a real interaction with the interview approaches, we conducted an experiment with $N = 69$ female participants (only females in order to prevent introducing an additional independent variable “interviewee gender”). In a 2×2 design (male vs. female interviewer; videoconference vs. highly-automated interview), participants experienced a structured interview where they responded to five interview questions. Participants came to the laboratory and either interacted with one of the virtual characters or with one of the human interviewers via videoconference. All participants received negative feedback regarding their interview performance and then responded to the same items like in the current study (adapted to the live interaction). Results for the independent variable “interview approach” were similar to the results in the current study as participants perceived the highly-automated interview as less fair ($F[1, 65] = 7.90, p < .01, \eta^2p = .10$) and they perceived less behavioral control ($F[1, 65] = 7.16, p < .01, \eta^2p = .10$). In contrast to the current study, there was no significant effect for privacy concerns ($F[1, 65] = 1.27, p = .26, \eta^2p = .02$), but we found that participants perceived less social presence

Discussion

This study compared the acceptance of a highly-automated interview to a videoconference interview in different contexts. Results show that the high level of automation impaired acceptance of the interview. Furthermore, our findings suggest that people seem to be sensitive to the context in which automated tools are used. The participants were more critical about the highly-automated interview in a high-stake context (personnel selection) than about the same tool for a low-stake context (training) (but note that after correcting for potential alpha-error accumulation the interaction effects, albeit in a range of small to medium effect sizes, were not significant anymore and should thus be interpreted cautiously). Following, we will discuss the results for the different acceptance variables in more detail.

First, the findings indicate that participants were more skeptical of people's ability to control a situation in which they are confronted with a highly-automated interview in comparison to a videoconference interview. This perceived lack of control seems to be more pronounced for automated high-stake interviews. Supporting our claim that the highly-automated interview may be perceived as offering less social bandwidth, interactivity, and transparency as defined by Potosky (2008), and because there is an automated decision (cf. Ötting & Maier, 2018), these results could indicate that participants were concerned about

($F[1, 65] = 15.58, p < .01, \eta^2p = .19$), lower interpersonal treatment ($F[1, 65] = 6.50, p < .05, \eta^2p = .09$) and higher consistency ($F[1, 65] = 10.67, p = .01, \eta^2p = .14$) within the highly-automated interview. The gender of the interviewer and the interaction between interviewer gender and interview approach did not affect the conclusions based on the main effect for the interview approach. These results support the overall conclusion based on the findings of the current study but also indicate that certain hypotheses of the current study might only hold for real interactions. The follow-up study was also pre-registered on AsPredicted (<http://aspredicted.org/blind.php?x=b4cn3y>) and its detailed results can be made available upon request.

reduced influence on the interview and its outcomes. One especially impactful influence on interviews is the use of impression management (see also Barrick, Shaffer, & DeGrassi, 2009; Blacksmith et al., 2016). This could explain the especially low perceived control in automated high-stake interviews as impression management gains importance in high-stake situations (cf. Holden & Passey, 2010; Levashina & Campion, 2006). In the case of a videoconference interview, applicants gain some control over the situation by knowing how to influence the interviewer (e.g., by ingratiation, smiling, nodding etc.; cf., Barrick et al., 2009) and the decision (as impression management impacts interviewer decisions; cf., Peck & Levashina, 2017). In an automatic decision scenario, feelings of “knowing how to influence” are likely reduced as interviewees might have no insight into which variables are used to determine if they are invited to the next selection stage and because people tend to believe that humans have less influence on decisions in such scenarios (Arkes et al., 2007; Nolan et al., 2016). This result is especially interesting in comparison to the results of Langer et al. (2017) who found no difference in controllability when comparing digital interviews to videoconference interviews. In their study, participants knew that a human interviewer will evaluate their interview answers, so it might be possible that the lower controllability in the current study was due to the fact that there was a highly-automated decision. However, we caution that the study by Langer et al. (2017) was different from our study in many other ways, which could also have increased their participants’ perceived controllability.

Furthermore, participants assessed the automated high-stake interview as being more ambiguous than videoconference interviews or automated low-stakes interviews. On the one hand, this may indicate that Potosky’s (2008) aspect of transparency appears to be relatively low for automated high-stake interviews. On the other hand, this is in line with former research that suggested that feelings of uncertainty and ambiguity are especially prevalent in novel situations where people do not know what to do, what to feel, or how to judge the

situations (Langer & König, 2018; Tene & Polonetsky, 2015). This might also account for the result that the low-stake interview was perceived as being less ambiguous, as it might be more familiar to use automation for low-stake scenarios such as training or work scheduling (see also M. K. Lee, 2018; Zyda, 2005). However, we were surprised by the fact that there was no difference in emotional creepiness between the groups, not even for automated high-stake interviews. One reason for this might be that people only observed the situation; thus they might not have experienced enough emotional immersion into the situation. This finding might also indicate that perceptions of the virtual character in the current study were not affected by potential negative emotional effects caused by the uncanny valley (Kätsyri et al., 2015; Langer & König, 2018; Mori, 1970; Mori et al., 2012).

One of the most important outcomes of the current study is that participants evaluated automated high-stake interviews as being particularly unfair. A possible explanation stems from the fact that we decided to describe a highly-automated interview process. Specifically, the interview tool we describe gathers data from interviewees, evaluates this data, and provides feedback for interviewees. It seems possible that participants might have interpreted that a computer gains decision power over a human being, which may be seen as less fair than a person being in charge of conducting high-stake interviews and making high-stake selection decisions (Binns et al., 2018; M. K. Lee, 2018; Ötting & Maier, 2018). To be clear, the issue at hand might be the lack of transparency of automated high-stake situations in addition to a highly-automated decision agent (Mittelstadt, Allo, Taddeo, Wachter, & Floridi, 2016). In the case of an interpersonal high-stake situation, people might be better able to understand the process and the decisions, whereas this is much harder when interacting with highly-automated tools and when being confronted with highly-automated decisions. Similar to reduced fairness evaluations, it seems that lower transparency was at least partly accountable for reduced perceived behavioral control and increased ambiguity, indicating that

differences in this aspect of Potosky's framework impacted the acceptance variables in the current study. These are important findings for the research area of explainable artificial intelligence (Biran & Cotton, 2017) as they highlight the necessity of transparent, comprehensible and ethically sound highly-automated procedures.

The findings regarding privacy concerns may imply that highly-automated interviews lead to more concerns about what is happening with ones' private data but should be interpreted very cautiously as the effect for privacy concerns was relatively low. These results tentatively support the hypothesized differences in surveillance as defined by Potosky (2008) and the findings by Langer et al. (2017) who also found that a more automated version of an interview led to higher privacy concerns (i.e., digital interviews). Potentially, higher privacy concerns could be elicited by higher uncertainty of who will access the data in interviews without any human influence (McCole et al., 2010).

We measured overall acceptance of the procedure as overall evaluation of the interview procedures, and mediation results for this outcome raise further concern about the acceptance of highly-automated interviews. Results indicate that the overall evaluation of the highly-automated interviews suffered from lower fairness perceptions and less perceived social presence. This supports the suggested differences for Potosky's (2008) aspects of social bandwidth and interactivity and that these differences detrimentally affect the overall evaluation of the highly-automated tool. Additionally, these results add to former personnel selection research showing that technology-mediated interviews (i.e., videoconference interviews) may be less accepted than face-to-face interviews (Blacksmith et al., 2016; Chapman, Uggerslev, & Webster, 2003; Sears, Zhang, Wiesner, Hackett, & Yuan, 2013) but they could at least be more accepted than completely automated interviews. In other words, Blacksmith et al. (2016) concluded that videoconference interviews are less accepted than face-to-face interviews because they impact the interpersonal communication. Our results

imply that highly-automated interviews are even less accepted than videoconference interviews possibly because they reduce the human influence on the interview even further (i.e., highly-automated interviews eliminate interpersonal communication and reduce human influence on the process and the decision; see also Langer et al., 2017).

Comparing the results of the ANOVAs (effects for perceived behavioral control, creepiness and privacy concerns) with the mediation results (indirect effects for fairness and social presence) seem to produce a diffuse picture as significant variables in the one analysis were not significant in the other. It is therefore important to note that within the ANOVA, fairness ($p = .051$) and social presence ($p = .086$) were approaching conventional significance threshold and showed small to medium effect sizes, whereas in the mediation analysis they significantly mediated the indirect effect. The reverse was true for perceived behavioral control, creepy ambiguity, and privacy concerns, which showed stronger effects and significant results within the ANOVA but were not significant in the mediation analysis (see confidence intervals in Table 4). Supporting various previous research (Bauer et al., 2006; Blacksmith et al., 2016; Tene & Polonetsky, 2015), this means that perceived behavioral control, creepiness, and probably also privacy concerns are important to distinguish highly-automated interviews from videoconference interviews. However, when it comes to consequences regarding overall acceptance of the procedure, social presence and fairness seem to be more important factors.

Considering the results for social presence, it is somewhat surprising that there were negligible effects of the interview approach for the conceptually related variable interpersonal treatment. A reason for this could be that the dialogue within the video was very friendly as the interviewer showed interpersonal connectedness, and reacted to nervousness shown by the applicant. This is one of the few promising results for highly-automated interviews as this could mean that even if such situations are perceived as less social present, carefully designed

highly-automated interactions (and behavior of virtual agents) can induce feelings of human-like interpersonal treatment (see also Gratch et al., 2007; K. M. Lee & Nass, 2003).

It was also surprising that we did not find any significant difference for consistency. A reason for this could be that people in the highly-automated interview might have thought that the computer character can autonomously ask questions thus reducing comparability between applicants, this would then again point towards the fact that people ascribe human features to computers (Gratch et al., 2007; Nass, Steuer, & Tauber, 1994).

As a final contribution, the results of the current study imply that it seems necessary to examine acceptance of highly-automated tools in different contexts. Specifically, participants were presented with the same highly-automated tool with equal technical features but still evaluations of this tool differed slightly to moderately based on the context. On the one hand this supports research by M. K. Lee (2018) who showed the contextual sensitivity of acceptance of automatic processes (automatic decisions are less accepted for human tasks such as hiring in comparison to mechanical tasks such as work scheduling). On the other hand, our findings preliminary indicate that an alternative explanation for the findings of her study could be differing stakes between the situations. Within the high-stakes context of automated selection interviews, people seem to perceive automation as less favorable than for the low-stakes context interview training. Consequently, this could possibly point towards the stakes of a situation as another contextual factor influencing acceptance of automation and opens avenues for future research on contextual influences regarding acceptance of automated tools.

Limitations

At least four limitations should be mentioned. First, our experimental design does not allow to control for the effects of the decision agent or the use of a virtual interviewer within the highly-automated interview (cf. Ötting & Maier, 2018). This means we do not know if

participants had reacted differently to a situation in which the interview procedure is automated but then a human decision agent provides feedback for the interview performance. It is also not possible to discern in which way the virtual character affected participants reactions. Therefore, future research could experimentally manipulate the way of data gathering and data evaluation to explore if participants might react especially negatively in cases where the entire interview process is automated (see Nolan et al., 2016, for a similar study from the perspective of hiring managers) and where there is no interviewer manifestation. However, it could also be possible that the distinction between the data collection and the decision afterwards makes no difference as applicants do not distinguish between the data collection and data evaluation method.

The first limitation is also due to the fact that we described one specific-automated interview tool and the results might be different for other versions of such tools. For instance, people might react differently to tools that do not include a virtual interviewer. Up to now, it is not clear how the results would differ. Adding a virtual character should add social presence (Nass et al., 1994), and interview tools without virtual characters might therefore be evaluated even worse. However, it is also possible that the virtual character is perceived as eerie, leading to an uncanny valley effect (Mori, 1970; Mori et al., 2012). Nevertheless, we chose to use this kind of highly-automated interview as it prototypically automates the entire interview process following the automation model by Parasuraman et al. (2000) and as we believe that an automatic interview with a virtual character is the closest automated adaptation of a videoconference interview. However, this should not cover the fact that this experiment can only be a starting point to examine how the automation of different parts of interviews affects people's reactions and behavior during such interviews.

Second, participants only watched a video of the interview situations. According to Nolan and colleagues (2016) this is an adequate first step as a proof of concept if the

proposed effects really exist. Compared to asking participants to imagine themselves being in a situation (as in, e.g., Atzmüller & Steiner, 2010; Shelton & Stewart, 2004), we followed M. K. Lee (2018) and introduced participants to an observer role, which, we believe, has two advantages: (a) Participants can judge the situation as if their friend had provided them with a detailed description of the situation she/he experienced (comparable to a word-of-mouth description of an application situation, cf. Van Hove & Lievens, 2009); and (b) this design allowed us to keep the information provided to participants in all conditions constant. Nevertheless, future studies should let participants interact directly with the respective technological tools in order to examine if the differences between highly-automated and videoconference interviews hold when people are really exposed to these approaches (see Footnote 3 for the results of a small follow-up study where we had participants interacting live with the highly-automated interview tool and compared this to live videoconference interviews; the results of this follow-up study support the overall conclusion of the main study but also indicate that parts of Hypothesis 1 might only apply to live interviews).

Third, participants were predominantly students, so results might not generalize to a more mature sample. However, a student sample was deemed an appropriate sample for the current study because basic job interview training might be more relevant for entry level jobs (cf., Langer et al., 2016) and because highly-automated solutions might be more useful to screen a large applicant pool (cf., Campion et al., 2016), which might be more likely for entry level positions compared to jobs in higher management.

Fourth, we used a new measure of overall acceptance of an HRM situation drawn from user experience research (i.e., the attractiveness scale of the user experience questionnaire UEQ, Laugwitz et al., 2008). Nevertheless, this measure proved to be useful to directly compare low and high-stake situations, something that would have been more problematic with more specific outcome measures (e.g., organizational attractiveness).

Main Practical Implications

Organizations should be careful when establishing similar highly-automated interviews as they seem to be less accepted than videoconference interviews. Implications of this study might be especially detrimental for high-stake interviews (e.g., in selection contexts), where word-of-mouth plays a crucial role (Van Hoye & Lievens, 2009). If applicants react negatively to automated interviews this might have aggravating outcomes for organizations using them. Not only could current applicants withdraw from the application process (Uggerslev, Fassina, & Kraichy, 2012), they could also advise their friends against applying for a job.

However, there are organizations that regularly use automated approaches. For them, and for organizations who would like to apply similar approaches, we recommend to think about ways to mitigate negative effects. For instance, it might be a good idea to provide participants with information about unfamiliar procedures, as this can increase transparency, decrease privacy concerns and improve participants' fairness perceptions (cf., McCarthy et al., 2017; Mittelstadt et al., 2016). However, information does not always lead to better acceptance, its effect in the context of automation seems to be especially complicated (Langer et al., 2018), and with growing automation and more complex technologies underlying this automation (e.g., deep learning algorithms), providing information can become a challenge (Mittelstadt et al., 2016; Zerilli, Knott, Maclaurin, & Gavaghan, 2018). Therefore, it is up to designers of automated tools as well as a challenge for interdisciplinary research to make these tools as controllable, social, and transparent as possible (see also the discussion on explainable artificial intelligence that is currently shaking the field of computer science; Biran & Cotton, 2017; T. Miller, Howe, & Sonenberg, 2017).

Future Research

It is still unanswered what organizations think of highly-automated interviews (see Nolan et al., 2016 for a first impression on how hiring managers might react to highly-automated selection). In addition, research should continue to evaluate the acceptance of highly-automated procedures. With ongoing technological development, people might become familiar with these procedures, thus increasing transparency and comprehensibility – and this could lead to entirely different reactions compared to the ones found in the current study. However, it is also likely that even more advanced highly-automated tools (e.g., tools based on deep machine learning) are even less transparent and comprehensible than current ones (T. Miller et al., 2017). Imagine an interview tool that recommends certain applicants to hiring managers. If the tool provides no explanation on its decision, hiring managers can follow the recommendation without being able to challenge it or to explain it to rejected applicants. Alternatively, they can ignore the recommendation but are equally unable to explain why they chose to do so. In any case, hiring managers as well as applicants would be exposed to a non-transparent, non-explainable and non-challengeable highly-automated decision, probably leading to especially negative reactions as well as serious legal, moral and ethical issues and challenges (Zerilli et al., 2018).

Conclusion

Applicant reaction research on highly-automated procedures in HRM is still in its infancy. The present study is one of the first studies examining acceptance of a highly-automated interview approach. Our findings caution organizations to apply such procedures with care especially within high-stake scenarios. Moreover, we hope that the current study motivates further research to investigate perceptions but also the validity of advanced technologies used for HRM contexts in order to improve understanding and design of these technologies.

References

- Ajzen, I. (2002a). Constructing a TPB questionnaire: Conceptual and methodological considerations. Retrieved from http://chuang.epage.au.edu.tw/ezfiles/168/1168/attach/20/pta_41176_7688352_57138.pdf
- Ajzen, I. (2002b). Perceived behavioral control, self-efficacy, locus of control, and the theory of planned behavior. *Journal of Applied Social Psychology, 32*, 665–683. <https://doi.org/10.1111/j.1559-1816.2002.tb00236.x>
- An, L., Yang, S., & Bhanu, B. (2015). Efficient smile detection by Extreme Learning Machine. *Neurocomputing, 149*, 354–363. <https://doi.org/10.1016/j.neucom.2014.04.072>
- Arkes, H. R., Shaffer, V. A., & Medow, M. A. (2007). Patients derogate physicians who use a computer-assisted diagnostic aid. *Medical Decision Making, 27*, 189–202. <https://doi.org/10.1177/0272989X06297391>
- Atzmüller, C., & Steiner, P. M. (2010). Experimental vignette studies in survey research. *Methodology, 6*, 128–138. <https://doi.org/10.1027/1614-2241/a000014>
- Baltrušaitis, T., Robinson, P., & Morency, L.-P. (2016). Openface: An open source facial behavior analysis toolkit. In *2016 IEEE Winter Conference on Applications of Computer Vision (WACV)*, (pp. 1–10). Lake Placid, NY: IEEE. <https://doi.org/10.1109/wacv.2016.7477553>
- Barrick, M. R., Shaffer, J. A., & DeGrassi, S. W. (2009). What you see may not be what you get: Relationships among self-presentation tactics and ratings of interview and job performance. *Journal of Applied Psychology, 94*, 1394–1411. <https://doi.org/10.1037/a0016532>

- Barry, B., & Fulmer, I. S. (2004). The medium and the message: The adaptive use of communication media in dyadic influence. *Academy of Management Review*, *29*, 272–292. <https://doi.org/10.2307/20159033>
- Bauer, T. N., Truxillo, D. M., Sanchez, R. J., Craig, J. M., Ferrara, P., & Campion, M. A. (2001). Applicant reactions to selection: Development of the Selection Procedural Justice Scale (SPJS). *Personnel Psychology*, *54*, 387–419. <https://doi.org/10.1111/j.1744-6570.2001.tb00097.x>
- Bauer, T. N., Truxillo, D. M., Tucker, J. S., Weathers, V., Bertolino, M., Erdogan, B., & Campion, M. A. (2006). Selection in the information age: The impact of privacy concerns and computer experience on applicant reactions. *Journal of Management*, *32*, 601–621. <https://doi.org/10.1177/0149206306289829>
- Binns, R., Van Kleek, M., Veale, M., Lyngs, U., Zhao, J., & Shadbolt, N. (2018). “It’s reducing a human being to a percentage”; Perceptions of justice in algorithmic decisions. *ArXiv:1801.10408 [Cs]*, 1–14. <https://doi.org/10.1145/3173574.3173951>
- Biran, O., & Cotton, C. (2017). Explanation and justification in machine learning: A survey. In *Proceedings of the 17th international joint conference on artificial intelligence IJCAI* (pp. 8–13). Melbourne, Australia. Retrieved from http://www.intelligentrobots.org/files/IJCAI2017/IJCAI-17_XAI_WS_Proceedings.pdf#page=8
- Blacksmith, N., Willford, J. C., & Behrend, T. S. (2016). Technology in the employment interview: A meta-analysis. *Personnel Assessment and Decisions*, *2*, 2. <https://doi.org/10.25035/pad.2016.002>
- Bobko, P., Roth, P. L., & Buster, M. A. (2007). The usefulness of unit weights in creating composite scores: A literature review, application to content validity, and meta-

analysis. *Organizational Research Methods*, *10*, 689–709.

<https://doi.org/10.1177/1094428106294734>

Boersma, P., & Van Heuven, V. (2001). Speak and unSpeak with PRAAT. *Glott International*, *5*, 341–347.

Brenner, F. S., Ortner, T. M., & Fay, D. (2016). Asynchronous video interviewing as a new technology in personnel selection: The applicant's point of view. *Frontiers in Psychology*, *7*, 863. <https://doi.org/10.3389/fpsyg.2016.00863>

Buehl, A.-K., Melchers, K. G., Macan, T., & Kühnel, J. (2018). Tell me sweet little lies: How does faking in interviews affect interview scores and interview validity? *Journal of Business and Psychology*, advance online publication. <https://doi.org/10.1007/s10869-018-9531-3>

Campion, M. C., Campion, M. A., Campion, E. D., & Reider, M. H. (2016). Initial investigation into computer scoring of candidate essays for personnel selection. *Journal of Applied Psychology*, *101*, 958–975. <https://doi.org/10.1037/apl0000108>

Chamorro-Premuzic, T., Akhtar, R., Winsborough, D., & Sherman, R. A. (2017). The datafication of talent: How technology is advancing the science of human potential at work. *Current Opinion in Behavioral Sciences*, *18*, 13–16.
<https://doi.org/10.1016/j.cobeha.2017.04.007>

Chapman, D. S., Uggerslev, K. L., & Webster, J. (2003). Applicant reactions to face-to-face and technology-mediated interviews: A field investigation. *Journal of Applied Psychology*, *88*, 944–953. <https://doi.org/10.1037/0021-9010.88.5.944>

Colquitt, J. A., & Zipay, K. P. (2015). Justice, fairness, and employee reactions. *Annual Review of Organizational Psychology and Organizational Behavior*, *2*, 75–99.
<https://doi.org/10.1146/annurev-orgpsych-032414-111457>

- Davis, F. D., Bagozzi, R. P., & Warshaw, P. R. (1989). User acceptance of computer technology: A comparison of two theoretical models. *Management Science*, *35*, 982–1003. <https://doi.org/10.1287/mnsc.35.8.982>
- DeGroot, T., & Kluemper, D. (2007). Evidence of predictive and incremental validity of personality factors, vocal attractiveness and the situational interview. *International Journal of Selection and Assessment*, *15*, 30–39. <https://doi.org/10.1111/j.1468-2389.2007.00365.x>
- Faliagka, E., Tsakalidis, A., & Tzimas, G. (2012). An integrated e-recruitment system for automated personality mining and applicant ranking. *Internet Research*, *22*, 551–568. <https://doi.org/10.1108/10662241211271545>
- Frauendorfer, D., Schmid Mast, M., Nguyen, L., & Gatica-Perez, D. (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*, 231–245. <https://doi.org/10.1007/s10919-014-0173-5>
- Gebhard, P., Mehlmann, G., & Kipp, M. (2011). Visual SceneMaker: A tool for authoring interactive virtual characters. *Journal on Multimodal User Interfaces*, *6*, 3–11. <https://doi.org/10.1007/s12193-011-0077-1>
- Gebhard, P., Schneeberger, T., André, E., Baur, T., Damian, I., König, C. J., & Langer, M. (2018). Serious games for training social skills in job interviews. *IEEE Transactions on Games*. <https://doi.org/doi.org/10.1109/TG.2018.2808525>
- Gilliland, S. W. (1993). The perceived fairness of selection systems: An organizational justice perspective. *Academy of Management Review*, *18*, 694–734. <https://doi.org/10.2307/258595>
- Gratch, J., Wang, N., Okhmatovskaia, A., Lamothe, F., Morales, M., van der Werf, R. J., & Morency, L.-P. (2007). Can virtual humans be more engaging than real ones? In

- International Conference on Human-Computer Interaction* (pp. 286–297). Berlin, Germany: Springer. Retrieved from http://link.springer.com/chapter/10.1007/978-3-540-73110-8_30
- Guchait, P., Ruetzler, T., Taylor, J., & Toldi, N. (2014). Video interviewing: A potential selection tool for hospitality managers – A study to understand applicant perspective. *International Journal of Hospitality Management*, *36*, 90–100. <https://doi.org/10.1016/j.ijhm.2013.08.004>
- Guo, X., Polanía, L. F., & Barner, K. E. (2018). Smile detection in the wild based on transfer learning. *ArXiv:1802.02185 [Cs]*. Retrieved from <http://arxiv.org/abs/1802.02185>
- Hausknecht, J. P., Day, D. V., & Thomas, S. C. (2004). Applicant reactions to selection procedures: An updated model and meta-analysis. *Personnel Psychology*, *57*, 639–683. <https://doi.org/10.1037/e518612013-214>
- Highhouse, S., Lievens, F., & Sinar, E. F. (2003). Measuring attraction to organizations. *Educational and Psychological Measurement*, *63*, 986–1001. <https://doi.org/10.1177/0013164403258403>
- HireVue. (2018). HireVue OnDemand. Retrieved August 5, 2018, from <https://www.hirevue.com/products/video-interviewing/ondemand>
- Hoff, K. A., & Bashir, M. (2015). Trust in automation: Integrating empirical evidence on factors that influence trust. *Human Factors*, *57*, 407–434. <https://doi.org/10.1177/0018720814547570>
- Holden, R. R., & Passey, J. (2010). Socially desirable responding in personality assessment: Not necessarily faking and not necessarily substance. *Personality and Individual Differences*, *49*, 446–450. <https://doi.org/10.1016/j.paid.2010.04.015>
- Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics*, *2*, 65–70.

- Intel. (2018). Intel RealSense technology. Retrieved from <https://software.intel.com/en-us/realsense/d400>
- Jansen, A., König, C. J., Stadelmann, E. H., & Kleinmann, M. (2012). Applicants' self-presentational behavior: What do recruiters expect and what do they get? *Journal of Personnel Psychology, 11*, 77–85. <https://doi.org/10.1027/1866-5888/a000046>
- Kätsyri, J., Förger, K., Mäkäriäinen, M., & Takala, T. (2015). A review of empirical evidence on different uncanny valley hypotheses: Support for perceptual mismatch as one road to the valley of eeriness. *Frontiers in Psychology, 6*, 390. <https://doi.org/10.3389/fpsyg.2015.00390>
- Langer, M., & König, C. J. (2018). Introducing and testing the Creepiness of Situation Scale (CRoSS). *Frontiers in Psychology, 9*, 2220. <https://doi.org/10.3389/fpsyg.2018.02220>
- Langer, M., König, C. J., & Fitili, A. (2018). Information as a double-edged sword: The role of computer experience and information on applicant reactions towards novel technologies for personnel selection. *Computers in Human Behavior, 81*, 19–30. <https://doi.org/10.1016/j.chb.2017.11.036>
- Langer, M., König, C. J., Gebhard, P., & André, E. (2016). Dear computer, teach me manners: Testing virtual employment interview training. *International Journal of Selection and Assessment, 24*, 312–323. <https://doi.org/10.1111/ijsa.12150>
- Langer, M., König, C. J., & Krause, K. (2017). Examining digital interviews for personnel selection: Applicant reactions and interviewer ratings. *International Journal of Selection and Assessment, 25*, 371–382. <https://doi.org/10.1111/ijsa.12191>
- Langer, M., Schmid Mast, M., Meyer, B., Maass, W., & König, C. J. (2019). Research in the era of sensing technologies and wearables. In R. N. Landers (Ed.), *Cambridge handbook of technology and employee behavior* (pp. 806–835). New York, NY: Cambridge University Press. Retrieved from 10.1017/9781108649636.030

- Laugwitz, B., Held, T., & Schrepp, M. (2008). Construction and evaluation of a user experience questionnaire. In *Symposium of the Austrian HCI and Usability Engineering Group* (pp. 63–76). Berlin, Germany: Springer. Retrieved from http://link.springer.com/chapter/10.1007/978-3-540-89350-9_6
- Lee, K. M., & Nass, C. (2003). Designing social presence of social actors in human computer interaction. In *Proceedings of the CHI 2003 conference on human factors in computing systems* (pp. 289–296). Fort Lauderdale, FL. <https://doi.org/10.1145/642611.642662>
- Lee, M. K. (2018). Understanding perception of algorithmic decisions: Fairness, trust, and emotion in response to algorithmic management. *Big Data & Society*, 5, 205395171875668. <https://doi.org/10.1177/2053951718756684>
- Lee, S. M., Kim, Y., & Lee, J. (1995). An empirical study of the relationships among end-user information systems acceptance, training, and effectiveness. *Journal of Management Information Systems*, 12, 189–202. <https://doi.org/10.1080/07421222.1995.11518086>
- Levashina, J., & Campion, M. A. (2006). A model of faking likelihood in the employment interview. *International Journal of Selection and Assessment*, 14, 299–316. <https://doi.org/10.1111/j.1468-2389.2006.00353.x>
- Malhotra, N. K., Kim, S. S., & Agarwal, J. (2004). Internet users' information privacy concerns (IUIPC): The construct, the scale, and a causal model. *Information Systems Research*, 15, 336–355. <https://doi.org/10.1287/isre.1040.0032>
- Maurer, T. J., Solamon, J. M., Andrews, K. D., & Troxtel, D. D. (2001). Interviewee coaching, preparation strategies, and response strategies in relation to performance in situational employment interviews: An extension of Maurer, Solamon, and Troxtel

- (1998). *Journal of Applied Psychology*, 86, 709–717. <https://doi.org/10.1037//0021-9010.86.4.709>
- McCarthy, J. M., Bauer, T. N., Truxillo, D. M., Campion, M. C., Van Iddekinge, C. H., & Campion, M. A. (2017). Using pre-test explanations to improve test-taker reactions: Testing a set of “wise” interventions. *Organizational Behavior and Human Decision Processes*, 141, 43–56. <https://doi.org/10.1016/j.obhdp.2017.04.002>
- McCarthy, J. M., & Goffin, R. (2004). Measuring job interview anxiety: Beyond weak knees and sweaty palms. *Personnel Psychology*, 57, 607–637. <https://doi.org/10.1111/j.1744-6570.2004.00002.x>
- McCole, P., Ramsey, E., & Williams, J. (2010). Trust considerations on attitudes towards online purchasing: The moderating effect of privacy and security concerns. *Journal of Business Research*, 63, 1018–1024. <https://doi.org/10.1016/j.jbusres.2009.02.025>
- Middendorf, C. H., & Macan, T. H. (2002). Note-taking in the employment interview: Effects on recall and judgments. *Journal of Applied Psychology*, 87, 293–303. <https://doi.org/10.1037//0021-9010.87.2.293>
- Miller, F. A., Katz, J. H., & Gans, R. (2018). The OD imperative to add inclusion to the algorithms of artificial intelligence. *OD Practitioner*, 50(1), 6–12.
- Miller, T., Howe, P., & Sonenberg, L. (2017). Explainable AI: Beware of inmates running the asylum. In *Proceedings of the 17th international joint conference on artificial intelligence IJCAI* (pp. 36–42). Melbourne, Australia. Retrieved from http://www.intelligentrobots.org/files/IJCAI2017/IJCAI-17_XAI_WS_Proceedings.pdf#page=36
- Mittelstadt, B. D., Allo, P., Taddeo, M., Wachter, S., & Floridi, L. (2016). The ethics of algorithms: Mapping the debate. *Big Data & Society*, 3, 205395171667967. <https://doi.org/10.1177/2053951716679679>

- Mori, M. (1970). Bukimi no tani [The uncanny valley]. *Energy*, 7, 33–35.
- Mori, M., MacDorman, K., & Kageki, N. (2012). The uncanny valley. *IEEE Robotics & Automation Magazine*, 19, 98–100. <https://doi.org/10.1109/MRA.2012.2192811>
- Naim, I., Tanveer, M. I., Gildea, D., & Hoque, M. E. (2018). Automated analysis and prediction of job interview performance. *IEEE Transactions on Affective Computing*, 9, 191–204. <https://doi.org/10.1109/TAFFC.2016.2614299>
- Nass, C., Steuer, J., & Tauber, E. R. (1994). Computers are social actors. In *Proceedings of the SIGCHI conference on Human factors in computing systems* (pp. 72–78). ACM. Retrieved from <http://dl.acm.org/citation.cfm?id=191703>
- Nolan, K., Carter, N., & Dalal, D. (2016). Threat of technological unemployment: Are hiring managers discounted for using standardized employee selection practices? *Personnel Assessment and Decisions*, 2, 4. <https://doi.org/10.25035/pad.2016.004>
- Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. *Science*, 349, 1–8. <https://doi.org/10.1126/science.aac4716>
- Ötting, S. K., & Maier, G. W. (2018). The importance of procedural justice in Human-Machine-Interactions: Intelligent systems as new decision agents in organizations. *Computers in Human Behavior*, 89, 27–39. <https://doi.org/10.1016/j.chb.2018.07.022>
- Parasuraman, R., Sheridan, T. B., & Wickens, C. D. (2000). A model for types and levels of human interaction with automation. *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans*, 30, 286–297. <https://doi.org/10.1109/3468.844354>
- Peck, J. A., & Levashina, J. (2017). Impression management and interview and job performance ratings: A meta-analysis of research design with tactics in mind. *Frontiers in Psychology*, 8, 201. <https://doi.org/10.3389/fpsyg.2017.00201>

- Peeters, H., & Lievens, F. (2006). Verbal and nonverbal impression management tactics in behavior description and situational interviews. *International Journal of Selection and Assessment, 14*, 206–222.
- Potosky, D. (2008). A conceptual framework for the role of the administration medium in the personnel assessment process. *Academy of Management Review, 33*, 629–648.
<https://doi.org/10.5465/amr.2008.32465704>
- Precire. (2017). Precire Technologies. Retrieved August 5, 2017, from
<https://www.precire.com/de/start/>
- Ruben, M. A., Hall, J. A., & Schmid Mast, M. (2015). Smiling in a job interview: When less is more. *Journal of Social Psychology, 155*, 107–126.
<https://doi.org/10.1080/00224545.2014.972312>
- Ryan, A. M., & Ployhart, R. E. (2000). Applicants' perceptions of selection procedures and decisions: A critical review and agenda for the future. *Journal of Management, 26*, 565–606. [https://doi.org/10.1016/s0149-2063\(00\)00041-6](https://doi.org/10.1016/s0149-2063(00)00041-6)
- Schmid Mast, M., Frauendorfer, D., Gatica-Perez, D., Choudhury, T., & Odobez, J.-M. (2017). A step towards automatic applicant selection: Predicting job performance based on applicant nonverbal interview behavior. *Manuscript Submitted for Publication.*
- Sears, G., Zhang, H., Wiesner, W. D., Hackett, R. W., & Yuan, Y. (2013). A comparative assessment of videoconference and face-to-face employment interviews. *Management Decision, 51*, 1733–1752. <https://doi.org/10.1108/MD-09-2012-0642>
- Shelton, N. J., & Stewart, R. E. (2004). Confronting perpetrators of prejudice: The inhibitory effects of social costs. *Psychology of Women Quarterly, 28*, 215–223.
<https://doi.org/10.1111/j.1471-6402.2004.00138.x>

- Sitzmann, T. (2011). A meta-analytic examination of the instructional effectiveness of computer-based simulation games. *Personnel Psychology, 64*, 489–528.
<https://doi.org/10.1111/j.1744-6570.2011.01190.x>
- Smith, H. J., Milberg, S. J., & Burke, S. J. (1996). Information privacy: Measuring individuals' concerns about organizational practices. *Management Information Systems Quarterly, 20*, 167–196. <https://doi.org/10.2307/249477>
- Stone, D. L., Deadrick, D. L., Lukaszewski, K. M., & Johnson, R. (2015). The influence of technology on the future of human resource management. *Human Resource Management Review, 25*, 216–231. <https://doi.org/10.1016/j.hrmr.2015.01.002>
- Stone, D. L., Lukaszewski, K. M., Stone-Romero, E. F., & Johnson, T. L. (2013). Factors affecting the effectiveness and acceptance of electronic selection systems. *Human Resource Management Review, 23*, 50–70. <https://doi.org/10.1016/j.hrmr.2012.06.006>
- Tene, O., & Polonetsky, J. (2015). A theory of creepy: Technology, privacy, and shifting social norms. *Yale Journal of Law and Technology, 16*, 59–102.
- Uggerslev, K. L., Fassina, N. E., & Kraichy, D. (2012). Recruiting through the stages: A meta-analytic test of predictors of applicant attraction at different stages of the recruiting process. *Personnel Psychology, 65*, 597–660.
<https://doi.org/10.1111/j.1744-6570.2012.01254.x>
- Van Hove, G., & Lievens, F. (2009). Tapping the grapevine: A closer look at word-of-mouth as a recruitment source. *Journal of Applied Psychology, 94*, 341–352.
<https://doi.org/10.1037/a0014066>
- Wagner, J., Lingensfelder, F., Baur, T., Damian, I., Kistler, F., & André, E. (2013). The social signal interpretation (SSI) framework: Multimodal signal processing and recognition in real-time. In *Proceedings of the 21st ACM international conference on Multimedia* (pp. 831–834). Retrieved from <http://dl.acm.org/citation.cfm?id=2502223>

- Walter, N., Ortbach, K., & Niehaves, B. (2015). Designing electronic feedback: Analyzing the effects of social presence on perceived feedback usefulness. *International Journal of Human-Computer Studies*, 76, 1–11. <https://doi.org/10.1016/j.ijhcs.2014.12.001>
- Warszta, T. (2012). *Application of Gilliland's model of applicants' reactions to the field of web-based selection*. [Unpublished dissertation]. Christian-Albrechts Universität Kiel, Germany. Retrieved from http://macau.uni-kiel.de/receive/dissertation_diss_00008734
- Zerilli, J., Knott, A., Maclaurin, J., & Gavaghan, C. (2018). Transparency in algorithmic and human decision-making: Is there a double standard? *Philosophy & Technology*, advance online publication. <https://doi.org/10.1007/s13347-018-0330-6>
- Zyda, M. (2005). From visual simulation to virtual reality to games. *Computer*, 38, 25–32. <https://doi.org/10.1109/mc.2005.297>

Table 1.
Correlations and Cronbach's Alpha for the Study Variables.

Scale	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.
1. Social Presence	.87											
2. Interpersonal Treatment	.57**	.90										
3. Perceived Self-Efficacy	.22*	.17	.79									
4. Perceived Controllability	.35**	.37**	.50**	.65								
5. Consistency	-.10	.11	.03	.01	.81							
6. Fairness	.39**	-.49**	-.38**	.54**	.36**	.92						
7. Emotional Creepiness	.37**	-.42**	-.41**	-.41**	.01	-.42**	.88					
8. Creepy Ambiguity	.30**	-.30**	-.35**	-.39**	.11	.75**	.28**	.85				
9. Privacy Concerns	-.25**	-.05	-.12	-.05	-.20*	.13	.12	-.16	.81			
10. High-Stake Evaluations	.09	-.07	-.05	-.04	.01	.18*	.25**	.13	-.02	.86		
11. Attractiveness	.57**	.47**	.46**	.54**	.11	.61**	.56**	-.46**	-.24**	-.19*	.90	
12. Context	-.08	-.01	.07	.05	.16	-.09	-.09	-.07	.25**	-.28**	.21*	-
13. Interview Type	.16	.02	.26**	.30**	.04	-.16	-.24**	-.19*	.15	.06	.18*	-.04

Note. Coding of context: -1 = high-stake context, 1 = low-stake context. Coding of interview type: -1 = highly-automated interview, 1 = videoconference interview. $N = 123$. Numbers in the diagonal represent Cronbach's alpha of the scales.

* $p < .05$, ** $p < .01$.

Table 2.

ACCEPTANCE OF HIGHLY-AUTOMATED INTERVIEWS

46

Means, Standard Deviations, Single ANOVA Results (including partial η^2) and Hypotheses for the Dependent Variables.

Variable	Condition				ANOVA					
	VC-LS	VC-HS	AI-LS	AI-HS	AI vs. VC		HS vs. LS		Interaction	
	<i>M</i> (<i>SD</i>)	<i>M</i> (<i>SD</i>)	<i>M</i> (<i>SD</i>)	<i>M</i> (<i>SD</i>)	<i>F</i> (1,119)	η_p^2	<i>F</i> (1,119)	η_p^2	<i>F</i> (1,119)	η_p^2
Social Presence	4.07 (1.18)	4.39 (1.11)	3.85 (1.12)	3.87 (1.33)	3.00	.03	0.65	.01	0.47	.00
Interpersonal Treatment	3.88 (0.70)	3.99 (0.71)	3.95 (0.73)	3.86 (0.77)	0.07	.00	0.01	.00	0.54	.00
Perceived Self-Efficacy	5.04 (0.89)	4.94 (1.02)	4.51 (1.17)	4.25 (1.37)	9.02**	.07	0.81	.00	0.14	.00
Perceived Controllability	4.63 (0.85)	4.89 (0.93)	4.41 (0.94)	3.92 (0.94)	12.99**	.10	0.53	.00	5.31*	.04
Consistency	3.37 (0.63)	3.15 (0.91)	3.40 (0.83)	3.08 (0.96)	0.01	.00	3.11	.03	0.10	.00
Fairness	3.50 (0.74)	3.45 (0.94)	3.63 (0.77)	2.71 (1.02)	3.88	.03	9.35**	.07	7.62**	.06
Emotional Creepiness	3.19 (1.17)	3.19 (1.28)	3.35 (1.48)	3.89 (1.41)	3.19	.03	1.24	.01	1.24	.01
Creepy Ambiguity	3.44 (1.15)	3.21 (1.26)	3.57 (1.23)	4.31 (1.21)	7.89**	.06	1.34	.01	4.94*	.04
Privacy Concerns	4.61 (0.85)	4.80 (0.97)	5.03 (1.08)	5.15 (1.04)	4.69*	.04	0.78	.01	0.04	.00
Attractiveness	-0.09 (1.00)	-0.22 (1.14)	-0.17 (0.98)	-0.99 (1.12)	4.82*	.04	6.22*	.05	3.31	.03

Note. VC = videoconference interview, AI = highly-automated interview, LS = low-stake context, HS = high-stake context. $n_{VC-LS} = 30$, $n_{VC-HS} = 32$, $n_{AI-LS} = 32$, $n_{AI-HS} = 29$.

* $p < .05$, ** $p < .01$.

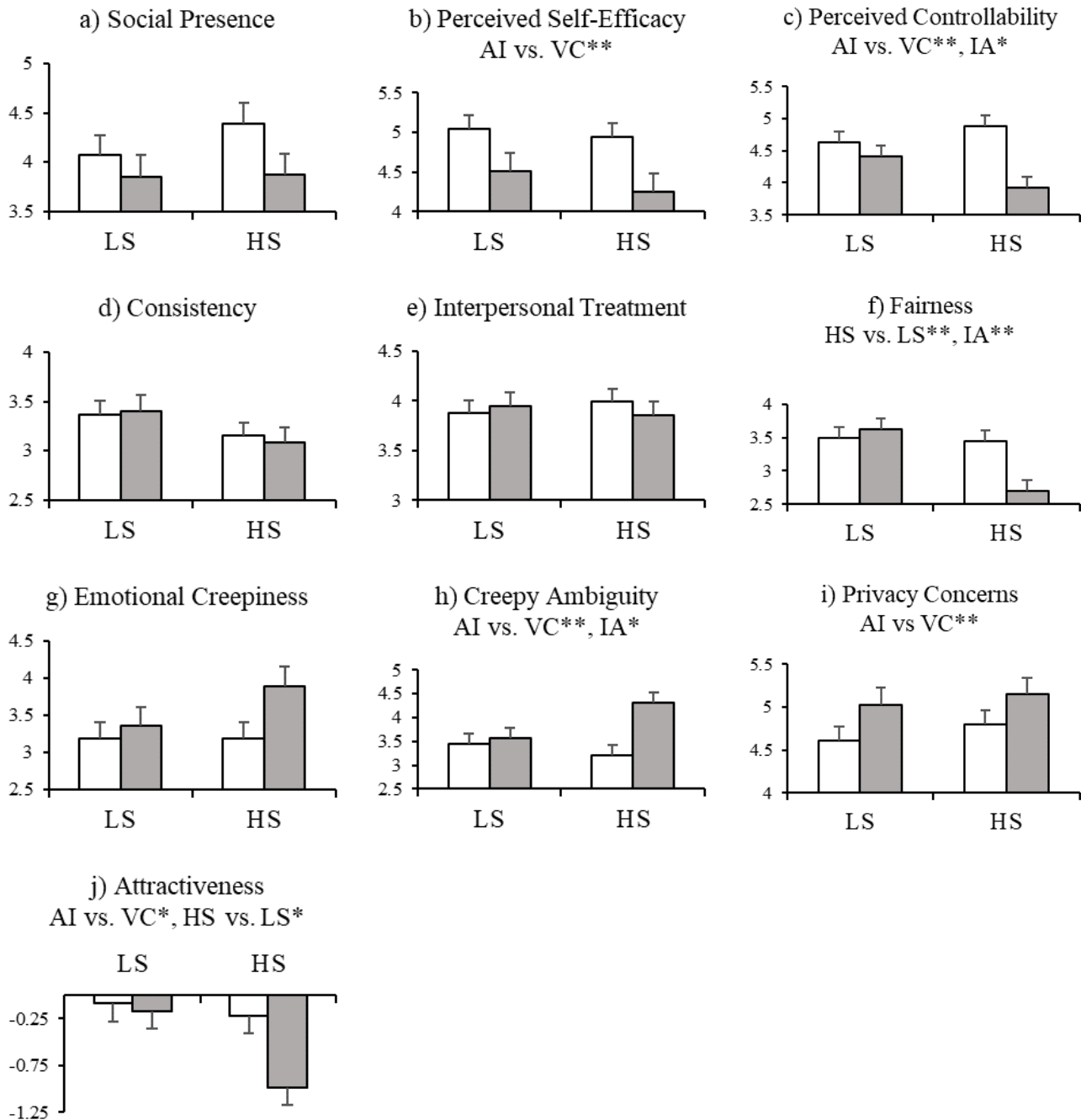


Figure 1. Results of the experimental groups for single variables. LS = low-stake context, HS = high-stake context, AI vs. VC = there is a main effect between highly-automated and videoconference interviews, HS vs. LS = there is a main effect between the high-stake and the low-stake context, IA = there is an interaction effect between the two independent variables. □ = videoconference interviews, ■ = highly-automated interviews.

* $p < .05$, ** $p < .01$.

Table 3.

Regression Results for the Mediation of the Hypothesized Mediators between Highly-Automated vs. Videoconference Interviews and Attractiveness of the Procedure

Model	R^2	Coefficient	SE	p	95% Confidence Interval
Model complete	.61	-	-	<.01	-
Social Presence → Attractiveness		0.28	0.08	<.01	[0.13, 0.43]
Interpersonal Treatment → Attractiveness		-0.01	0.13	.93	[-0.26, 0.24]
Perceived Self-Efficacy → Attractiveness		0.12	0.07	.08	[-0.02, 0.26]
Perceived Controllability → Attractiveness		0.16	0.09	.09	[-0.03, 0.34]
Consistency → Attractiveness		0.05	0.09	.59	[-0.13, 0.23]
Fairness → Attractiveness		0.32	0.10	<.01	[0.12, 0.52]
Emotional Creepiness → Attractiveness		-0.13	0.08	.11	[-0.30, 0.03]
Creepy Ambiguity → Attractiveness		-0.07	0.08	.41	[-0.23, 0.10]
Privacy Concerns → Attractiveness		-0.08	0.07	.28	[-0.22, 0.07]
AI vs. VC → Attractiveness		-0.05	0.07	.53	[-0.19, 0.10]

Note. AI = highly-automated interview, VC = videoconference interview. Coding of the variable AI vs. VC: -1 = highly-automated interview, 1 = videoconference interview. The 95% confidence interval for the effects is obtained by the bias-corrected bootstrap with 50,000 resamples.

Table 4.

Results for the Indirect Effects of Highly-Automated vs. Videoconference Interviews over the Hypothesized Mediators on Attractiveness of the Procedure

Model	IE_{med}	SE_{Boot}	95% Confidence Interval
Complete indirect effect	.22	0.07	[0.09, 0.36]
AI vs. VC → Social Presence → Attractiveness	.05	0.03	[0.001, 0.12]
AI vs. VC → Interpersonal Treatment → Attractiveness	.00	0.01	[-0.02, 0.02]
AI vs. VC → Perceived Self-Efficacy → Attractiveness	.03	0.02	[-0.002, 0.09]
AI vs. VC → Perceived Controllability → Attractiveness	.04	0.03	[-0.01, 0.12]
AI vs. VC → Consistency → Attractiveness	.00	0.01	[-0.01, 0.02]
AI vs. VC → Fairness → Attractiveness	.04	0.03	[0.001, 0.12]
AI vs. VC → Emotional Creepiness → Attractiveness	.03	0.02	[-0.01, 0.09]
AI vs. VC → Creepy Ambiguity → Attractiveness	.02	0.03	[-0.03, 0.08]
AI vs. VC → Privacy Concerns → Attractiveness	.01	0.02	[-0.01, 0.05]

Note. AI = highly-automated interviews, VC = videoconference interviews. Coding of the variable AI vs. VC: -1 = highly-automated interviews, 1 = videoconference interviews. The 95% confidence interval for the effects is obtained by the bias-corrected bootstrap with 50,000 resamples. IE_{med} = completely standardized indirect effect of the mediation. SE_{Boot} = standard error of the bootstrapped effect sizes.

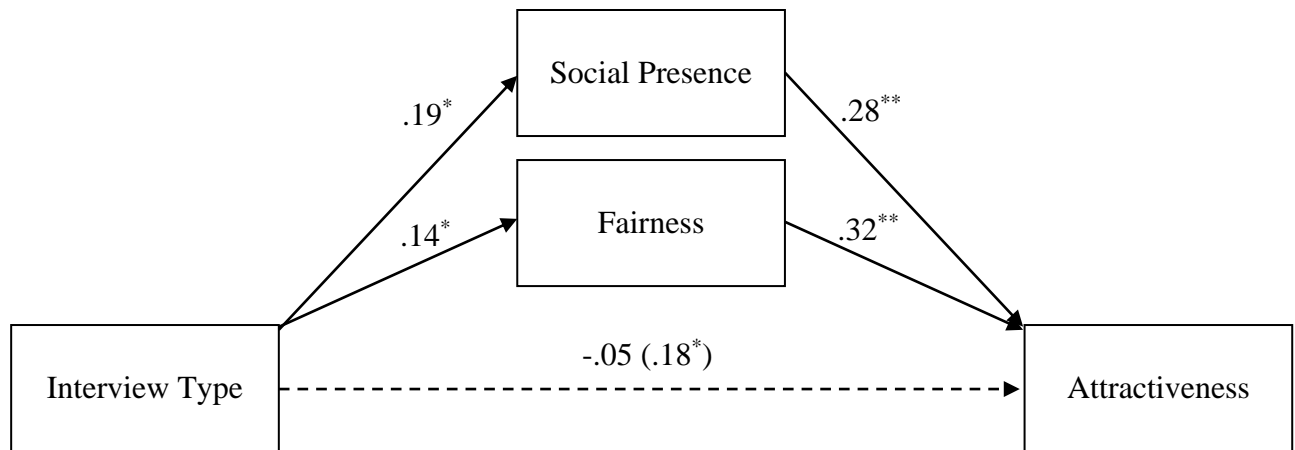


Figure 2. Mediation of social presence and fairness between highly-automated vs. videoconference interviews and attractiveness of the procedure. The number in brackets indicates the zero-order correlation between interaction type and attractiveness. Coding of interview type: -1 = highly-automated interview, 1 = videoconference interview. $^*p < .05$, $^{**}p < .01$.