Artificial Intelligence as Colleague and Supervisor: Successful and Fair Interactions Between Intelligent Technologies and Employees at Work

by

Sonja K. Ötting, M.Sc.

A Dissertation Submitted in Partial Fulfilment

of the Requirements for the Degree of

Doctor of Science (Dr. rer. nat.)

Department of Psychology

Faculty of Psychology and Sports Science

Bielefeld University

November 2020

Supervisor and primary examiner:	Prof. Dr. Günter W. Maier
	Department of Psychology
	Bielefeld University
Secondary examiner:	Prof. Dr. Cornelius König
	Department of Psychology
	Saarland University

Acknowledgements

First of all, I want to most sincerely thank Prof. Dr. Günter W. Maier, who gave me this opportunity to further develop my academic skills and dive into this intriguing research topic. Prof. Maier always provided me with sophisticated and constructive feedback, while giving me trust and autonomy to make my own decisions. He created a working environment that helped me grow and develop my skills as a scientist and I am deeply thankful for it. I would also like to thank Prof. Dr. Cornelius König, who agreed to serve as secondary examiner, and Prof. Dr. Gerd Bohner, who agreed to preside over the examination commission. Thank you very much for your time and effort.

Many thanks also to my fellow PhD students. You are the best colleagues I could have ever imagined. Thank you for all your advice, the ears you lend me whenever I struggled, and the many (analogue and virtual) tea(m) times. A special thanks goes to Barbara Steinmann. When I had to fight through difficult times, you helped and supported me to trust in myself and work up the courage to engage in this endeavour of being a researcher.

I am deeply thankful for the love and encouragement of my family. You were always there for me, helped me find my way through every challenge and had my back when I made an important decision. I am most grateful for the loving support of my husband, Christoph. You were exceptional at distracting me when stress threatened to gain the upper hand and putting me back on track before procrastination got out of hand. You were always understanding and thoughtful and with your calmness every problem seemed solvable.

To you who cannot be with me anymore, without you, I would not be who I am and I miss you badly.

Summary

Employees increasingly share workplaces and tasks with artificial intelligence (AI). Intelligent technologies have been developing so rapidly that they can take on the role of a co-worker (e.g., a robot that works in a shared workspace) or even a supervisor (e.g., an algorithm that makes decisions). Both types of relations between AI and employee affect employee motivation, well-being, and performance. In three studies, the present work therefore examines AI as robotic co-workers and as supervisors. More specifically, I investigated which robot design features make human-robot interaction (HRI) at work most successful and how and why effects of procedural justice differ depending on whether humans or AI act as decision agent.

In Study 1, we focussed on AI as co-worker and meta-analytically integrated 81 studies on the relation of five robot design features (i.e., feedback and visibility of the interface, adaptability and autonomy of the controller, and human likeness of the appearance) with seven indicators of successful HRI (i.e., task performance, cooperation, satisfaction, acceptance, trust, mental workload, and situation awareness). Results showed that the features of interface and controller significantly affected successful HRI, while human likeness did not. Moderation analyses revealed that only design features of the controller had significant specific effects in addition to those on task performance and satisfaction: Adaptability affected cooperation and acceptance, and autonomy affected mental workload.

In Studies 2 and 3, we focussed on AI as supervisor and examined and compared procedural justice effects of human and AI decision agents on employee attitudes and behaviour. To this end, we conducted two vignette experiments in each study. In Study 2, we investigated whether the type of decision agent (human vs. AI) influenced the effects of procedural justice on employee attitudes and behaviour. The results showed no differences in effect sizes between humans or AI as decision agent, emphasising the importance of

procedural justice for both decision agents. In Study 3, we compared strength and specificity of four mediators of procedural justice effects, investigated differences between decision agents and examined responsibility as explaining mechanism for these differences. The results for both types of decision agents showed trust as strongest mediator for effects on attitudes, and negative affect as strongest mediator for effects on behaviour. When comparing the two types of decision agents, trust as mediator was less pronounced for AI compared to human decisions, whereas no difference between the two types of decision agents was found for negative affect. Additionally, we confirmed the responsibility that is attributed to a decision agent as underlying mechanism for these differences.

In summary, the present work extends the understanding of employee interactions with AI as co-worker and supervisor at work by integrating theories from industrial and organisational psychology as well as engineering and information science. The results provide valuable insights for theory development in HRI and organisational justice concerning the integration and investigation of context factors, of effects of robot design characteristics on successful HRI and of characteristics of the decision agents that might influence justice effects. Moreover, the results provide recommendations for engineers, AI designers and human resource practitioners on what to bear in mind when planning to develop and implement AI in the workplace.

Table of contents

Introduction

Search engines that answer any question that could possibly arise at work, apps that schedule and allocate tasks and shifts (Franklin et al., 2014; Machado et al., 2016), algorithms that screen job candidates (Dineen et al., 2004; Liem et al., 2018), and robots that work side-by side with employees to accomplish shared tasks (Gombolay et al., 2014)—technologies equipped with artificial intelligence (AI), in various forms of appearance, are on the rise in the workplace. The estimation of the amount of employees that use some form of AI at work has risen from 32% in 2018 to 50% in 2019 (Oracle & Future Workplace, 2019). Employees already interact with intelligent technologies in their everyday working life, and this will increase even more in the future.

Less advanced technologies, such as personal computers or automated production lines, are an integral part of many workplaces. Intelligent technologies, such as intelligent robots or algorithms in applications, however, will make an even bigger impact (Kauffeld & Maier, 2020; Steil & Maier, 2020) because they can fulfil tasks that used to be exclusively assigned to human employees and are able to directly interact with employees (Cascio & Montealegre, 2016). Similar to relationships between employees in organisations (Dutton & Ragins, 2007; Shanock et al., 2012), intelligent technologies can interact with employees in different roles, the most important being co-worker (e.g., in form of a robot that works in a shared workspace; Onnasch et al., 2016) and supervisor (e.g., in form of an algorithm that makes leadership decisions; Wesche & Sonderegger, 2019). Both types of relationships, when investigated between humans, were shown to affect employee motivation, well-being, and performance (Basford & Offermann, 2012; Sherony & Green, 2002). It is therefore of eminent importance to explore the effects of these interactions with intelligent technologies to be able to design workplaces and AI in a way that benefits employees as well as the organisation.

However, industrial and organisational (IO) psychology, specifically research on work design and the design of organisational decisions, does not yet mirror the importance of intelligent technologies with an equally prominent place in its theories. In work design theories, technology is only incorporated as a mere tool, not as interaction partner (Morgeson & Humphrey, 2006). Work design research merely started to investigate possible effects of advanced technologies in the workplace (Bharadwaj et al., 2020; Parker & Grote, 2020) and organisational justice research only recently recognized intelligent technologies as possible decision agents (Brockner & Wiesenfeld, 2020). Research on human-technology interactions, on the other hand, produces a large amount of empirical studies but seldom focuses on the workplace as an important and critical context for these interactions. In addition, most studies lack a sound theoretical foundation and do not include insights from IO psychology, which are essential in order to describe and understand interactions at work. Linking theory, research methods, and empirical insights from IO psychology and human-technology interactions will enrich both disciplines, help to develop theory, and provide vital insights for practitioners.

In order to address these research gaps, the aim of this dissertation is twofold. The first aim is to investigate AI as robotic co-workers and how different robot design features influence the success of human-robot interaction (HRI) at work. The second aim is to investigate AI as supervisors and how the effects and explaining mechanisms of procedurally just decisions differ between humans and AI as decision agents. This dissertation therefore contributes to the literature in three ways. First, I identified those robot design features that enable successful HRI at work. By drawing from engineering and information science as well as psychological sources, this extends the understanding of human-AI interactions at work. With this, I answer the calls for interdisciplinary research that are growing ever louder (Rhoten & Parker, 2004; Zhu & Fu, 2019). Second, I investigated differences between human and AI decision agents concerning the perception of fairness, and its effects and explaining mechanisms. Doing this, I conducted research that accounts for the important role of technology in the workplace (Parker & Grote, 2020) and with that this work further develops organisational justice theories. Third, I compared the strength and specificity of explaining mechanisms of procedural justice effects, which has been demanded multiple times by prominent researchers (Colquitt et al., 2013; Colquitt & Zipay, 2015). Fourth, I combined advanced meta-analytical methods (Cheung, 2015), experimental research designs (Aguinis & Bradley, 2014), and the replication of results (Kepes & McDaniel, 2013) to be able to provide diverse and rigorous research and give reliable recommendations for both research and application in work, decision and AI design.

Theoretical Background

A technology can broadly be defined as artificially intelligent if it can achieve human-level performance in some cognitive task (Negnevitsky, 2005) or more specifically, artificial intelligence is "an agent's ability to achieve goals in a wide range of environments" (Legg & Hutter, 2007). An agent is a software representation of a real entity, capable of deciding and acting with a certain degree of autonomy (Sammut & Webb, 2017). In order to be capable of deciding and acting, the agent needs sensors to represent the environment it interacts with or data to accurately make decisions (Legg & Hutter, 2007). Often the agents are specialized to a specific goal (such as finding the best applicant for a position) with certain cognitive task (such as calculating scores for each applicant) that can be applied in a range of environments (such as different positions in a range of organisations) (Negnevitsky, 2005). In the workplace, artificially intelligent agents are predominantly implemented in the form of intelligent robots (with the purpose of, e.g., planning the assembly of a product or welding a work piece) or in the form of intelligent software applications (with the purpose of, e.g., deciding about task and shift scheduling or applicant selection).

Intelligent Technologies as Co-Workers

Intelligent technologies that are supposed to function as a co-worker need certain abilities to be able to share a workspace with employees and fulfil allocated tasks. The technology needs to sense its physical environment (in order to safely and successfully interact with it) and needs a physical embodiment (in order to fulfil tasks where physical actions are needed). Intelligent robots offer both of these abilities (Coiffet & Chirouze, 1983).

While a robot in general is a multifunctional multi-link programmable device, enabled to fulfil predetermined tasks (e.g., IFR Statistical Department, 2019; International Organization for Standardization, 2012; Spong et al., 2020), an intelligent robot is able to sense its environment and respond to changes in it, in order to perform diverse tasks (Coiffet & Chirouze, 1983). The intelligent robot can exhibit flexibly programmable behaviour and make use of sensor data and complex internal control systems to interact with its environment (Fong et al., 2003). Because of a certain autonomy, artificial cognition, and its physical embodiment, an intelligent robot is likely perceived as an independent entity and attributed with intentions and agency (Broadbent, 2017; Hancock et al., 2011; Young et al., 2011).

Human-robot interaction is defined as the process of a human and a robot working together to accomplish a common goal (Goodrich & Schultz, 2007). The literature describes two forms of HRI: cooperation and collaboration (Onnasch et al., 2016). Cooperation describes an interaction with a common goal, but tasks that are clearly divided between employee and robot. Employee and robot work independently, each on their allocated tasks, to achieve the common goal. Collaboration describes an interaction where employee and robot share tasks, work interdependently and use synergies to achieve the common goal. In the context of work, for both forms of HRI, the common goal is to successfully execute job assignments. Successful HRI in this sense therefore describes the attainment of task-related goals. Numerous research approaches and disciplines have provided indicators of successful

HRI, among them intuitive interaction research (Blackler, Desai et al., 2018), the technology acceptance model (Venkatesh & Bala, 2008), cognitive engineering (Parasuraman et al., 2008), or user-centred design (Norman, 1988). They approach successful HRI from different theoretical perspectives, but identify a range of similar indicators. Taken together, these approaches describe behavioural (task performance and cooperation with the robot), attitudinal (satisfaction, acceptance, and trust), and cognitive (mental workload and situation awareness) indicators of successful HRI.

Designing Robots to Enable Successful HRI

One essential enabler of successful HRI at work is task-related communication between robot and employee (Fong et al., 2003), which can occur in explicit (verbal, written, or through signals and gestures) and implicit form (through motion, behaviour, form, and appearance). Explicit communication can be facilitated through the design of the robot interface. Via the interface, information is provided by the employee to the robot (e.g., input to control the robot), and vice versa (e.g., about the current status of the robot) (Goodrich & Schultz, 2007). The two interface design features assumed most influential for successful HRI are visibility and feedback (Blackler, Desai et al., 2018; Norman, 1988). Visibility refers to the action possibilities (so-called affordances) the interface has to offer (Maier & Fadel, 2009; You & Chen, 2007). One example of high affordance visibility is the use of joysticks that are pushed forward to move a robot forward (Adamides et al., 2017). Feedback describes continuous, sufficient, and useful information about the results of actions, the robot's internal states or its environment (Hartson, 2003). This can be visual (Chen et al., 2014), haptic (Diaz et al., 2014), or auditory feedback (Mavridis et al., 2015).

Implicit communication can be facilitated through the design of the robot's controller and its appearance. The controller represents the algorithms and software that operate a robot and that allow the employee to monitor or control the robot's movements and communication

(International Organization for Standardization, 2012). High adaptability and autonomy are assumed as the most important design features of the controller (Beer et al., 2014; Graaf & Ben Allouch, 2013; Heerink et al., 2010). Adaptability is defined as the controller's ability to adapt to the changing needs of the user, in the sense of personalization (Graaf & Ben Allouch, 2013). This can be, for instance, adapting stiffness of the robot's joints according to user needs in a certain task (Duchaine et al., 2012; Gopinathan et al., 2017; Muxfeldt et al., 2017), or proactively selecting tasks through anticipation of user intent (Hoffman & Breazeal, 2007; Huang & Mutlu, 2016). Autonomy of the robot describes its ability to perform tasks without human intervention (International Organization for Standardization, 2012). High degrees of autonomy of the robot (Manzey et al., 2012) can be described as higher levels of robot responsibility (Sheridan & Verplank, 1978) as well as higher stages of information processing (Parasuraman et al., 2000).

The appearance is the feature that makes the first impression, through the shape and behaviour of the robot. The most important feature of robot appearance is assumed to be human likeness (Broadbent et al., 2009). Human likeness of a robot, in shape or in behaviour (e.g., human-like body, speech, or movements), is related to an attribution of human characteristics, such as social agency, intentions, or mental states (Duffy, 2003). This tendency, called anthropomorphism, is believed to lead to more successful interactions with robots through an increased familiarity (Duffy, 2003; Fink, 2012).

These robot design features and their effects on individual indicators of successful HRI have been investigated in many experiments. However, these experiments often only investigated individual features in specific contexts and with specific tasks. In order to identify which robot design feature has the strongest effects on successful HRI, meta-analytical research is needed to summarize and compare overall and specific effects of each design feature. Meta-analytical research can provide reliable information for researchers in search of explanations for how, why and when HRI is successful and guidelines for practitioners in search of best-practice design and implementation of robots for successful interactions at work.

Intelligent Technologies as Supervisors

Advances in AI development have brought about the possibility of intelligent technologies that no longer work alongside employees but function in a supervisory role (Wesche & Sonderegger, 2019; Langer et al., in press). These technologies already do and increasingly will make decisions and lead employees, i.e., allocate shifts or tasks, set the working pace, decide about recruitment of new employees or development of existing personnel (e.g., Machado et al., 2016; Mlekus et al., 2019; Naim et al., 2016). These decisions have a major impact on employees' everyday work, their daily routines, and even their careers and therefore carry the risk to reduce employees' satisfaction, motivation, or performance (e.g., Truxillo et al., 2017; Wolbeck, 2019). In order to avert this risk and facilitate beneficial attitudes and behaviour towards the organisation, an incorporation of principles of organisational justice has shown to be effective (Colquitt et al., 2013; Phillips, 2002; Wolbeck, 2019).

Justice in Organisations

The perception of fair treatment in a decision situation is contingent on the principles of organisational justice (Colquitt & Zipay, 2015; Greenberg, 2011). The concept of organisational justice is composed of four dimensions (Colquitt & Rodell, 2015): distributive, procedural, interpersonal, and informational justice. Distributive justice describes that a decision outcome is perceived as fair if certain allocation principles are used. Procedural justice describes that a decision-making procedure is perceived as fair if the recipients have a voice during the process or influence over the outcome, and if justice criteria (such as consistency, lack of bias, accuracy, correctability, and ethicality) are considered.

Interpersonal justice describes that the decision agent's behaviour toward the recipient is perceived as fair if it is respectful, polite, and dignified. Finally, informational justice describes that the information used to explain how the decision was formed is perceived as fair if the information is adequate, truthful, well-reasoned, specific, and timely.

Organisational justice of a decision enhances beneficial attitudes and behaviour, such as job satisfaction, organisational commitment, cooperation, or organisational citizenship behaviour, and reduces potentially harmful behaviour, such as counterproductive work behaviour (Cohen-Charash & Spector, 2001; Colquitt et al., 2013; Cremer & Tyler, 2005). The question of how the perception of fairness shapes employee attitudes and behaviour has been researched in a multitude of theories that propose possible explaining mechanisms for justice effects. The four most prominent mediators in these theories are positive and negative affect, and trust in the supervisor and identification with the work group (Colquitt & Zipay, 2015). Positive and negative affect describe independent dimensions of subjective feeling states (Watson, 2000). Several theories propose affect as a mediator of justice effects (e.g., appraisal theories, uncertainty management theory or affective events theory; for an overview see Cropanzano et al., 2020) and describe positive affect as reaction to fair decisions and negative affect as reaction to unfair decisions. Trust in the supervisor is the most common indicator of social exchange quality (Colquitt et al., 2014). Social exchange theory describes interactions as social exchanges; one interaction partner offers a certain benefit in exchange for reciprocation from the other (Blau, 1964). Just decisions can cause deeper relationships, with the result that employees are more likely to reciprocate (Organ, 1990). Identification with the group is proposed as a mediator in the group engagement model (Tyler & Blader, 2003). When employees are treated in a just manner, they will feel respected because they are proud to belong to a group that treats others fairly. They subsequently develop a stronger identification with the group, and consequently show more cooperative behaviour (e.g., organisational citizenship behaviour).

All four mediators individually have been confirmed in empirical research (e.g., Jiang et al., 2017; Khan et al., 2013; Soenen & Melkonian, 2017). However, even though organisational justice researchers have urgently called for parallel investigation of mediators from different theoretical perspectives (Colquitt et al., 2013), experimental studies that investigate two or more mediators are exceptionally rare and do not allow for a direct comparison of mediator strength (e.g., Chen et al., 2015).

Intelligent Technologies as Decision Agents

Even though artificial intelligence is emerging as new a decision agent in organisations, AI research and organisational justice research are seldom connected. In research and application of AI, some components of organisational justice have already been recognized, even though the term justice is not explicitly used. There are intelligent algorithms that allocate goods in a mathematically just way (Goldman & Procaccia, 2015; Lee & Baykal, 2017) or apply justice principles to task allocation in human–robot collaboration teams (El Mesbahi et al., 2014). There are intelligent robots that communicate in a polite and respectful manner (Fussell et al., 2008), communicate reasons for algorithmic decisions (Muggleton et al., 2018), or start to include morale and ethics into decisions (Kahn et al., 2013; Wallach, 2010). Whereas distributive, interpersonal and informational justice and even ethics have been considered, justice of the actual decision process has often been left out (Robert et al., 2020). This is surprising because procedural justice not only positively influences employee attitudes and behaviour (Cohen-Charash & Spector, 2001; Colquitt et al., 2013), employees are also less affected by unfavourable outcomes when the decision procedures are perceived to be fair (Brockner et al., 2009; Brockner & Wiesenfeld, 2005).

Organisational justice research only recently recognized the inclusion of AI as decision (or justice) agents as a major topic in future organisational justice research (Brockner & Wiesenfeld, 2020). However, although there are various theories that explain how procedural justice affects employee attitudes and behaviour (Colquitt & Zipay, 2015), and some even differentiate between organisation and supervisor as decision agents (Blader & Tyler, 2003; Rupp et al., 2014), researchers still know little about intelligent technologies as decision agents or even how specific characteristics of decision agents influence justice effects. In a literature review, Marques and colleagues (2017) summarised research on the impact of human decision agents for justice effects, showing a surprising lack of empirical studies. The few empirical studies that have investigated characteristics of human decision agents on procedural justice effects have focused on leadership behaviour, such as transformational leadership or passion (Cremer, 2006; Cremer & den Ouden, 2009) or leader–follower similarity (Cornelis et al., 2011). Only a few studies directly compared how employees react to human versus AI decisions (e.g., Harriott et al., 2013; Hinds et al., 2004; Lee & Baykal, 2017). However they did not investigate perceptions of fairness. Those who did investigate fairness perceptions, did not investigate employees or the effects of justice on employee attitudes and behaviour (Marcinkowski et al., 2020; Schlicker et al., 2019).

Differences Between Intelligent Technologies and Humans as Just Decision Agents

Research on the interactions between humans and intelligent technologies shows a grave ambiguity concerning the question of whether there are differences human to human and human to AI interactions. On the one hand, it can be argued that procedural justice is equally important for employee attitudes and behaviour in situations with AI as decision agents as in situations with human decision agents. This reasoning is mainly based on the Computers-Are-Social-Actors theory (Nass & Moon, 2000), which established the hypothesis that humans show similar social reactions to actions made by a human or a computer. The theory originally stems from social psychology, where a number of social interaction phenomena (such as applying gender biases or engaging in polite behaviour: Nass et al., 1997; Nass et al., 1999) could be transferred to human–computer interactions. It is widely used (e.g., Edwards et al., 2014; Lee et al., 2006), however, it developed and had its peak before intelligent technologies evolved.

On the other hand, research suggests that there are differences in how decisions made by humans and those made by AIs influence justice effects and how these are transferred. For example, a vignette study with students showed that the effect of procedural justice on the perception of the university's reputation differed significantly between human and AI decision agents (Marcinkowski et al., 2020). In addition, research has shown differences in emotional reactions (especially negative affect) and trust towards human interaction partners in comparison to machine-like interaction partners (Visser et al., 2016; Walter et al., 2014). One study could show that there are, however, fewer differences when the machine is more human-like (Kulms & Kopp, 2019).

Organisational justice theories propose three purposes of justice that determine the strength of its effects and explain why these effects might differ between human and AI as decision agents (Cropanzano et al., 2001). First, justice can serve instrumental purposes because it assures employees that their behaviour in the organisation will result in the reciprocation of benefits for them. AIs as interaction partners were shown to cause less reciprocation than human interaction partners (Lee & Liang, 2015). They might therefore not provide enough assurance to be able to provide future benefits and procedural justice effects are diminished. Second, justice can serve interpersonal purposes because it indicates inclusion and status in the group, fostering identification with the group, which then causes beneficial behaviour towards the organisation. Humans were shown to indicate less identification with the group in interactions with AIs than in interactions with humans (Peña et al., 2019). Therefore procedural justice may lead to less beneficial behaviour when AIs are the decision agent. Finally, justice can serve moral purposes because people generally seek out moral behaviour and decisions. Research has shown that people expect higher moral standards from robots than from humans (Voiklis et al., 2016). Procedural justice effects might therefore be diminished when an AI makes a decision because they might be perceived as less moral compared to human decision agents.

Further research additionally suggests that there might not only be differences in how strong justice effects are but also in how justice effects are mediated. Organisational justice research as well as research on AI as decision agent have proposed judgements of responsibility as explaining mechanism for differences in how justice effects are mediated. Fairness theory (Folger & Cropanzano, 2001) focuses on how the responsibility attributed to the decision agent affects fairness perceptions and effects. Three elements are central to an attribution of responsibility. First, an aversive state occurs, which raises the question of *would* an alternative state have felt different, if the decision had been different. Second, the discretionary conduct of a person is considered, which raises the question of *could* the person have acted differently and therefore have caused a different outcome. Finally, moral principles are judged, which raises the question of *should* the person have acted differently. In research on AI decision agents, the concept of judgements of responsibility is applied using attributional theory (e.g., Weiner, 1995). According to attributional theory, judgements of responsibility are formed through the assessment of causal dimensions such as intentionality and controllability (Weiner, 1995, 2006). Intentionality refers to whether the decision was made purposefully or unintended. Controllability refers to whether the decision was preventable or inevitable. These judgements of responsibility form an employee's reaction towards the AI and its decisions (Britt & Garrity, 2006; Wickens et al., 2011).

Applied to the four most common mediators of procedural justice effects (positive and negative affect, trust, and identification), differences can be assumed as described in the following. First, positive affective reactions are more likely to occur in a just decision. In a situation with a just decision, affective reactions are more likely based on the event, not the agent (Malle & Scheutz, 2014), and different agents should therefore not influence justice effects. Second, negative affective reactions are more likely to occur in unfair decisions and are followed by an investigation of the agent's responsibility (Folger & Cropanzano, 2001; Weiner, 1995). Negative affective reactions only occur when harm has been done, a moral norm has been violated, or when there was an intention to harm (Cropanzano et al., 2000). Because an AI is not likely to be held morally responsible or accountable (Voiklis et al., 2016), negative affective reactions might be less likely to occur in a situation with an AI as a decision agent (van der Woerdt & Haselager, 2017, 2019). Third, trust in a reciprocation from a decision agent cannot develop when intentionality and control are lacking. When beneficial behaviour seems unintended or even coincidental, it is less likely to form trust that the decision agent will reciprocate this behaviour (Blau, 1964). Empirical studies show that AI decision agents are less likely attributed with an intention to harm or benefit someone (Voiklis et al., 2016; Xie et al., 2019). Therefore, trust in an AI decision agent might develop less likely. Finally, research on identification and group membership showed that intelligent technologies can be perceived as legitimate group members (Häring et al., 2014; Kuchenbrandt et al., 2013). However, humans are more likely to identify with the group in interactions with human partners than in interactions with artificial partners (Peña et al., 2019). It might therefore be likely that employees identify less with a group where an AI is the decision agent than with a group where a human makes the decisions.

In summary, employees' negative affective reaction, their trust in the decision agent and their identification with the work group are expected to differ depending on the procedural justice of human or AI decisions.

Aims and Outline of the Present Work

By addressing the important role of intelligent technologies as potential co-workers and supervisors in the workplace, the present work contributes to the advancement of interdisciplinary literature linking HRI, AI, and organisational justice. As such, this dissertation has two aims. The first aim is to advance knowledge on intelligent technologies in the role of co-workers by investigating whether and how different robot design features influence the success of human-robot interactions at work (Study 1). The second aim is to advance knowledge on characteristics of the decision agent in organisational justice literature by investigating how effects and explaining mechanisms of procedurally just decisions differ between humans and AI as supervisors (Studies 2 and 3).

The first research aim is based on the assumption that an intelligent robot is a good co-worker when its design enhances successful HRI, which means that employee and robot successfully achieve task-related goals. In Study 1, we gave an overview on design features of a robot's interface, controller and appearance that are assumed to facilitate communication between employee and robot and hence contribute to successful HRI. Beyond that, we meta-analytically integrated experimental studies that investigate effects of individual robot design features and compared their overall strength as well as identified specific effects on the indicators of successful HRI.

The second research aim is based on the observation that intelligent technologies will increasingly often make organisational decisions and the assumption that with this emergence of new decision agents, principles of procedural justice need a reassessment. Previous research either does not investigate differences between human and AI decision agents or yields highly ambiguous results. Therefore, Study 2 and 3 investigate more closely whether procedural justice effects differ between decision agents. Specifically, in Study 2 we investigated differences concerning direct effects of procedural justice on employee attitudes and behaviour in two common decision situations in organisations. Study 3 then builds on these results, as we investigated differences between decision agents concerning indirect effects of justice on employee attitudes and behaviour as well as possible explaining mechanisms.

Study 1 – Let's Work Together: A Meta-Analysis on Robot Design Features that Enable Successful Human–Robot Interaction at Work

In the light of workplaces where employees increasingly often share workspace and tasks with an intelligent robot, the aim of Study 1 was to explore which robot design features make human-robot interactions in the workplace most successful. To this end, we systematically searched studies on the influence of robot design features on successful HRI from engineering and information science (IEEE Xplore and ACM Digital Library) as well as psychological (PsychInfo, Web of Science) search engines. The systematic selection of literature and a structured coding procedure led to 81 included studies, containing 380 effect sizes. Mean effects were calculated using three-level meta-analysis (Cheung, 2014, 2015) to handle dependencies of multiple effect sizes in one study (Cheung & Chan, 2004). We calculated mean effect sizes for the relation of each robot design feature (i.e., features of interface, controller, and appearance) with successful HRI in general as well as moderation analyses for specific effects of each design feature on each of the indicators of successful HRI (i.e., task performance, cooperation, satisfaction, acceptance, trust, mental workload, and situation awareness).

The meta-analytical results showed that sufficient feedback through the interface, clear visibility of affordances, and adaptability and autonomy of the controller positively affected successful HRI with medium-sized effects ($d = .50$, $p = .003$; $d = .44$, $p < .001$; $d = .48$, $p = .005$; $d = .58$, $p = .011$, respectively). Appearance did not have a significant effect on successful HRI ($d = .24$, $p = .579$).

The moderation analyses revealed that certain indicators of successful HRI were influenced by all design features, whereas others were influenced only by specific design features. All four design features of interface and controller positively affected task performance and user satisfaction. For task performance, all features had comparable medium-sized effects ($d = .44$ to $.62$). For user satisfaction, feedback had a large effect ($d = 1.38$, $p = .003$), whereas visibility and adaptability had comparable medium-sized effects ($d = .60$, $p = .018$; $d = .73$, $p < .001$). Autonomy was only represented by one effect size and did not show a significant effect. We could show specific effects, besides those on performance and satisfaction, only for the design features of the controller. Adaptability had an additional effect on cooperation ($d = 0.83$, $p = .006$) and acceptance ($d = 0.86$, $p < .001$). Autonomy was the only indicator with a significant effect on mental workload ($d = 2.17$, $p < .001$). Here, it has to be noted that mental workload was reverse coded during all analyses, so that for all outcomes high values signify the desirable direction. Lastly, none of the design features had an effect on trust or situation awareness.

These meta-analytical results show that robot design at work needs to consider multiple features of interface and controller to achieve successful HRI that covers not only task performance and satisfaction, but also cooperation, acceptance, and mental workload. Additionally, the results revealed the need for further empirical research because not all assumed relationships between robot design features and successful HRI could be tested (e.g., specific effects of human likeness) and some heterogeneity and unexplained variance remained.

Study 2 – The Importance of Procedural Justice in Human–Machine Interactions: Intelligent Systems as New Decision Agents in Organizations

Study 2 takes the focus from intelligent robots as co-workers that interact side-by-side with employees to AI as supervisor that makes decisions for or about employees. Even though decades of justice research investigated procedural justice and its effects (Brockner & Wiesenfeld, 2020; Greenberg, 1987), these insights might not be easily transferred to AI decision making. The aim of Study 2 therefore was to explore direct effects of procedural justice and their interaction with the type of decision agent (human vs. AI, the latter in appearance of tablet computer or robot) on employee attitudes and behaviour. Specifically, we predicted that the type of decision agent would moderate the relationship between procedural justice and employee attitudes and behaviour, with the relationship being strongest when the decision agent is a human team leader, medium when the decision agent is a humanoid robot, and weakest when the agent is a computer system. The hypotheses were investigated using a between-subjects design in two online experimental vignette studies ($N_1 = 149$ and $N_2 = 145$) that described two common decision situations in organisations (i.e., the allocation of new tasks and the allocation of further vocational training). Hypotheses were tested using multiple regression analyses with contrast coding of the experimental manipulation and moderation analysis (Cohen et al., 2003; Hayes, 2018).

Results of both samples showed significant effects of procedural justice on job satisfaction ($B = .97$ and $.88$, $p < .001$), commitment ($B = .38$ and $.30$, $p < .001$), cooperation ($B = .49$ and $.43$, $p < .001$), organisational citizenship behaviour ($B = .17$ and $.13$, $p < .001$ and $p = .002$), and counterproductive work behaviour ($B = -.20$ and $.10$, $p < .001$ and $p = .041$), regardless of which decision agent made the decisions. The effect sizes are comparable to meta-analytical findings on procedural justice effects (Colquitt et al., 2001; Colquitt et al., 2013). These results confirm the pivotal importance of procedural justice in the workplace for both human and AI decision agents and further emphasise the importance of designing AI that is capable of making procedurally just decisions.

In addition to these main results, a difference between decision situations became obvious. All effect sizes were larger for the task allocation decision than for the allocation of vocational training. As procedural justice was more important for employee attitudes and behaviour in a situation with a decision concerning the allocation of tasks, Study 3 used this scenario to further investigate procedural justice effects in the context of AI as decision agent.

Study 3 – How Procedural Justice Works: Artificial Intelligence as New Decision Agent and the Mediation of Justice Effects

Study 3 built on the results of Study 2 and aimed to further extend the investigation of possible differences between human and AI decision agents. Justice research, over the decades, developed various theories on how just decisions affect employees. These explaining mechanisms are seldom compared concerning strength and transmitting effects to specific attitudes and behaviour, and again need reassessment in the new context of AI decisions. In Study 3, we therefore compared the strength and specificity of the mediators affect, trust, and identification for procedural justice effects on employee attitudes and behaviour and investigated differences between human and AI decision agents (the latter in appearance of a human-like and a machine-like robot). In addition, we examined responsibility as explaining mechanism for these differences. We manipulated procedural justice and type of decision agent in an experimental vignette in two samples of 229 and 132 employees. The second sample was used to replicate and extend the main results from the first sample. Hypotheses were tested using multiple regression analyses with contrast coding of the experimental manipulation, mediation analysis with parallel mediators and moderated mediation analysis (Cohen et al., 2003; Hayes, 2018).

In both experiments, the parallel mediation analysis revealed that trust was the strongest mediator for the effect of procedural justice on attitudes (i.e., job satisfaction and commitment) and negative affect was the strongest mediator for the effect of procedural justice on behaviour (i.e., organisational citizenship behaviour and counterproductive work behaviour). In addition, the index of moderated mediation showed that trust as mediator was less pronounced for AI decisions compared to human decisions (job satisfaction: $B_{experiment1}$ = .30, 95% $\text{CI}_{experiment1}$ = [.16, .46] and $B_{experiment2}$ = .17, $\text{CI}_{experiment2}$ = [.03, .35]; commitment: $B_{experiment1}$ = .12, $\text{CI}_{experiment1}$ = [.05, .20] and $B_{experiment2}$ = .11, $\text{CI}_{experiment2}$ = [.02, .23]), whereas no differences could be found for negative affect. There were no differences between the two AI decision agents (the human-like and the machine-like robot). Results concerning identification were ambiguous: Differences between decision agents could be shown only in the second, smaller, sample (commitment: B = .12, CI = [.02, .27]).

Additionally, in a further analysis we could confirm judgements of responsibility of a decision agent as underlying mechanism for differences found between human and AI decision agents. Employees perceive AIs as less intentional and in control of decisions and effects of procedural justice on negative affect, trust, and identification are less pronounced for AI decision agents. However, judgements of responsibility could not explain all differences between human and AI decision agents, this characteristic of a decision agent therefore proved useful to explain differences in how procedural justice affects employee attitudes and behaviour, but there have to be other important characteristics as well. Our experiments therefore show that there are differences between human and AI decision agents and some of this variance can be explained by judgements of responsibility, but future research needs to investigate characteristics of decision agents more closely.

General Discussion

The present work addresses the important role of intelligent technologies as potential co-workers and supervisors at work using a variety of robust research methods. In three studies, I investigated which robot design features make HRI at work most successful and how and why effects of procedural justice differ between humans and AI as decision agent. The present work therefore extends the understanding of human-AI interactions at work and with its interdisciplinary focus contributes to the advancement of the fields of HRI, AI, work design, and organisational justice.

In Study 1, we focussed on AI as co-worker and we meta-analytically investigated which robot design features most strongly contributed to the success of HRI at work and whether there are specific effects of the design features on individual indicators of successful HRI. The results of this meta-analysis revealed that feedback and visibility of the interface and adaptability and autonomy of the controller had comparable medium-sized effects on successful HRI at work, whereas human likeness did not have a significant effect. The moderation analyses revealed that certain indicators of successful HRI were influenced by all design features, whereas others were only influenced by specific design features. As assumed, all four design features of interface and controller positively affected task performance and—with the exception of autonomy—user satisfaction. With regard to task performance, all features had comparable medium-sized effects. With regard to user satisfaction, feedback had a large effect and visibility and adaptability had comparable medium-sized effects, whereas autonomy was only represented by one effect size that did not show a significant effect. In summary, all four features of interface and controller are reasonably good design choices when aiming for improved performance, whereas feedback might be preferable in order to benefit user satisfaction.

Beyond the effects on performance and satisfaction, we could not find any specific effects for the design features of the interface. The assumed effects of visibility on mental workload and of feedback on acceptance and situation awareness could not be confirmed. However, the results showed specific effects beyond those on performance and satisfaction for the design features of the controller. Adaptability had an additional effect on cooperation and acceptance. Among the investigated features, it therefore showed the broadest effects. However, some of these effects need to be interpreted with caution because they are based on a small number of effect sizes (e.g., cooperation or trust). Autonomy had an additional effect on mental workload. Most theories assume that good robot design generally reduces mental workload (e.g., Blackler, Popovic, & Desai, 2018; Onnasch et al., 2014). Yet, autonomy is the one design feature that specifically aims at reducing mental workload, especially in task-related interactions (Breazeal, 2004), which could now be confirmed by our analyses.

In Study 2, we shifted the focus from AI as co-worker on AI as supervisor and we investigated whether procedural justice perceptions and their effect on employee attitudes and behaviour differ between humans and AI as decision agents. The results from both experiments in Study 2 consistently showed significant effects of procedural justice on employee attitudes and behaviour, independent of the type of decision agent. The effect sizes of these relationships were comparable to previous meta-analytical findings on justice effects (Colquitt et al., 2001; Colquitt et al., 2013). Additionally, the manipulation checks showed that even the perception of procedural justice did not differ between decision agents. Together, this shows that neither perceptions of procedural justice nor their effects vary depending on the type of decision agent. This demonstrates that the established importance of procedural justice in the workplace not only applies to interactions with human supervisors but with intelligent technology as supervisor as well.

In Study 3, our aim was to further investigate the mechanisms of procedural justice effects. Even though, in Study 2, we could demonstrate that procedural justice is important for the effects of both human and AI decisions, research suggested differences in how these effects are transmitted. Therefore, we investigated whether the type of decision agent (human vs. AI) makes a difference concerning mediating mechanisms in procedural justice effects by investigating differences in strength and specificity of mediators from prominent justice theories. Additionally, we examined whether judgements of responsibility might be one characteristic of a decision agent that explains differences in justice effects. Both experiments in Study 3 showed that the effects of procedural justice on attitudes and behaviour are explained by specific mediators. Overall, with respect to attitudes, trust was the strongest mediator; whereas regarding behaviours, negative affect was the strongest mediator. Concerning differences between human and AI decision agents, trust was a weaker mediator for AI decision agents; whereas no differences occurred for negative affect as mediator. Finally, intentionality and controllability of a decision (the two facets of judgements of responsibility) were confirmed as underlying mechanisms for differences between human and AI decision agents. The employees perceive AIs as less intentional and in control of decisions. Therefore, effects of procedural justice on negative affect, identification, and trust are less pronounced.

Theoretical Implications

Several results obtained in the present work emphasise the importance of considering the context when investigating interactions between AI and employees at work. With regard to AI as co-worker, the focus on task-related interactions at work and the type of task are contextual factors that might explain unexpected nonsignificant results in our meta-analysis. With regard to AI as supervisor, our experiments showed that the type of decision and the characteristics of a decision agent are contextual factors that might explain differences in procedural justice effects. All four contextual factors will be described in the following.

First, with regard to the focus on task-related interactions at work, the studies included in our meta-analysis focus on task-related interactions, such as navigating, manipulating, or cooperatively solving a task with a robot because we investigated interactions at work. However, previous research on the effects of human likeness of robots was mostly conducted in the context of social interactions (Duffy, 2003; Fink, 2012; Gong, 2008), such as making conversation. A human-like appearance was shown to be preferred if it matched the sociability required in a job (Goetz et al., 2003). Therefore, human likeness might not be an important influencing factor for the success of task-related interactions at work. Further support for this argument is provided by the results of Studies 2 and 3: In all four experiments, human likeness of an intelligent technology did not affect employee attitudes and behaviour. Still, as the results of the meta-analysis are based on a rather limited number of effect sizes and our experiments were among the first to investigate differences between decision agents, more research is needed to clarify whether and when human likeness influences successful HRI.

Second, with regard to the type of task that is approached by employee and robot, the studies included in our meta-analysis investigate HRI at work in various settings and for a range of different tasks. Both the effects of interface features and the effects on trust and situation awareness (where we unexpectedly obtained nonsignificant results) have been shown to be highly sensitive to task characteristics: The framing of a task in the organisational context is an important factor influencing trust (e.g., Hoff & Bashir, 2015), and task quantity and complexity strongly influence situation awareness (e.g., Endsley, 2000) and the effects of interface features such as feedback (e.g., Burke et al., 2006). Additionally, the studies included in our meta-analysis mostly investigated participants with no or limited experience with the technologies used. This novelty of a task that needs to be accomplished in interaction with an intelligent technology might be responsible for nonsignificant effects of visibility on mental workload and of feedback on situation awareness. If a task requires employees to perform many unfamiliar actions, the benefits of familiarity through visible affordances might just not come into effect and the benefit of detailed feedback about robot states and environment might be cancelled out by the additional demand to process this information.

Third, with regard to the type of decision, the results from Study 2 showed that the effects of procedural justice were larger in the context of a decision about the allocation of tasks than in the context of a decision about the allocation of further vocational training. A theoretical explanation for these differences might lie in the level of abstraction (concrete or abstract) and psychological distance (proximal or distal), which determine the impact an environment has on employees (Lewin, 1943; Soderberg et al., 2015). A proximal and concrete environment or situation has a higher impact on employee attitudes and behaviour than a distal and abstract situation (Becker, 2012). Differences in procedural justice effects might therefore occur because the allocation of tasks might be perceived as more concrete due to a higher contiguity to the actual work task and environment, and more proximal because changes in tasks are more present and immediate than a future, single training would be.

Fourth, with regard to the characteristics of a decision agent, the result from Study 3 could show that whether a decision agent is judged personally responsible for a decision significantly influences how procedural justice effects affect employee attitudes and behaviour. Even though there are justice theories that differentiate between organisation and supervisor as decision agents (Blader & Tyler, 2003; Rupp et al., 2014), there is a surprising lack of studies that investigate how characteristics of a decision agent influence justice effects

(Marques et al., 2017). The present work is one of the first to examine differences between human and AI as decision agent concerning justice perceptions and effects (for two recent exceptions see Marcinkowski et al., 2020; Schlicker et al., 2019) and, to our knowledge, the first to derive specific characteristics of a decision agent (i.e., intentionality and controllability) from this comparison that influence procedural justice effects. These characteristics not only explain differences between human and AI decision agents, they might as well explain differences between other sources of justice and therefore improve the ability to predict employee reactions to procedural justice (Cojuharenco et al., 2017; Rupp et al., 2014).

The present work also has several implications for justice literature. First, it comprises studies that are among the first to experimentally compare mediators from social exchange theory, affective events theory and the group engagement model, three of the most prominent justice theories. These comparisons provide important insights on differential effects of these mediators. Most justice theories propose explanations for justice effects explicitly on behavioural outcomes; for example, exchange behaviour such as OCB in social exchange theory (Blau, 1964) or cooperation in the group engagement model (Tyler & Blader, 2003). They often simply assume that the same holds true for effects on attitudes. Our results indicate that some mediators are better suited to explain the effects of procedural justice on attitudes than those on behaviour. A much more differentiated investigation of mediators and outcomes is needed.

Second, our experimental studies showed that the justice of a decision process is an important influencing factor for employee attitudes and behaviour even when intelligent technologies make decisions. Our experiments also showed that characteristics of a decision agent can influence how procedural justice affects employee attitudes and behaviour. However, justice theories account for this in an insufficient manner. Empirical studies researching the source of justice are underrepresented and existing theories do not incorporate characteristics of a decision agent that might be responsible for differences in how justice works (Rupp et al., 2014). In the present work, I therefore investigated intentionality and controllability, in order to confirm that these characteristics can explain differences between agents. The results of Study 3 showed that it is of particular importance for a decision agent to be perceived as having made the decision intentionally. Therefore, including characteristics of the decision agent, such as intentionality, as moderator into justice theories will enrich justice literature.

Practical Implications

The results obtained in the present work have several implications for organisations that aim to implement intelligent technology and for designers of these technologies who want to ensure successful and fair interactions.

Robot designers and organisations that want to apply an intelligent robot working side-by-side with employees should carefully choose design features that fit the targeted work context. The most versatile design feature investigated in our meta-analysis is adaptability in the sense of a personalization to the employee's needs. It positively affects performance, cooperation, satisfaction, and acceptance. Yet, in our meta-analysis, we could not show significant effects of adaptability on cognitive indicators of successful HRI (i.e., mental workload and situation awareness). The only feature that effectively reduces mental workload, according to our results, is autonomy. Yet, more autonomy does not necessarily influence positive attitudes towards the robot. Robot designers should carefully consider and balance these specific effects with requirements and implementation costs in the specific context, to find the best possible fit.

Organisations should make sure that any decision (made by human or AI supervisors) is made through just procedures. With regard to human supervisors, this can be achieved through justice training (Richter et al., 2016; Skarlicki & Latham, 2005; Whitman et al., 2012). With regard to AI supervisors, it is important to ensure that engineers and designers have the necessary abilities to design just decisions. Procedural justice training is needed that is adapted to this particular target group and the special context of AI design and that is evaluated in intervention studies. In addition, Study 3 showed that the decision agent has to convey an impression of purposefully intending a decision. To achieve this, the communication of decisions needs to transparently and clearly describe who made the decision. Therefore, organisational representatives should carefully consider how decision procedures and their communication are designed, especially before deciding which AI to implement as a decision agent.

AI designers have to factor in justice rules early in the AI design process because the implementation of just procedures into AI decision making has to consider more than simply writing text modules used to communicate a decision. Designers need to find ways to enable an AI to consider employee opinions and preferences and have to make sure that the decisions adhere to established procedural justice rules such as being unbiased, correctable, accurate, consistent, and explainable (Colquitt, 2001). Töniges et al. (2017) already proposed several recommendations for the implementation of justice rules in intelligent technologies. With regard to procedural justice rules, this includes that the AI needs to present a means for the user to raise objections to the decision process, to make suggestions for improvements, or to make corrections. This could, for example, be ensured by providing natural, conversation-like dialogues. These recommendations should be further refined and tested in practical use.

Strengths and Limitations

The present work is characterised by certain strengths as well as limitations, which will be addressed in the following.

A major strength is the combination of multiple perspectives on AI in the workplace. I integrated research from the fields of psychology as well as engineering and information science by drawing on theories from work design, organisational justice, HRI, and AI literature and, in our meta-analysis, by using scientific search engines from all research fields. By doing this, I provided interdisciplinary research that is urgently needed to investigate complex situations in organisations that touch multiple disciplines (Rhoten & Parker, 2004; Zhu & Fu, 2019).

A second major strength is that I examined AI in two highly important roles in the workplace by investigating AI as robotic co-worker and AI as algorithmic supervisor. Often, research engaging in interactions (or relationships) at work focus solely on hierarchical interactions between supervisor and employees and largely underestimates the important role of interactions between co-workers (Basford & Offermann, 2012). In the present work, I therefore provided insights into several requirements intelligent technologies should meet when they are implemented as both employees' co-workers and supervisors.

A further strength is the application of robust and advanced research methods. First, I used advanced meta-analytical methods to appropriately account for dependencies in the primary studies (Cheung, 2015). Doing this, I could prevent a major loss of information and an underestimation of the degree of heterogeneity that would result from the use of usual strategies to deal with these dependencies. Second, I used experimental vignette methods to balance the benefits of a controlled experimental environment with a realistic reflection of a situation that is still rarely found in organisational everyday life (Aguinis & Bradley, 2014). Choosing a correlational design instead would have had the disadvantage not to be able to eliminate alternative influencing factors. A further benefit of this approach is that it allows for the controlled investigation of technology that is as yet rarely applied in organisations. With the increasing advancement of intelligent technologies, new abilities emerge at very short notice. It is therefore important to continuously examine new abilities and roles and their impact on employees, their workplace, and organisations early because it is of vital importance for AI designers to gain access to this kind of information early in the development phase. Third, replication of the results of the vignette studies showed their reproducibility and heeded the call for more replications in IO-Psychology (Kepes & McDaniel, 2013), which allowed for a greater confidence in the results.

The present work has several limitations that need to be addressed. First, although we could include a large number of studies in our meta-analysis, some subsamples are based on a rather small number of studies or participants within studies, and may be subject to second-order sampling error (Hunter & Schmidt, 2015). Hence, some of the specific analyses may not provide reliable information, especially those that are based on very few effect sizes (e.g., the effect of autonomy on satisfaction or of visibility on situation awareness). Researchers and practitioners should therefore interpret these effects with caution even though meta-analytical results are still preferable to other methods that integrate research results because they allow for a quantification of effect sizes.

Second, the results of our meta-analysis revealed significant heterogeneity between studies. Even though splitting successful HRI into different indicators reduced heterogeneity and increased the amount of variance explained by the predictors, the meta-analysis is partly unable to explain why effects of robot design features differ between studies. This variation might be explained by differences between tasks (e.g., task difficulty: Gopinathan et al., 2017), or by individual differences (e.g., expertise: Hoff & Bashir, 2015). In this meta-analysis, we could not calculate moderation analyses for these factors because the original studies did not include the data needed to code the moderators. Future empirical research should therefore investigate these possible moderators or report necessary information to provide data for future meta-analyses.

Third, experimental research is often criticized as lacking external validity, ultimately compromising the generalizability of the results (Scandura & Williams, 2000). However, experimental research is essential in order to investigate causal relationships. Therefore, in Studies 2 and 3, I chose to utilise the experimental vignette methodology, which offers a unique way to retain experimental benefits while at the same time maximizing realism and external validity (Aguinis & Bradley, 2014). Experimental vignette studies are a common method in organisational justice and AI research (e.g., Dineen et al., 2004; Kwon et al., 2016; Zweig & Webster, 2002). To provide further realism, thoroughly designed vignettes were applied that incorporated visual material and vividly described detailed situations. Furthermore, I exclusively included employees in our samples and chose a context that is likely to be part of any of these employees' daily life (i.e., a weekly routine meeting).

Fourth, the imagined interaction in these experiments gives an impression of only one justice related event for which the participants rate the likeliness of certain attitudes and behaviour. In interactions over a longer period of time or a sequence of decisions, possible differences between decision agents might be clearer. First (unconscious) reactions to intelligent technologies might be different from reactions of employees after a longer time period. After some time, employees might become more aware of dissimilarities between humans and AI as supervisors. Additionally, organisational justice research shows that fairness perceptions can change over time (Konradt et al., 2016; Streicher et al., 2012). The experience of justice events changes and interacts with the global fairness perception of the entity—the decision agent (Jones & Skarlicki, 2012). This dynamic perspective on fairness perceptions is a rapidly developing but still under-researched perspective. Additional laboratory experiments are needed that cover longer interaction periods, maybe even with multiple decision situations, to distinguish effects of first mindless reactions and fairness perceptions that are formed over longer time periods. One possible method for these laboratory experiments are the so-called Wizard of Oz experiments, where the AI is remote controlled to simulate actual autonomous behaviour and reactions (Dahlbäck et al., 1993; Riek, 2012).

Lastly, the data in Studies 2 and 3 have been collected solely in Germany. Because Germany has specific characteristics both concerning culture (see Hofstede et al., 2010 for the cultural profile) and the labour market (e.g., works councils and employee organisation; Jenkins & Blyton, 2008) the studies' results might be constrained concerning their generalizability. With regard to the specific culture, a number of systematic reviews and meta-analyses investigated cultural differences concerning justice effects (Gelfand et al., 2007; Shao et al., 2012; Silva & Caetano, 2016). With regard to procedural justice, power distance seems to be important: Effects on attitudes and behaviour are higher in countries with low power distance (e.g., Germany) compared to countries with high power distance (e.g., China) (Gelfand et al., 2007; Shao et al., 2012). Therefore, our results should be replicated in countries with different cultural profiles. With regard to the characteristics of the German labour market, a correlational study using large-scale survey data found that the existence of works councils increases perceptions of wages as being fair (Pfeifer, 2014). Yet, empirical studies investigating the influence of specific characteristics of the German labour market are exceptionally rare. Therefore, future studies should further investigate the role of characteristics of the labour market for procedural justice effects.

Directions for Future Research

Even though our research on AI as co-workers could provide valuable insights on the effect of robot design features on indicators of successful HRI at work, it also revealed one major shortcoming of present HRI research: the need for conceptual and empirical research on explanations for these effects. Interlinking research fields more closely in the future could help to provide a more profound theoretical basis for further empirical research. In the following, I will briefly point out two possible starting points for this.

First, one theory that might be particularly useful for explaining effects of design features on successful HRI is the self-determination theory (SDT: Gagné & Deci, 2005; Ryan & Deci, 2000; Stone et al., 2009). SDT is one of the most broadly applied theories in the context of motivation at work, and has already been theoretically applied to human–technology interactions (Szalma, 2014). One tenet of the theory is that environmental factors (e.g., work design, or interactions) have the potential to satisfy basic needs (need for competence, need for autonomy, and need for relatedness), subsequently benefitting performance, satisfaction, and trust (Deci et al., 2017; Deci & Ryan, 2012). With increasing HRI at work, robot design features might become important environmental factors. The satisfaction of basic needs might, in turn, function as a mediating mechanism that explains the impact of robot design features on successful HRI (Szalma, 2014). For instance, a robot with high visibility of affordances might convey a sense of easy usage, satisfying the user's need to feel competent, therefore fostering their satisfaction and task performance.

Second, little is known about the interaction of different indicators of successful HRI, or whether some indicators might even act as mediating mechanisms for others. Some of the abovementioned theories have direct or indirect assumptions about this: Cognitive engineering closely links mental workload and situation awareness, with low to moderate mental workload leading to the highest awareness (Wickens et al., 2008). One of the main assumptions of intuitive interaction research is the effect of high familiarity on performance and satisfaction (Blackler et al., 2003), which is assumed to be mediated by mental workload (Blackler, Desai et al., 2018). Yet, empirical examination of these assumptions is rare and needs further development. Also, in order to determine a fully comprehensive picture of employee reactions, HRI research in the context of work can benefit from including additional indicators of successful HRI. Potential indicators could, for example, be derived from work design literature (e.g., work motivation: see Humphrey et al., 2007), or occupational health literature (e.g., work-related strain: see Ganster & Rosen, 2013; Griffin & Clarke, 2011).

The present research on AI as supervisor furthermore revealed several gaps in current research. First, even though I could show that intelligent technology is perceived as less intentional and in control of a decision, judgements of responsibility only partially explained the differences between decision agents. There likely are other characteristics of AI decision agents that can cause or compensate differences in how justice works. One possible explanation that future research could investigate is the attributed competence of the decision agent. An AI decision agent might not be perceived as competent to make a certain decision (in the sense of professional, social, and moral competence) as a human supervisor would be (Bartneck et al., 2009; Malle, 2016; Scheunemann et al., 2020). van der Woerdt and Haselager (2017), for example, showed that a perceived lack of ability of a robot reduced feelings of disappointment and blame towards the robot. Besides investigating the attribution of competence, it should be investigated whether employees ascribe certain characteristics to intelligent technologies (such as high consistency and lack of bias, but no possibility to voice opinions or concerns) or to human agents (such as high bias through sympathy, but the possibility to be convinced in a conversation), which might influence employee perceptions.

Second, future justice research needs to investigate what determines employees' perceptions of the decision agent's intentionality and control over a decision. Interdisciplinary research from the fields of philosophy and information science already started to address this topic (see e.g., Zhu, 2009; Ziemke et al., 2015), proposing, for example, autonomy and adaptability of the AI as factors that influence the perception of intentionality. However, these factors, their interplay, and their effects on employee perception have not yet been investigated empirically.

Finally, not only procedural justice but also the other justice dimensions should be investigated in the context of AI as decision agents. Especially informational justice poses challenges to the design of decision algorithms: It requires the communication and explanation of reasons for and rules of the decision (Greenberg, 1993). However, often decision algorithms are based on big databases and machine learning, and are unable to infer reasons or explain how the decision was made (for an exception, see: Schmid et al., 2017). Therefore, future research needs to either find ways for AI to provide these reasons and explanations or to compensate for the perceived injustice when informational justice rules are violated.

Conclusion

Intelligent robots and artificial decision agents are being developed at a high pace. With rising numbers of AI implementations at work, employees will experience interactions with AI co-workers and supervisors with increasing frequency. In order to enable organisations to make informed decisions about whether and how to implement intelligent technologies and designers and engineers to develop suitable AI, researchers need to investigate consequences and mechanisms of these interactions. The present work uses an interdisciplinary approach to close these knowledge gaps and provides vital insights into what is needed to create successful and fair interactions between AI and employees at work. These insights not only extend HRI, AI, work design, and organisational justice literature and provide important recommendations for AI designers, engineers and human resource practitioners. They also reveal essential paths for future research, which will help to understand interactions with AI as co-worker and supervisor even better and to benefit employees and organisations alike.

References

Adamides, G., Katsanos, C., Parmet, Y., Christou, G., Xenos, M., Hadzilacos, T., & Edan, Y. (2017). HRI usability evaluation of interaction modes for a teleoperated agricultural robotic sprayer. *Applied Ergonomics*, *62*, 237–246. https://doi.org/10.1016/j.apergo.2017.03.008

Aguinis, H., & Bradley, K. J. (2014). Best practice recommendations for designing and implementing experimental vignette methodology studies. *Organizational Research Methods*, *17*, 351–371. https://doi.org/10.1177/1094428114547952

Bartneck, C., Kulić, D., Croft, E. A., & Zoghbi, S. (2009). Measurement instruments for the anthropomorphism, animacy, likeability, perceived intelligence, and perceived safety of robots. *International Journal of Social Robotics*, *1*(1), 71–81. https://doi.org/10.1007/s12369-008-0001-3

Basford, T. E., & Offermann, L. R. (2012). Beyond leadership: The impact of coworker relationships on employee motivation and intent to stay. *Journal of Management & Organization*, *18*(6), 807–817. https://doi.org/10.5172/jmo.2012.18.6.807

Beer, J. M., Fisk, A. D., & Rogers, W. A. (2014). Toward a framework for levels of robot autonomy in human-robot interaction. *Journal of Human-Robot Interaction*, *3*(2), 74–99. https://doi.org/10.5898/JHRI.3.2.Beer

Bharadwaj, A., Parent-Rocheleau, X., Cascio, W. F., Bienefeld, N., & Kahlert, A. (2020). Work design, technology, and artificial intelligence: Keeping humans at the center of work. *Academy of Management Proceedings*, *2020*(1), 12748. https://doi.org/10.5465/AMBPP.2020.12748symposium

Blackler, A. L., Desai, S., McEwan, M., Popovic, V., & Diefenbach, S. (2018). Perspectives on the nature of intuitive interaction. In A. L. Blackler (Ed.), *Intuitive interaction: Research and application.* CRC Press. https://eprints.qut.edu.au/117564/

Blackler, A. L., Popovic, V., & Desai, S. (2018). Research methods for intuitive interaction. In A. L. Blackler (Ed.), *Intuitive interaction: Research and application* (pp. 65–88). CRC Press.

Blackler, A. L., Popovic, V., & Mahar, D. P. (2003). The nature of intuitive use of products: An experimental approach. *Design Studies*, *24*(6), 491–506. https://doi.org/10.1016/S0142-694X(03)00038-3

Blader, S. L., & Tyler, T. R. (2003). A four-component model of procedural justice: Defining the meaning of a "fair" process. *Personality & Social Psychology Bulletin*, *29*(6), 747–758.

Blau, P. M. (1964). *Exchange and power in social life*. John Wiley & Sons Inc.

Breazeal, C. (2004). Social interactions in HRI: The robot view. *IEEE Transactions on Systems, Man and Cybernetics, Part C (Applications and Reviews)*, *34*(2), 181–186. https://doi.org/10.1109/TSMCC.2004.826268

Britt, T. W., & Garrity, M. J. (2006). Attributions and personality as predictors of the road rage response. *British Journal of Social Psychology*, *45*(1), 127–147. https://doi.org/10.1348/014466605X41355

Broadbent, E. (2017). Interactions with robots: The truths we reveal about ourselves. *Annual Review of Psychology*, *68*, 627–652. https://doi.org/10.1146/annurev-psych-010416-043958

Broadbent, E., Stafford, R., & MacDonald, B. (2009). Acceptance of healthcare robots for the older population: Review and future directions. *International Journal of Social Robotics*, *1*(4), 319–330. https://doi.org/10.1007/s12369-009-0030-6

Brockner, J., & Wiesenfeld, B. M. (2005). How, when, and why does outcome favorability interact with procedural fairness? In J. Greenberg & J. A. Colquitt (Eds.), *Handbook of organizational justice* (pp. 525–554). Erlbaum.

Brockner, J., & Wiesenfeld, B. M. (2020). Organizational justice is alive and well and living elsewhere (but not too far away). In E. A. Lind (Ed.), *Frontiers in social psychology. Social psychology and justice* (pp. 213–242). Routledge.

Brockner, J., Wiesenfeld, B. M., & Diekmann, K. A. (2009). Towards a "fairer" conception of process fairness: Why, when and how more may not always be better than less. *The Academy of Management Annals*, *3*(1), 183–216. https://doi.org/10.1080/19416520903047228

Burke, J. L., Prewett, M. S., Gray, A. A., Yang, L., Stilson, F. R. B., Coovert, M. D., Elliot, L. R., & Redden, E. (2006). Comparing the effects of visual-auditory and visual-tactile feedback on user performance. In F. Quek, J. Yang, D. Massaro, A. Alwan, & T. J. Hazen (Eds.), *ICMI '06: Proceedings of the 8th International Conference on Multimodal Interfaces* (p. 108). ACM Press. https://doi.org/10.1145/1180995.1181017

Cascio, W. F., & Montealegre, R. (2016). How technology is changing work and organizations. *Annual Review of Organizational Psychology and Organizational Behavior*, *3*(1), 349–375. https://doi.org/10.1146/annurev-orgpsych-041015-062352

Chen, J. Y. C., Oden, R. V. N., & Merritt, J. O. (2014). Utility of stereoscopic displays for indirect-vision driving and robot teleoperation. *Ergonomics*, *57*(1), 12–22. https://doi.org/10.1080/00140139.2013.859739

Chen, S.-Y., Wu, W.-C., Chang, C.-S., Lin, C.-T., Kung, J.-Y., Weng, H.-C., Lin, Y.-T., & Lee, S.-I. (2015). Organizational justice, trust, and identification and their effects on organizational commitment in hospital nursing staff. *BMC Health Services Research*, *15*, 363. https://doi.org/10.1186/s12913-015-1016-8

Cheung, M. W. L. (2014). MetaSEM: An R package for meta-analysis using structural equation modeling. *Frontiers in Psychology*, *5*, 1521. https://doi.org/10.3389/fpsyg.2014.01521

Cheung, M. W. L. (2015). *Meta-analysis: A structural equation modeling approach*. Wiley.

Cheung, S. F., & Chan, D. K. S. (2004). Dependent effect sizes in meta-analysis: Incorporating the degree of interdependence. *Journal of Applied Psychology*, *89*(5), 780–791. https://doi.org/10.1037/0021-9010.89.5.780

Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). *Applied multiple regression/correlation analysis for the behavioral sciences* (Third edition). Routledge.

Cohen-Charash, Y., & Spector, P. E. (2001). The role of justice in organizations: A meta-analysis. *Organizational Behavior and Human Decision Processes*, *86*, 278–321. https://doi.org/10.1006/obhd.2001.2958

Coiffet, P., & Chirouze, M. (1983). Robots and robots in use. In P. Coiffet & M. Chirouze (Eds.), *An introduction to robot technology* (pp. 17–24). Springer Netherlands. https://doi.org/10.1007/978-94-011-6100-8_2

Cojuharenco, I., Marques, T., & Patient, D. (2017). Tell me who, and I'll tell you how fair: A model of agent bias in justice reasoning. *Group & Organization Management*, *42*(5), 630–656. https://doi.org/10.1177/1059601117729607

Colquitt, J. A. (2001). On the dimensionality of organizational justice: A construct validation of a measure. *Journal of Applied Psychology*, *86*, 386–400. https://doi.org/10.1037/0021-9010.86.3.386

Colquitt, J. A., Baer, M. D., Long, D. M., & Halvorsen-Ganepola, M. D. K. (2014). Scale indicators of social exchange relationships: A comparison of relative content validity. *Journal of Applied Psychology*, *99*(4), 599–618. https://doi.org/10.1037/a0036374

Colquitt, J. A., Conlon, D. E., Wesson, M. J., Porter, C. O. L. H., & Ng, K. Y. (2001). Justice at the millennium: A meta-analytic review of 25 years of organizational justice research. *Journal of Applied Psychology*, *86*, 425–445. https://doi.org/10.1037//0021-9010.86.3.425

Colquitt, J. A., & Rodell, J. B. (2015). Measuring justice and fairness. In R. S. Cropanzano & M. L. Ambrose (Eds.), *Oxford library of psychology. The oxford handbook of justice in the workplace* (pp. 187–202). Oxford University Press.

Colquitt, J. A., Scott, B. A., Rodell, J. B., Long, D. M., Zapata, C. P., Conlon, D. E., & Wesson, M. J. (2013). Justice at the millennium, a decade later: A meta-analytic test of social exchange and affect-based perspectives. *Journal of Applied Psychology*, *98*, 199–236. https://doi.org/10.1037/a0031757

Colquitt, J. A., & Zipay, K. P. (2015). Justice, fairness, and employee reactions. *Annual Review of Organizational Psychology and Organizational Behavior*, *2*(1), 75–99. https://doi.org/10.1146/annurev-orgpsych-032414-111457

Cornelis, I., van Hiel, A., & Cremer, D. de (2011). Birds of a feather: Leader-follower similarity and procedural fairness effects on cooperation. *European Journal of Work and Organizational Psychology*, *20*, 388–415. https://doi.org/10.1080/13594321003630055

Cremer, D. de (2006). When authorities influence followers' affect: The interactive effect of procedural justice and transformational leadership. *European Journal of Work and Organizational Psychology*, *15*(3), 322–351. https://doi.org/10.1080/13594320600627662

Cremer, D. de, & den Ouden, N. (2009). "When passion breeds justice": Procedural fairness effects as a function of authority's passion. *European Journal of Social Psychology*, *39*, 384–400. https://doi.org/10.1002/ejsp.529

Cremer, D. de, & Tyler, T. R. (2005). Managing group behavior: The interplay between procedural justice, sense of self, and cooperation. In M. P. Zanna (Ed.), *Advances in* *experimental social psychology* (Vol. 37, pp. 151–218). Elsevier. https://doi.org/10.1016/S0065-2601(05)37003-1

Cropanzano, R. S., Ambrose, M. L., & van Wagoner, P. (2020). Organizational justice and workplace emotion. In E. A. Lind (Ed.), *Frontiers in social psychology. Social psychology and justice* (pp. 243–283). Routledge.

Cropanzano, R. S., Rupp, D. E., Mohler, C. J., & Schminke, M. (2001). Three roads to organizational justice. In J. Ferris (Ed.), *Research in personnel and human resources management* (pp. 1–113). JAI Press.

Cropanzano, R. S., Weiss, H. M., Suckow, K. J., & Grandey, A. A. (2000). Doing justice to workplace emotion. In N. M. Ashkanasy, C. E. Härtel, & W. J. Zerbe (Eds.), *Emotions in the workplace: Research, theory, and practice* (pp. 49–62). Quorum Books/Greenwood Publishing Group.

Dahlbäck, N., Jönsson, A., & Ahrenberg, L. (1993). Wizard of Oz studies: Why and how. *Knowledge-Based Systems*, *6*(4), 258–266. https://doi.org/10.1016/0950-7051(93)90017-N

Deci, E. L., Olafsen, A. H., & Ryan, R. M. (2017). Self-determination theory in work organizations: The state of a science. *Annual Review of Organizational Psychology and Organizational Behavior*, *4*(1), 19–43. https://doi.org/10.1146/annurev-orgpsych-032516-113108

Deci, E. L., & Ryan, R. M. (2012). Self-determination theory. In P. A. M. van Lange (Ed.), *Handbook of theories of social psychology* (pp. 416–437). SAGE. https://doi.org/10.4135/9781446249215.n21

Diaz, I., Gil, J. J., & Louredo, M. (2014). A haptic pedal for surgery assistance. *Computer Methods and Programs in Biomedicine*, *116*(2), 97–104. https://doi.org/10.1016/j.cmpb.2013.10.010

Dineen, B. R., Noe, R. A., & Wang, C. (2004). Perceived fairness of web-based applicant screening procedures: Weighing the rules of justice and the role of individual differences. *Human Resource Management*, *43*, 127–145. https://doi.org/10.1002/hrm.20011

Duchaine, V., St-Onge, B. M., Gao, D., & Gosselin, C. (2012). Stable and intuitive control of an intelligent assist device. *IEEE Transactions on Haptics*, *5*(2), 148–159. https://doi.org/10.1109/ToH.2011.49

Duffy, B. R. (2003). Anthropomorphism and the social robot. *Robotics and Autonomous Systems*, *42*(3-4), 177–190. https://doi.org/10.1016/S0921-8890(02)00374-3

Dutton, J. E., & Ragins, B. R. (Eds.). (2007). *LEA's organization and management series. Exploring positive relationships at work: Building a theoretical and research foundation*. Lawrence Erlbaum Associates Publishers.

Edwards, C., Edwards, A., Spence, P. R., & Shelton, A. K. (2014). Is that a bot running the social media feed? Testing the differences in perceptions of communication quality for a human agent and a bot agent on Twitter. *Computers in Human Behavior*, *33*, 372–376. https://doi.org/10.1016/j.chb.2013.08.013

El Mesbahi, M., Elmarzouqi, N., & Lapayre, J.-C. (2014). Fairness properties for collaborative work using human-computer interactions and human-robot interactions based environment: "Let us be fair". In N. Streitz, & P. Markopoulos (Eds.), *Lecture Notes in Computer Science: Vol. 8530. Distributed, ambient, and pervasive interactions.* (pp. 319–328). Springer, Cham. https://doi.org/10.1007/978-3-319-07788-8_30

Endsley, M. R. (2000). Theoretical underpinnings of situation awareness: A critical review. In M. R. Endsley & D. J. Garland (Eds.), *Situation awareness: Analysis and measurement* (pp. 3–28). CRC Press.

Fink, J. (2012). Anthropomorphism and human likeness in the design of robots and human-robot interaction. In S. S. Ge, O. Khatib, J.-J. Cabibihan, R. Simmons, & M.-A. Williams

(Eds.), *Social robotics: Vol. 7621.* (pp. 199–208). Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-34103-8_20

Folger, R., & Cropanzano, R. S. (2001). Fairness theory: Justice as accountability. In J. Greenberg & R. S. Cropanzano (Eds.), *Advances in organizational justice* (pp. 1–55). Stanford University Press.

Fong, T., Thorpe, C., & Baur, C. (2003). Collaboration, dialogue, human-robot interaction. In R. A. Jarvis & A. Zelinsky (Eds.), *Springer Tracts in Advanced Robotics: Vol. 6. Robotics research: The tenth international symposium.* (pp. 255–266). Springer. https://doi.org/10.1007/3-540-36460-9_17

Franklin, S., Madl, T., D'Mello, S., & Snaider, J. (2014). LIDA: A systems-level architecture for cognition, emotion, and learning. *IEEE Transactions on Autonomous Mental Development*, *6*, 19–41. https://doi.org/10.1109/TAMD.2013.2277589

Fussell, S. R., Kiesler, S., Setlock, L. D., & Yew, V. (2008). How people anthropomorphize robots. In T. Fong, K. Dautenhahn, M. Scheutz, & Y. Demiris (Eds.), *HRI '08: Proceedings of the 3rd International Conference on Human-Robot Interaction* (pp. 145–152). ACM Press. https://doi.org/10.1145/1349822.1349842

Gagné, M., & Deci, E. L. (2005). Self-determination theory and work motivation. *Journal of Organizational Behavior*, *26*(4), 331–362. https://doi.org/10.1002/job.322

Ganster, D. C., & Rosen, C. C. (2013). Work stress and employee health. *Journal of Management*, *39*(5), 1085–1122. https://doi.org/10.1177/0149206313475815

Gelfand, M. J., Erez, M., & Aycan, Z. (2007). Cross-cultural organizational behavior. *Annual Review of Psychology*, *58*, 479–514. https://doi.org/10.1146/annurev.psych.58.110405.085559

Goetz, J., Kiesler, S., & Powers, A. (2003, October). Matching robot appearance and behavior to tasks to improve human-robot cooperation. In *RO-MAN '03: Proceedings of the 12th IEEE International Workshop on Robot and Human Interactive Communication,* Millbrae, CA, USA.

Goldman, J., & Procaccia, A. D. (2015). Spliddit. *ACM SIGecom Exchanges*, *13*(2), 41–46. https://doi.org/10.1145/2728732.2728738

Gombolay, M. C., Gutierrez, R. A., Sturla, G. F., & Shah, J. A. (2014). Decision-making authority, team efficiency and human worker satisfaction in mixed human-robot teams. In D. Fox, L. E. Kavraki, & H. Kurniawati (Chairs), *Robotics: Science and Systems X,* Berkeley, USA.

Gong, L. (2008). How social is social responses to computers? The function of the degree of anthropomorphism in computer representations. *Computers in Human Behavior*, *24*, 1494–1509. https://doi.org/10.1016/j.chb.2007.05.007

Goodrich, M. A., & Schultz, A. C. (2007). Human-robot interaction: A survey. *Foundations and Trends® in Human-Computer Interaction*, *1*(3), 203–275. https://doi.org/10.1561/1100000005

Gopinathan, S., Ötting, S. K., & Steil, J. J. (2017). A user study on personalized stiffness control and task specificity in physical human-robot interaction. *Frontiers in Robotics and AI*, *4*, Article 58. https://doi.org/10.3389/frobt.2017.00058

Graaf, M. M. A. de, & Ben Allouch, S. (2013). Exploring influencing variables for the acceptance of social robots. *Robotics and Autonomous Systems*, *61*(12), 1476–1486. https://doi.org/10.1016/j.robot.2013.07.007

Greenberg, J. (1987). A taxonomy of organizational justice theories. *Academy of Management Review*, *12*(1), 9–22. https://doi.org/10.5465/amr.1987.4306437

Greenberg, J. (1993). The social side of fairness: Interpersonal and informational classes of organizational justice. In R. S. Cropanzano (Ed.), *Series in applied psychology. Justice in the workplace: Approaching fairness in human resource management* (pp. 79–103). Lawrence Erlbaum Associates, Inc.

Greenberg, J. (2011). Organizational justice: The dynamics of fairness in the workplace. In S. Zedeck (Ed.), *APA handbook of industrial and organizational psychology: Vol. 3. Maintaining, expanding, and contracting the organization* (pp. 271–327). American Psychological Association. https://doi.org/10.1037/12171-008

Griffin, M. A., & Clarke, S. (2011). Stress and well-being at work. In S. Zedeck (Ed.), *APA handbook of industrial and organizational psychology: Vol. 3. Maintaining, expanding, and contracting the organization* (pp. 359–397). American Psychological Association. https://doi.org/10.1037/12171-010

Hancock, P. A., Billings, D. R., Schaefer, K. E., Chen, J. Y. C., Visser, E. J. de, & Parasuraman, R. (2011). A meta-analysis of factors affecting trust in human-robot interaction. *Human Factors*, *53*(5), 517–527. https://doi.org/10.1177/0018720811417254

Häring, M., Kuchenbrandt, D., & André, E. (2014). Would you like to play with me? How robots' group membership and task features influence human-robot interaction. In G. Sagerer, M. Imai, T. Belpaeme, & A. Thomaz (Eds.), *HRI'14, HRI '14: Proceedings of the 2014 ACM/IEEE International Conference on Human-Robot Interaction* (pp. 9–16). ACM Press.

Harriott, C. E., Zhang, T., & Adams, J. A. (2013). Assessing physical workload for human–robot peer-based teams. *International Journal of Human-Computer Studies*, *71*(7-8), 821–837. https://doi.org/10.1016/j.ijhcs.2013.04.005

Hartson, R. (2003). Cognitive, physical, sensory, and functional affordances in interaction design. *Behaviour & Information Technology*, *22*(5), 315–338. https://doi.org/10.1080/01449290310001592587

Hayes, A. F. (2018). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach* (Second edition). *Methodology in the social sciences*. The Guilford Press. https://ebookcentral.proquest.com/lib/potsdamuni/detail.action?docID=5109647

Heerink, M., Kröse, B., Evers, V., & Wielinga, B. (2010). Assessing acceptance of assistive social agent technology by older adults: The Almere Model. *International Journal of Social Robotics*, *2*(4), 361–375. https://doi.org/10.1007/s12369-010-0068-5

Hinds, P. J., Roberts, T. L., & Jones, H. (2004). Whose job is it anyway? A study of human-robot interaction in a collaborative task. *Human–Computer Interaction*, *19*(1), 151–181. https://doi.org/10.1207/s15327051hci1901&2_7

Hoff, K. A., & Bashir, M. (2015). Trust in automation: Integrating empirical evidence on factors that influence trust. *Human Factors*, *57*(3), 407–434. https://doi.org/10.1177/0018720814547570

Hoffman, G., & Breazeal, C. (2007). Effects of anticipatory action on human-robot teamwork efficiency, fluency, and perception of team. In *HRI '07: Proceedings of the 2nd ACM/IEEE International Conference on Human-Robot Interaction* (pp. 1–8). ACM Press. https://doi.org/10.1145/1228716.1228718

Hofstede, G., Hofstede, G. J., & Minkov, M. (2010). *Cultures and organizations: Software of the mind*. Intercultural cooperation and its importance for survival (3rd edition). McGraw-Hill.

Huang, C.-M., & Mutlu, B. (2016). Anticipatory robot control for efficient human-robot collaboration. In *HRI '16: Proceedings of the Eleventh ACM/IEEE International*

Conference on Human Robot Interaction (pp. 83–90). IEEE Press. http://dl.acm.org/citation.cfm?id=2906831.2906846

Humphrey, S. E., Nahrgang, J. D., & Morgeson, F. P. (2007). Integrating motivational, social, and contextual work design features: A meta-analytic summary and theoretical extension of the work design literature. *Journal of Applied Psychology*, *92*(5), 1332–1356. https://doi.org/10.1037/0021-9010.92.5.1332

(2019). *World robotics 2019: Industrial robots*. Frankfurt am Main. IFR Statistical Department.

International Organization for Standardization (2012). *Robots and robotic devices* (ISO 8373:2012).

Jenkins, J., & Blyton, P. (2008). Works councils. In P. Blyton, N. Bacon, J. Fiorito, & E. Heery (Eds.), *The SAGE handbook of industrial relations* (pp. 346–357). SAGE.

Jiang, Z., Gollan, P. J., & Brooks, G. (2017). Relationships between organizational justice, organizational trust and organizational commitment: Across-cultural study of China, South Korea and Australia. *The International Journal of Human Resource Management*, *28*(7), 973–1004. https://doi.org/10.1080/09585192.2015.1128457

Jones, D. A., & Skarlicki, D. P. (2012). How perceptions of fairness can change. *Organizational Psychology Review*, *3*, 138–160. https://doi.org/10.1177/2041386612461665

Kahn, P. H., Jr., Gary, H. E., & Shen, S. (2013). Social and moral relationships with robots: Genetic epistemology in an exponentially increasing technological world. *Human Development*, *56*(1), 1–4. https://doi.org/10.1159/000345544

Kauffeld, S., & Maier, G. W. (2020). Digitalisierte Arbeitswelt [Digitalized world of work]. *Gruppe. Interaktion. Organisation. Zeitschrift Für Angewandte Organisationspsychologie (GIO)*, *51*(1), 1–4. https://doi.org/10.1007/s11612-020-00508-y

Kepes, S., & McDaniel, M. A. (2013). How trustworthy is the scientific literature in industrial and organizational psychology? *Industrial and Organizational Psychology*, *6*(3), 252–268. https://doi.org/10.1111/iops.12045

Khan, A. K., Quratulain, S., & Crawshaw, J. R. (2013). The mediating role of discrete emotions in the relationship between injustice and counterproductive work behaviors: A study in Pakistan. *Journal of Business and Psychology*, *28*(1), 49–61. https://doi.org/10.1007/s10869-012-9269-2

Konradt, U., Garbers, Y., Erdogan, B., & Bauer, T. N. (2016). Patterns of change in fairness perceptions during the hiring process. *International Journal of Selection and Assessment*, *24*, 246–259. https://doi.org/10.1111/ijsa.12144

Kuchenbrandt, D., Eyssel, F., Bobinger, S., & Neufeld, M. (2013). When a robot's group membership matters. *International Journal of Social Robotics*, *5*(3), 409–417. https://doi.org/10.1007/s12369-013-0197-8

Kulms, P., & Kopp, S. (2019). More human-likeness, more trust? In F. Alt, A. Bulling, & T. Döring (Eds.), *ICPS: ACM international conference proceeding series, Muc '19: Tagungsband Mensch und Computer 2019* (pp. 31–42). The Association for Computing Machinery, Inc. https://doi.org/10.1145/3340764.3340793

Kwon, M., Jung, M. F., & Knepper, R. A. (2016). Human expectations of social robots. In *HRI '16: Proceedings of the Eleventh ACM/IEEE International Conference on Human Robot Interaction* (pp. 463–464). IEEE Press. http://dl.acm.org/citation.cfm?id=2906831.2906928

Langer, M., König, C. J., & Busch, V. (2020). Changing the means of managerial work: Effects of automated decision-support systems on personnel selection tasks. *Journal of Business and Psychology*. https://doi.org/10.1007/s10869-020-09711-6

Lee, K. M., Jung, Y., Kim, J., & Kim, S. R. (2006). Are physically embodied social agents better than disembodied social agents? The effects of physical embodiment, tactile interaction, and people's loneliness in human–robot interaction. *International Journal of Human-Computer Studies*, *64*(10), 962–973. https://doi.org/10.1016/j.ijhcs.2006.05.002

Lee, M. K., & Baykal, S. (2017). Algorithmic mediation in group decisions. In C. P. Lee, S. Poltrock, L. Barkhuus, M. Borges, & W. A. Kellogg (Eds.), *CSCW '17: Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing* (pp. 1035–1048). ACM Press. https://doi.org/10.1145/2998181.2998230

Lee, S. A., & Liang, Y. J. (2015). Reciprocity in computer-human interaction: Source-based, norm-based, and affect-based explanations. *Cyberpsychology, Behavior and Social Networking*, *18*, 234–240. https://doi.org/10.1089/cyber.2014.0458

Legg, S., & Hutter, M. (2007). Universal intelligence: A definition of machine intelligence. *Minds and Machines*, *17*(4), 391–444. https://doi.org/10.1007/s11023-007-9079-x

Lewin, K. (1943). Defining the 'field at a given time'. *Psychological Review*, *50*, 292–310. https://doi.org/10.1037/h0062738

Liem, C. C. S., Langer, M., Demetriou, A., Hiemstra, A. M. F., Sukma Wicaksana, A., Born, M. P., & König, C. J. (2018). *Psychology meets machine learning: Interdisciplinary perspectives on algorithmic job candidate screening*. Springer International Publishing. https://doi.org/10.1007/978-3-319-98131-4_9

Machado, L., Prikladnicki, R., Meneguzzi, F., Souza, C. R. B. de, & Carmel, E. (2016). Task allocation for crowdsourcing using AI planning. In *CSI-SE '16: Proceedings of the 3rd*

International Workshop on CrowdSourcing in Software Engineering (pp. 36–40). ACM Press. https://doi.org/10.1145/2897659.2897666

Maier, J. R. A., & Fadel, G. M. (2009). Affordance based design: A relational theory for design. *Research in Engineering Design*, *20*(1), 13–27. https://doi.org/10.1007/s00163-008-0060-3

Malle, B. F. (2016). Integrating robot ethics and machine morality: The study and design of moral competence in robots. *Ethics and Information Technology*, *18*(4), 243–256. https://doi.org/10.1007/s10676-015-9367-8

Malle, B. F., & Scheutz, M. (2014). Moral competence in social robots. In *ETHICS '14: Proceedings of the IEEE 2014 International Symposium on Ethics in Engineering, Science, and Technology* (8:1-8:6). IEEE Press. http://dl.acm.org/citation.cfm?id=2960587.2960597

Manzey, D., Reichenbach, J., & Onnasch, L. (2012). Human performance consequences of automated decision aids. *Journal of Cognitive Engineering and Decision Making*, *6*(1), 57–87. https://doi.org/10.1177/1555343411433844

Marcinkowski, F., Kieslich, K., Starke, C., & Lünich, M. (2020). Implications of AI (un-)fairness in higher education admissions. In M. Hildebrandt, C. Castillo, E. Celis, S. Ruggieri, L. Taylor, & G. Zanfir-Fortuna (Eds.), *FAT '20: Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency* (pp. 122–130). ACM. https://doi.org/10.1145/3351095.3372867

Marques, T., Patient, D., & Cojuharenco, I. (2017). The "who" of organizational justice: Source effects on justice judgements. In C. Moliner, R. S. Cropanzano, & V. Martínez-Tur (Eds.), *Organizational justice: International perspectives and conceptual advances* (pp. 58–84). Routledge.

Mavridis, N., Pierris, G., Gallina, P., Moustakas, N., & Astaras, A. (2015). Subjective difficulty and indicators of performance of joystick-based robot arm teleoperation with auditory feedback. In *ICAR '15: Proceedings of the 17th International Conference on Advanced Robotics* (pp. 91–98). IEEE Press. https://doi.org/10.1109/ICAR.2015.7251439

Mlekus, L., Lehmann, J., & Maier, G. W. (2019). *New work situations call for familiar work design methods: Effects and mediating mechanisms of task rotation in a technology-supported workplace* [Manuscript submitted for publication]. Department of Psychology, Bielefeld University.

Morgeson, F. P., & Humphrey, S. E. (2006). The Work Design Questionnaire (WDQ): Developing and validating a comprehensive measure for assessing job design and the nature of work. *Journal of Applied Psychology*, *91*(6), 1321–1339. https://doi.org/10.1037/0021-9010.91.6.1321

Muggleton, S., Schmid, U., Zeller, C., Tamaddoni-Nezhad, A., & Besold, T. (2018). Ultra-strong machine learning: Comprehensibility of programs learned with ILP. *Machine Learning, 107, 1119-1140*. https://doi.org/10.1007/s10994-018-5707-3

Muxfeldt, A., Gopinathan, S., Coenders, T., & Steil, J. J. (2017). A user study on human-robot-interactive recovery for industrial assembly problems. In *RO-MAN '17: Proceedings of the 26th IEEE International Symposium on Robot and Human Interactive Communication* (pp. 824–830). IEEE Press. https://doi.org/10.1109/ROMAN.2017.8172398

Naim, I., Tanveer, M. I., Gildea, D., & Hoque, E. (2016). Automated analysis and prediction of job interview performance. *IEEE Transactions on Affective Computing*. https://doi.org/10.1109/TAFFC.2016.2614299

Nass, C. I., & Moon, Y. (2000). Machines and mindlessness: Social responses to computers. *Journal of Social Issues*, *56(1)*, 81–103. https://doi.org/10.1111/0022-4537.00153

Nass, C. I., Moon, Y., & Carney, P. (1999). Are people polite to computers? Responses to computer-based interviewing systems. *Journal of Applied Social Psychology*, *29(5)*, 1093–1109. https://doi.org/10.1111/j.1559-1816.1999.tb00142.x

Nass, C. I., Moon, Y., & Green, N. (1997). Are machines gender neutral? Gender-stereotypic responses to computers with voices. *Journal of Applied Social Psychology*, *27(10)*, 864–876. https://doi.org/10.1111/j.1559-1816.1997.tb00275.x

Negnevitsky, M. (2005). *Artificial intelligence: A guide to intelligent systems* (2nd ed.). Addison-Wesley.

Norman, D. A. (1988). *The psychology of everyday things*. Basic Books.

Onnasch, L., Maier, X., & Jürgensohn, T. (2016). *Mensch-Roboter-Interaktion: Eine Taxonomie für alle Anwendungsfälle* (baua: Fokus). Bundesanstalt für Arbeitsschutz und Arbeitsmedizin.

Onnasch, L., Wickens, C. D., Li, H., & Manzey, D. (2014). Human performance consequences of stages and levels of automation: An integrated meta-analysis. *Human Factors*, *56*(3), 476–488. https://doi.org/10.1177/0018720813501549

Oracle & Future Workplace (2019). *AI@Work Study 2019.* https://www.oracle.com/a/ocom/docs/applications/hcm/ai-at-work-ebook.pdf

Organ, D. W. (1990). The motivational basis of organizational citizenship behavior. *Research in Organizational Behavior*, *12*(1), 43–72.

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*(3), 286–297. https://doi.org/10.1109/3468.844354

Parasuraman, R., Sheridan, T. B., & Wickens, C. D. (2008). Situation awareness, mental workload, and trust in automation: Viable, empirically supported Cognitive Engineering Constructs. *Journal of Cognitive Engineering and Decision Making*, *2*(2), 140–160. https://doi.org/10.1518/155534308X284417

Parker, S. K., & Grote, G. (2020). Automation, algorithms, and beyond: Why work design matters more than ever in a digital world. *Applied Psychology.* Advance online publication. https://doi.org/10.1111/apps.12241

Peña, J., Ghaznavi, J., Brody, N., Prada, R., Martinho, C., Santos, P. A., Damas, H., & Dimas, J. (2019). Effects of human vs. computer-controlled characters and social identity cues on enjoyment. *Journal of Media Psychology*, *31,* 35-47. https://doi.org/10.1027/1864-1105/a000218

Pfeifer, C. (2014). Determinants of fair own wage perceptions: the moderating effect of works councils and performance evaluations. *Applied Economics Letters*, *21*(1), 47–50. https://doi.org/10.1080/13504851.2013.837572

Phillips, J. M. (2002). Antecedents and consequences of procedural justice perceptions in hierarchical decision-making teams. *Small Group Research*, *33*(1), 32–64. https://doi.org/10.1177/104649640203300102

Rhoten, D., & Parker, A. (2004). Education: Risks and rewards of an interdisciplinary research path. *Science*, *306*(5704), 2046. https://doi.org/10.1126/science.1103628

Richter, M., König, C. J., Koppermann, C., & Schilling, M. (2016). Displaying fairness while delivering bad news: Testing the effectiveness of organizational bad news training in the layoff context. *Journal of Applied Psychology*, *101(6)*, 779–792. https://doi.org/10.1037/apl0000087

Riek, L. (2012). Wizard of Oz studies in HRI: A systematic review and new reporting guidelines. *Journal of Human-Robot Interaction*, *1(1)*, 119–136. https://doi.org/10.5898/JHRI.1.1.Riek

Robert, L. P., Pierce, C., Marquis, L., Kim, S., & Alahmad, R. (2020). Designing fair AI for managing employees in organizations: A review, critique, and design agenda. *Human–Computer Interaction*, *17*(2), 1–31. https://doi.org/10.1080/07370024.2020.1735391

Rupp, D. E., Shao, R., Jones, K. S., & Liao, H. (2014). The utility of a multifoci approach to the study of organizational justice: A meta-analytic investigation into the consideration of normative rules, moral accountability, bandwidth-fidelity, and social exchange. *Organizational Behavior and Human Decision Processes*, *123*(2), 159–185. https://doi.org/10.1016/j.obhdp.2013.10.011

Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*, *55*(1), 68–78. https://doi.org/10.1037//0003-066X.55.1.68

Sammut, C., & Webb, G. I. (2017). *Encyclopedia of machine learning and data mining*. Springer, US. https://doi.org/10.1007/978-1-4899-7687-1

Scandura, T. A., & Williams, E. A. (2000). Research methodology in management: Current practices, trends, and implications for future research. *Academy of Management Journal*, *43*(6), 1248–1264. https://doi.org/10.5465/1556348

Scheunemann, M. M., Cuijpers, R. H., & Salge, C. (2020). Warmth and competence to predict human preference of robot behavior in physical human-robot interaction. In *RO-MAN '20: Proceedings of the 29th International Conference on Robot and Human Interactive Communication,* Naples, Italy. http://arxiv.org/pdf/2008.05799v1

Schlicker, N., Langer, M., Ötting, S. K., Baum, K., König, C. J., & Wallach, D. P. (2020). *What to expect from opening up 'black boxes'? Comparing perceptions of justice between human and automated agents* [Manuscript submitted for publication].

Schmid, U., Zeller, C., Besold, T., Tamaddoni-Nezhad, A., & Muggleton, S. (2017). How does predicate invention affect human comprehensibility? In J. Cussens & A. Russo (Eds.), *Lecture notes in artificial intelligence: Vol. 10326. Inductive Logic Programming: ILP 2016, 26th international conference* (pp. 52–67). Springer, Berlin, Heidelberg.

Shanock, L. R., Roch, S. G., & Mishra, V. (2012). Why we should care about exchange relationships with coworkers as well as supervisors: Both fellow employees and the organization benefit. *TPM-Testing, Psychometrics, Methodology in Applied Psychology*, *19*(4), 291–310.

Shao, R., Rupp, D. E., Skarlicki, D. P., & Jones, K. S. (2012). Employee justice across cultures: A meta-analytic review. *Journal of Management*, *39*(1), 263–301. https://doi.org/10.1177/0149206311422447

Sheridan, T. B., & Verplank, W. L. (1978). *Human and computer control of undersea teleoperators (technical report)*. Cambridge, MA. MIT, Man Machine Systems Laboratory.

Sherony, K. M., & Green, S. G. (2002). Coworker exchange: Relationships between coworkers, leader-member exchange, and work attitudes. *Journal of Applied Psychology*, *87*(3), 542–548. https://doi.org/10.1037/0021-9010.87.3.542

Silva, M. R., & Caetano, A. (2016). Organizational justice across cultures: A systematic review of four decades of research and some directions for the future. *Social Justice Research*, *29*, 257–287. https://doi.org/10.1007/s11211-016-0263-0

Skarlicki, D. P., & Latham, G. P. (2005). How can training be used to foster organizational justice? In J. Greenberg & J. A. Colquitt (Eds.), *Handbook of organizational justice* (pp. 499–522). Erlbaum.

Soderberg, C. K., Callahan, S. P., Kochersberger, A. O., Amit, E., & Ledgerwood, A. (2015). The effects of psychological distance on abstraction: Two meta-analyses. *Psychological Bulletin*, *141*(3), 525–548. https://doi.org/10.1037/bul0000005

Soenen, G., & Melkonian, T. (2017). Fairness and commitment to change in M&As: The mediating role of organizational identification. *European Management Journal*, *35*(4), 486–492. https://doi.org/10.1016/j.emj.2016.10.005

Spong, M. W., Hutchinson, S., & Vidyasagar, M. (2020). *Robot modeling and control* (2nd ed.). Wiley.

Steil, J. J., & Maier, G. W. (2020). Kollaborative Roboter: Universale Werkzeuge in der digitalisierten und vernetzten Arbeitswelt [Collaborative robots: universal tools in the digitalized and interconnected working environment]. In G. W. Maier, G. Engels, & E. Steffen (Eds.), *Handbuch Gestaltung digitaler und vernetzter Arbeitswelten* (pp. 323–346). Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-662-52979-9_15

Stone, D. N., Deci, E. L., & Ryan, R. M. (2009). Beyond talk: Creating autonomous motivation through Self-Determination Theory. *Journal of General Management*, *34*(3), 75–91. https://doi.org/10.1177/030630700903400305

Streicher, B., Jonas, E., Maier, G. W., Frey, D., & Spießberger, A. (2012). Procedural fairness and creativity: Does voice maintain people's creative vein over time? *Creativity Research Journal*, *24*(4), 358–363. https://doi.org/10.1080/10400419.2012.730334

Szalma, J. L. (2014). On the application of motivation theory to human factors/ergonomics: Motivational design principles for human-technology interaction. *Human Factors*, *56*(8), 1453–1471. https://doi.org/10.1177/0018720814553471

Töniges, T., Ötting, S. K., Wrede, B., Maier, G. W., & Sagerer, G. (2017). An emerging decision authority: Adaptive cyber-physical system design for fair human-machine interaction and decision processes. In H. Song, D. B. Rawat, S. Jeschke, & C. Brecher (Eds.), *Intelligent data-centric systems: Sensor collected intelligence. Cyber-physical systems: Foundations, principles and applications* (pp. 419–430). Elsevier Academic Press. https://doi.org/10.1016/B978-0-12-803801-7.00026-2

Truxillo, D. M., Bauer, T. N., & Garcia, A. M. (2017). Applicant reactions to hiring procedures. In H. W. Goldstein, E. D. Pulakos, J. Passmore, & C. Semedo (Eds.), *Wiley Blackwell handbooks in organizational psychology. The Wiley Blackwell handbook of the psychology of recruitment, selection and employee retention* (pp. 53–70). Wiley Blackwell.

Tyler, T. R., & Blader, S. L. (2003). The group engagement model: Procedural justice, social identity, and cooperative behavior. *Personality and Social Psychology Review*, *7*(4), 349–361. https://doi.org/10.1207/S15327957PSPR0704_07

van der Woerdt, S., & Haselager, P. (2017). Lack of effort or lack of ability? Robot failures and human perception of agency and responsibility. In T. Bosse & B. Bredeweg (Eds.), *Communications in Computer and Information Science: Vol. 765, BNAIC '16: Proceedings of the 28th Benelux Conference on Artificial Intelligence. Revised selected papers* (pp. 155–168). Springer. https://doi.org/10.1007/978-3-319-67468-1_11

van der Woerdt, S., & Haselager, P. (2019). When robots appear to have a mind: The human perception of machine agency and responsibility. *New Ideas in Psychology*, *54*, 93–100. https://doi.org/10.1016/j.newideapsych.2017.11.001

Venkatesh, V., & Bala, H. (2008). Technology Acceptance Model 3 and a research agenda on interventions. *Decision Sciences*, *39*(2), 273–315. https://doi.org/10.1111/j.1540-5915.2008.00192.x

Visser, E. J. de, Monfort, S. S., McKendrick, R., Smith, M. A. B., McKnight, P. E., Krueger, F., & Parasuraman, R. (2016). Almost human: Anthropomorphism increases trust resilience in cognitive agents. *Journal of Experimental Psychology: Applied*, *22*(3), 331–349. https://doi.org/10.1037/xap0000092

Voiklis, J., Kim, B., Cusimano, C., & Malle, B. F. (2016). Moral judgments of human vs. robot agents. In *RO-MAN '16: Proceedings of the 25th IEEE International Symposium on Robot and Human Interactive Communication* (pp. 775–780). IEEE Press. https://doi.org/10.1109/ROMAN.2016.7745207

Wallach, W. (2010). Robot minds and human ethics: The need for a comprehensive model of moral decision making. *Ethics and Information Technology*, *12*, 243–250. https://doi.org/10.1007/s10676-010-9232-8

Walter, S., Wendt, C., Böhnke, J., Crawcour, S., Tan, J.-W., Chan, A., Limbrecht, K., Gruss, S., & Traue, H. C. (2014). Similarities and differences of emotions in human-machine and human-human interactions: What kind of emotions are relevant for future companion systems? *Ergonomics*, *57*(3), 374–386. https://doi.org/10.1080/00140139.2013.822566

Watson, D. (2000). *Mood and temperament*. *Emotions and social behavior*. Guilford Press.

Weiner, B. (1995). *Judgments of responsibility: A foundation for a theory of social conduct*. Guilford Press.

Weiner, B. (2006). *Social motivation, justice, and the moral emotions: An attributional approach*. Psychology Press.

Wesche, J. S., & Sonderegger, A. (2019). When computers take the lead: The automation of leadership. *Computers in Human Behavior*, *101*, 197–209. https://doi.org/10.1016/j.chb.2019.07.027

Whitman, D. S., Caleo, S., Carpenter, N. C., Horner, M. T., & Bernerth, J. B. (2012). Fairness at the collective level: A meta-analytic examination of the consequences and boundary conditions of organizational justice climate. *Journal of Applied Psychology*, *97*(4), 776–791. https://doi.org/10.1037/a0028021

Wickens, C. M., Wiesenthal, D. L., Flora, D. B., & Flett, G. L. (2011). Understanding driver anger and aggression: Attributional theory in the driving environment. *Journal of Experimental Psychology: Applied*, *17*(4), 354–370. https://doi.org/10.1037/a0025815

Wickens, C. D., McCarley, J. S., Alexander, A. L., Thomas, L. C., Ambinder, M., & Zheng, S. (2008). Attention-situation awareness (A-SA) model of pilot error. In D. C. Foyle & B. L. Hooey (Eds.), *Human performance modeling in aviation* (pp. 213–239). CRC Press.

Wolbeck, L. (2019). *Fairness aspects in personnel scheduling.* https://doi.org/10.17169/REFUBIUM-26050

Xie, Y., Bodala, I. P., Ong, D. C., Hsu, D., & Soh, H. (2019). Robot capability and intention in trust-based decisions across tasks. In *HRI '19: Proceedings of the 14th ACM/IEEE International Conference on Human-Robot Interaction* (pp. 39–47). IEEE. https://doi.org/10.1109/HRI.2019.8673084

You, H.-c., & Chen, K. (2007). Applications of affordance and semantics in product design. *Design Studies*, *28*(1), 23–38. https://doi.org/10.1016/j.destud.2006.07.002

Young, J. E., Sung, J., Voida, A., Sharlin, E., Igarashi, T., Christensen, H. I., & Grinter, R. E. (2011). Evaluating human-robot interaction. *International Journal of Social Robotics*, *3*(1), 53–67. https://doi.org/10.1007/s12369-010-0081-8

Zhu, J. (2009). *Intentional systems and the artificial intelligence (AI) hermeneutic network: Agency and intentionality in expressive computational systems* [Dissertation, Georgia Institute of Technology]. http://hdl.handle.net/1853/34730

Zhu, Y., & Fu, K.-w. (2019). The relationship between interdisciplinarity and journal impact factor in the field of communication during 1997–2016. *Journal of Communication*, *69*(3), 273–297. https://doi.org/10.1093/JOC/JQZ012

Ziemke, T., Thill, S., & Vernon, D. (2015). Embodiment is a double-edged sword in human-robot interaction: Ascribed vs. intrinsic intentionality. In *HRI '15 - Extended abstracts: Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction Extended Abstracts.* ACM.

Zweig, D., & Webster, J. (2002). Where is the line between benign and invasive? An examination of psychological barriers to the acceptance of awareness monitoring systems. *Journal of Organizational Behavior*, *23*, 605–633. https://doi.org/10.1002/job.157

Statement of Authorship

Hiermit erkläre ich, dass ich die vorliegende Dissertation "Artificial intelligence as colleague and supervisor: Successful and fair interactions between intelligent technologies and employees at work" weder in der gegenwärtigen noch in einer anderen Fassung einer anderen Fakultät vorgelegt habe oder hatte.

Ich versichere, dass ich die Dissertation selbstständig und ohne unerlaubte Hilfe angefertigt sowie unter ausschließlicher Verwendung der von mir angegebenen Quellen verfasst und wörtlich oder sinngemäß aus der Literatur entnommene Textstellen kenntlich gemacht habe.

Ferner bestätige ich, dass ich den federführenden Beitrag zu den unter gemeinschaftlicher Autorenschaft entstandenen Manuskripten geleistet habe.

Bielefeld, im November 2020

Sonja K. Ötting

Overview of Published and Submitted Work

Study 1

Ötting, S. K., Masjutin, L., & Maier, G. W. (2020). Let's work together: A meta-analysis on robot design features that enable successful human–robot interaction at work. *Human Factors*. https://doi.org/10.117/0018720820966433.

Study 2

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

Study 3

Ötting, S. K., Ongsiek, O., Kahlert, E., Fronia, J., & Maier, G. W. (2019). *How procedural justice works: Artificial Intelligence as a new decision agent and the mediation of justice effects*. [Manuscript submitted for publication].

Manuscript of Study 3

This is a pre-peer review version of the manuscript, submitted at the Journal *Organizational Behavior and Human Decision Processes.*

How Procedural Justice Works: Artificial Intelligence as a New Decision Agent and the Mediation of Justice Effects

Sonja K. Ötting, Olga Ongsiek, Eyleen Kahlert, Jessica Fronia, & Günter W. Maier

Bielefeld University, Department of Psychology

Correspondence concerning this article should be addressed to Sonja K. Ötting, Bielefeld University, Universitätsstraße 25, 33615 Bielefeld, Germany. Email: soetting@uni-bielefeld.de

Acknowledgements

Funding: This work was supported by the North-Rhine Westfalian graduate school "Design of flexible working environments – Human Centered Cyber-Physical Systems in Industry 4.0" (NRW Forschungskolleg: Gestaltung von flexiblen Arbeitswelten – Menschenzentrierte Nutzung von Cyber-Physical Systems in Industrie 4.0) funded by the Ministerium für Kultur und Wissenschaft des Landes Nordrhein-Westfalen.

How Procedural Justice Works: Artificial Intelligence as a New Decision Agent and the Mediation of Justice Effects

Abstract

Artificial intelligence (AI) recently emerged as new decision agent in organizations, which might change how justice affects employees. We investigated differences between human and AI decision agents concerning procedural justice effects on attitudes and behavior (i.e., job satisfaction, commitment, organizational citizenship, and counterproductive work behaviors) and mediating mechanisms (i.e., positive and negative affect, trust, and identification). Using experimental vignettes on two samples of 229 and 132 employees, we manipulated procedural justice and decision agent. The results showed that trust was the strongest mediator for attitudes and is less pronounced for AI decisions, whereas negative affect was the strongest mediator for behavior, with no difference between decision agents. We confirmed responsibility of a decision agent as one underlying mechanism for these differences. These results are vital for designing and implementing AI. They also show that future research needs to investigate the characteristics of decision agents more closely.

Keywords: procedural justice, decision agent, source of justice, artificial intelligence, responsibility, moderated mediation

Imagine the following: in the future, an artificial intelligence (AI) will make critical decisions about employees' daily work. Unimaginable? It is already happening! AIs already make decisions about the allocation of tasks and shifts (Franklin et al., 2014), or demands for training (Langer et al., 2016). They have emerged as new leaders and decision agents in organizations, not only as bodiless algorithms in the human resources department, but also as robotic team member at shop floor level (Larson & DeChurch, 2020; Wesche & Sonderegger, 2019).

The question of how decisions in organizations should be designed to be perceived as fair, consequently leading to positive employee attitudes and behavior, has been the topic of organizational justice research for more than four decades now (Brockner & Wiesenfeld, 2020; Greenberg, 1987). However, even though justice research has shown that leaders should deeply care about justice when making decisions (Colquitt et al., 2013; Colquitt & Zipay, 2015), AI research has merely touched the subject of justice and it rarely incorporates the guidelines that organizational justice research can offer (Robert et al., 2020). Only recently, the inclusion of AI as decision (or justice) agents has been identified as a major topic in future organizational justice research (Brockner & Wiesenfeld, 2020). Justice researchers need to investigate how characteristics of AI decision agents and human decision agents differ, how these differences might impact organizational justice effects on employee attitudes and behavior, and whether there are differences between decision agents in how these effects are mediated. This research will provide vital insights for system designers who wish to design AI that makes just decisions and for organizations who want to successfully implement those technologies.

Although there are various theories that can explain how organizational justice affects employee attitudes and behavior (Colquitt & Zipay, 2015; Greenberg, 1987), and many differentiate between decision agents such as organization and supervisor (Rupp et al., 2014), academics still know little about how specific characteristics of a decision agent influence justice effects. Only a few studies have investigated perceptions of organizational justice in the context of an AI as decision agent (e.g., Ötting & Maier, 2018; Uhde et al., 2020), and none have approached how these differences might influence explaining mechanisms for justice effects.

In order to explain differences in reactions to decisions, both research in organizational justice (Folger & Cropanzano, 2001), as well as research in AI, and human-machine interactions (Wickens et al., 2011) uses the concept of responsibility. Fairness theory, for example, states that if decision agents are not perceived as responsible for their decision, then justice cannot unfold its effects (Folger & Cropanzano, 2001). Attributions of responsibility are made when the decision agent had control over the decision or a had clear intention to make this particular decision (Weiner, 1995, 2006). Research on human-machine interactions shows that artificial agents are less likely to be held morally responsible (Voiklis et al., 2016). Yet, despite using the same concept, these research streams have not yet been combined to answer the question of whether AI decision agents are indeed held less responsible and if justice effects are consequently less pronounced for decisions made by AIs.

The present study addresses these research gaps and it contributes to justice and AI literature in several ways. First, we include AIs as decision agents, and we also investigate their influence on justice effects and mediating mechanisms. Consequently, we will provide important knowledge on a job environment that is becoming increasingly popular, and thus follow Johns' (2006) recommendation for more context-sensitive research. Second, we provide, for the first time, an experimental investigation of four parallel mediators of justice effects, which has been demanded in prior research (Colquitt et al., 2013; Colquitt & Zipay, 2015) but not yet realized. We are therefore able to determine the most relevant mediators for specific attitudinal and behavioral outcomes. Third, we empirically investigate judgements of responsibility as an underlying explanation for the differences between human and AI decision agents, and we use research on AI to extend the justice literature on characteristics of the decision agent in general. Fourth, we join the benefits of a controlled experimental environment with a realistic reflection of a situation that is still rarely found in organizational everyday life by using the experimental vignette methodology to investigate employees from various organizations (Aguinis & Bradley, 2014). Finally, we provide a replication of the main findings in our study to show their reproducibility and to heed the call for more replications in IO-Psychology (Kepes & McDaniel, 2013).

How Procedural Justice Works in Organizations

A very important determinant of employee attitudes and behaviors in the context of human decision situations is whether employees perceive decision procedures to be fair (Cohen-Charash & Spector, 2001; Colquitt et al., 2013). Procedural justice refers to certain rules or standards concerning the decision procedure, which lead to employee perceptions of fairness (Goldman & Cropanzano, 2015). A just decision procedure provides a voice during the process or an influence on the outcome and considers further justice criteria (such as consistency, lack of bias, accuracy, correctability, and ethicality; for an overview see: Greenberg, 2011). Perceiving a procedure to be fair not only benefits positive attitudes and behavior (e.g., more job satisfaction, commitment, organizational citizenship behaviors [OCB] and less counterproductive work behaviors [CWB]), it also serves as a buffer when the outcome of a decision is perceived as unfavorable. Employees are less affected by unfavorable outcomes when the decision process is perceived as fair (Brockner et al., 2009).

Justice research over the decades has produced a number of theories that describe how the above described effects of justice on employee attitudes and behavior are transmitted (e.g., uncertainty management theory: Lind & van Den Bos, 2002, affective events theory:

Weiss & Cropanzano, 1996, social exchange theory: Blau, 1964, or the group engagement model: Tyler & Blader, 2003). The four most prominent explaining mechanisms from these theories are positive and negative affect, and trust and identification (Colquitt & Zipay, 2015). In the following, we will briefly introduce each mechanism.

Positive and negative affect are independent dimensions of subjective feeling states (Watson, 2000). Affect was introduced to the justice literature through appraisal theories (Cropanzano et al., 2020; Weiss et al., 1999), which described the interplay of favorability of outcome distribution and procedure of a decision, uncertainty management theory (Lind & van Den Bos, 2002), which proposed mainly negative affect as mediator for injustice effects and affective events theory (Weiss & Cropanzano, 1996), which explains how justice events shape the employee’s attitudes and behavior through positive and negative affect (Judge et al., 2006; Matta et al., 2014). Meta-analytic evidence shows significant average indirect effects of procedural justice on behavioral outcomes through state affect (Colquitt et al., 2013). Several empirical studies have found that negative affect mediates the relationship of justice with CWB (e.g., Matta et al., 2014; VanYperen et al., 2000) and positive affect with OCB (e.g., Lee & Allen, 2002; Yi & Gong, 2008). In the case of attitudes, positive and negative affect were shown to mediate the effects of procedural justice on job satisfaction (Lin, 2015) and commitment (Baccili, 2003).

Trust in the supervisor is the most common indicator of social exchange quality (Colquitt et al., 2014). Social exchange theory describes interactions in relationships as social exchanges, where one interaction partner offers a benefit in exchange for reciprocation from the other (Blau, 1964). Just decisions benefit social exchange quality, resulting in employees who are more likely to reciprocate, e.g., showing more beneficial behavior (Organ, 1990). Meta-analytical investigations show that procedural justice has a significant average indirect effect through social exchange quality on OCB but not on CWB (Colquitt et al., 2013,

p. 217). Trust in the supervisor has additionally been found to mediate the effects of procedural justice on both job satisfaction and commitment (Aryee et al., 2002; Chen et al., 2015; Jiang et al., 2017).

Identification with the group is introduced as explaining mechanisms for justice effects by the group engagement model (Tyler & Blader, 2003): When employees are treated in a just manner, they consequently feel respected because they are proud to belong to a group that treats others fairly. They subsequently develop a stronger identification with this group, and consequently show more cooperative behaviors (e.g., OCB). A number of studies have found this mediating effect for job satisfaction (Ngo et al., 2013; Yuan et al., 2016), commitment (e.g., Ngo et al., 2013; Soenen & Melkonian, 2017), OCB (e.g., Blader & Tyler, 2009; Olkkonen & Lipponen, 2006), and CWB (Ekmekcioglu & Aydogan, 2019).

Even though organizational justice researchers have urgently called for parallel investigation of mediators from different theoretical perspectives (Colquitt et al., 2013), experimental studies that combine two or even more mediators are exceptionally rare. Two meta-analyses included efforts to compare certain mediators of justice effects, comparing social exchange quality and affect (Colquitt et al., 2013) and trust and identification with the organization (Rupp et al., 2014). However, these analyses did not differentiate between positive and negative affect, cannot be used to compare affect and identification, were partly unable to statistically compare the strength of the investigated mediators, and only investigated a limited number of outcomes (mainly focusing on OCB). Because current research does not provide sufficient information to specify hypotheses on which mediator is the strongest mediator for certain justice-outcome relationships, we pose the following research question:

Research Question: Which mediators are the strongest mediators with regard to the indirect effect of procedural justice on employee attitudes and behavior?

How Procedural Justice Works for AI Decisions

The introduction of advanced technologies fitted with AI in the workplace (e.g., assistive or collaborative robots) brings new decision agents into organizations, even at shop floor level. These intelligent technologies are already used to make decisions in organizations (Naim et al., 2016; Wesche & Sonderegger, 2019) and robots fitted with AI will soon be able to closely interact and cooperate with employees (Steil & Maier, 2017, 2020). For researchers and practitioners to be able to suitably react to these changes, we need to investigate differences between human decision agents and these new artificial decision agents and to reassess justice theories in this new context of AI decision making.

Research on interactions between humans and machines (technology such as computers or robots) shows a grave ambiguity concerning the question of the differences between human and machine-like interaction partners. The Computers-Are-Social-Actors Theory (Nass & Moon, 2000) established the hypothesis that people show social reactions to computers to be similar to social reactions to humans by replicating paradigms from social psychology in the context of human-computer interactions. Furthermore, an experimental study could show that employees react to the procedural justice of AI decisions (computer and robot systems) in a similar way as to the procedural justice of human decisions (Ötting & Maier, 2018). Nevertheless, other research suggests that there are differences in how these human-machine interactions are processed. For example, a study in the context of education showed that the effect of procedural fairness perceptions differed significantly between human and AI decisions (Marcinkowski et al., 2020). In addition, research has shown differences in emotional reactions (especially negative affect) and trust towards human and machine-like interaction partners (Visser et al., 2016; Walter et al., 2014). However, Anthropomorphism, which is the tendency to ascribe human-like characteristics to something, therefore seems to partly settle differences between human and AI decision agents; differences seem to be reduced when a technology is more human-like (Kulms & Kopp, 2019).

One possible explanation for these differences in how reactions to justice are transmitted, might be the attribution of responsibility to a decision agent. Organizational justice research has already approached judgements of responsibility using fairness theory (Folger & Cropanzano, 2001). Fairness theory focuses on the implications of responsibility of the decision agent for the perception of fairness and its effect. Three elements are central to responsibility: first, the occurrence of an aversive state and the question of what an alternative state *would* have felt like, if the situation or the decision had been different; second, the discretionary conduct of a person and the question of *could* the person have acted differently and therefore have caused a different situation; and finally, the judgement of moral principles and the question of *should* the person have acted differently. The authors of fairness theory clearly state that: “If no one is to blame, there is no social injustice.” (Folger & Cropanzano, 2001, p. 2). Even though fairness theory has not yet been applied to AI decision agents, the concept of judgements of responsibility has been applied to human-machine interactions through attributional theory (e.g., Weiner, 1995). According to attributional theory, judgements of responsibility are formed through the assessment of causal dimensions such as intentionality and controllability (Weiner, 1995, 2006). Intentionality describes whether the decision was made purposefully or unintended. Controllability describes whether the decision was preventable or inevitable. These judgements form the reaction towards the AIs and their decisions (Britt & Garrity, 2006; Wickens et al., 2011). Given that AIs have started to rapidly grow in importance in organizations, this combination of research on effects of judgements of responsibility on the impact of procedural justice and research on the responsibility of AI is vitally important.

Procedural Justice Mediation Differs Between Decision Agents

In the following section, we will review the theoretical and empirical research that suggested differences between human and AI decision agents concerning the mediation of justice effects. We will then derive hypotheses for affect, trust, and identification. For affect, according to fairness and attributional theory, we need to differentiate between positive and negative affective reactions (Weiner, 1995). Because positive affective reactions are more likely to occur in a fair decision, the affective reaction is more likely to be based on the event and not the agent (Malle & Scheutz, 2014). Therefore, we do not expect to find any differences between the decision agents for positive affect.

Negative affective reactions are more likely to occur in an unfair decision, which presents a violation that is followed by an investigation of the agent's responsibility (Folger & Cropanzano, 2001; Weiner, 1995). Cropanzano et al. (2000) have applied this thought to affective reactions in justice situations. Referring to fairness theory, they state that negative reactions (e.g., anger) only occur when harm has been done, a moral norm has been violated, or when there was an intention to harm. Because AIs are not likely to be held morally responsible or accountable for having the intention to harm (Voiklis et al., 2016), negative affective reactions might be less likely to occur (van der Woerdt & Haselager, 2019).

Concerning differences in trust reactions between human and AI decision agents, empirical studies show that AI decision agents are less likely to be attributed with an intention to harm (or benefit) someone (Voiklis et al., 2016; Xie et al., 2019). Therefore, following social exchange theory (Blau, 1964), people will be unlikely to feel the need to reciprocate towards this decision agent.

Concerning identification with the group, research on group membership is ambiguous. Some studies indicate that robots can be perceived as legitimate group members and in-group robots are evaluated more positively than out-group robots (Häring et al., 2014; Kuchenbrandt et al., 2013). Therefore, reactions of pride about a just decision of an in-group robot and identification with this group might be possible. However, research has also shown that participants are more likely to identify with the group in interactions with human partners than in interactions with artificial partners (Peña et al., 2019). Therefore, we assume that the perception of identification with the group differs between human and AI decision agents. Hence, our first hypothesis states:

> Hypothesis 1: The type of decision agent moderates the mediation of procedural justice effects on employee attitudes and behavior (i.e. job satisfaction, commitment, OCB, CWB). With a human decision agent, the indirect effects of procedural justice on attitudes and behavior through a) negative affect, b) trust, and c) identification is stronger than with an AI decision agent.

The above described combination of fairness theory and attributional theory suggests that judgements of responsibility are the underlying mechanism responsible for differences between human and AI decision agents. The concept of judgements of responsibility can be subdivided into intentionality and controllability: Attributions of responsibility are made when the decision agent had a clear intention to make this particular decision and had control over it (Weiner, 1995, 2006).

AI decision agents are less likely to be held morally responsible or accountable, or perceived as having an intention (Malle et al., 2015; Obhi & Hall, 2011). Furthermore, agents perceived as more in control and as having more intention foster (for example) trust in the agent or a less negative affect (van der Woerdt & Haselager, 2017; Xie et al., 2019). To empirically test intentionality and controllability as an underlying mechanism for differences between human and AI decision agents, we propose the following hypotheses:

> Hypothesis 2: Intentionality moderates the mediation of procedural justice effects on employee attitudes and behavior (i.e. job satisfaction, commitment, OCB, CWB). The higher the perceived intentionality of the decision agent, the stronger are the indirect effects of procedural justice on attitudes and behavior through a) negative affect, b) trust, and c) identification.

> Hypothesis 3: Controllability moderates the mediation of procedural justice effects on employee attitudes and behavior (i.e. job satisfaction, commitment, OCB, CWB). The higher the perceived control of the decision agent over the decision, the stronger are the indirect effects of procedural justice on attitudes and behavior through a) negative affect, b) trust, and c) identification.

The Present Studies

We approached these hypotheses in two consecutive studies. In Study 1, we investigated the research question and the first hypothesis. We had two aims in Study 2: first, we strived to replicate, and thus strengthen, our main findings from Study 1; and second, we extended the first study by investigating Hypotheses 2 and 3 on judgements of responsibility to investigate the underlying reasons for any differences between the decision agents.

Study 1

In the first study, we tested whether the type of decision agents (human vs AI decision agents) influence how procedural justice effects are mediated. In addition, we compared mediator strength and specificity concerning the effects of procedural justice on attitudinal and behavioral outcomes. We manipulated procedural justice and type of decision agent, and we measured job satisfaction, commitment, OCB, and CWB as dependent variables. Positive and negative affect, trust, and identification were investigated as mediators.

Method

Participants and Design

This study was conducted in Germany in 2018. The target population was restricted to employees, recruited from diverse industries through social media. $N = 248$ finished the online experimental vignette study. We had to exclude 19 participants from the sample because they indicated that they were not employed. The final sample consisted of 229 employees (55.5 % female, 2.2 % unknown; $M_{\text{age}} = 32.04$, $SD_{\text{age}} = 11.30$, $min_{\text{age}} = 19$, $max_{age} = 65$). The participants worked an average 34.84 hours per week ($SD = 11.61$, $min = 4$, $max = 70$) and had at least half a year of work experience ($M = 10.38$, $SD = 11.85$, $max = 52$).

Using the experimental vignette methodology, we applied a 2 (fair/unfair) x 3 (human/AI: humanoid robot/AI: machine-like robot) factors between-subjects design. Our vignettes portray hypothetical situations with the aim of manipulating different levels of independent variables. Compared to laboratory experiments, which face the dilemma of sacrificing external for internal validity (Scandura & Williams, 2000), experimental vignettes can enhance internal and external validity simultaneously (Aguinis & Bradley, 2014). Authentic scenarios can be presented to further raise experimental realism, including textual and visual material. Furthermore, experimental vignettes are widely used in organizational justice research (see e.g., Trinkner et al., 2019; Uhde et al., 2020).

Procedure

After reading an informed consent page, the participants were randomly assigned to one of six hypothetical vignettes and asked to read the text thoroughly while imagining themselves in the described situation. Afterwards, the participants indicated the likelihood of engaging in several behaviors or experiencing several sentiments following the imagined situation. All data was obtained online and was fully anonymous. The study was approved by a university's ethics committee (no. 2017-055W).

The six vignettes described an everyday situation of an employee working in a manufacturing team in an automotive company. The decision agent was manipulated as either a human or an AI decision agent, the latter in two robotic variations. The vignettes described that either a human team leader, a humanoid robot (named Pepper), or a machine-like robot (named PX3000) is responsible for the team's organization. This included a decision about the scheduling and allocation of tasks, which was described in more detail. The manipulation was extended by a picture showing two employees in interaction with the respective decision agent. Procedural justice was manipulated by varying the dimensions voice, consistency, bias suppression, correctability, and accuracy in the decision process. The manipulation is based on other vignette studies in organizational justice research (Cornelis et al., 2011; Scott & Colquitt, 2007). Given that recent studies showed gender effects in justice evaluations (Caleo, 2016), we presented either a male or female team leader to each half of the human agent group, both in the vignette text and picture. The original German vignettes, an English translation, and the pictures we used can be found in Figures A1, A2, and A3 in the appendix.

Measures

Instructions, items, and answer scales have been slightly adjusted to address the decision agent in the aforementioned imaginative situation and to instruct participants not to rate their actual job, but to vividly imagine and rate the described situation. If not stated otherwise, we used a five-point answer scale ranging from 1 (*very unlikely*) to 5 (*very likely*). For those instruments where no German translations existed, we translated the items ourselves using the collaborative and iterative translation technique (Douglas & Craig, 2007).

For the manipulation check, procedural justice was tested with seven items of the corresponding subscale of the German version of Colquitt's (2001) scale (Maier et al., 2007), using a five-point answer scale ranging from 1 (*not at all*) to 5 (*entirely*). A sample item is: "Have those procedures been applied consistently?" The human likeness of the decision agent was tested with five items of a semantic differential scale for anthropomorphism, the amount of attributed human likeness (Bartneck et al., 2009). A sample differential ranges from 1 (*fake*) to 5 (*natural*).

Concerning mediators, affect was measured with 10 positive and 10 negative affect descriptions from the German version of the Positive and Negative Affect Schedule (Krohne et al., 1996); for example, "interested" was used for positive and "angry" was used for negative affect. Identification with the work group was measured using five items from a German adaption of the Organizational Commitment Scale (Felfe & Franke, 2012). A sample item is "I am proud to be part of this work group." Trust in the decision agent, was measured with a nine-item subscale from the German Workplace Trust Surveys (Lehmann-Willenbrock & Kauffeld, 2010). A sample item is "The decision agent acts as he/she had promised to act."

Concerning the outcomes, job satisfaction was measured using six items from the German job description inventory (Neuberger & Allerbeck, 1978). The participants answered on a seven-point Kunin scale ranging from 1 (*very dissatisfied*) to 7 (*very satisfied*). A sample item is "How satisfied are you with your colleagues?" Commitment was measured with the German nine-item version of the Organizational Commitment Questionnaire (Maier & Woschée, 2002; Porter & Smith, 1970). A sample item is "I am proud to say that I belong to this company." OCB was measured with 20 items of the corresponding scale of the German

FELA-S (Staufenbiel & Hartz, 2000). A sample item is: "I take the initiative to save the company from possible problems." CWB was measured with the nine-item scale from Robinson and O'Leary-Kelly (1998). A sample item is: "I would work badly, incorrectly, or slowly on purpose."

Statistical Approach

For manipulation checks and hypothesis test, we conducted regression analyses using PROCESS version 3.4 for SPSS (Hayes, 2018) with 95% bias-corrected bootstrap confidence intervals and 10,000 bootstrap samples. We used Model 1 for the manipulation checks, Model 4 for the parallel mediation models, and a custom model for the moderated mediation models (the syntax obtained from the corresponding author).

Results

Preliminary Analyses

To check the manipulation, we conducted regression analyses with procedural justice and type of decision agent as predictors, and perceived procedural justice and anthropomorphism as outcomes. The analyses for perceived procedural justice ($R^2 = .54$) showed significant effects of the procedural justice manipulation ($B = .70$, $SE = .05$, $p < .001$), and an unexpected significant interaction ($B = .24$, $SE = .09$, $p = .010$). Participants in the unfair condition indicated less procedural justice ($M = 1.89$, $SD = .68$) than participants in the fair condition ($M = 3.37$, $SD = .73$). The analyses for anthropomorphism of the decision agent ($R^2 = .35$) showed significant differences between human and AI decision agents ($B = .72$, $SE = .10$, $p < .001$), but not between the robotic variations of the AI decision agent ($B = -.02$, $SE = .14$, $p = .872$). Participants in the human condition indicated more anthropomorphism ($M = 3.05$, $SD = .86$) than participants in the AI conditions (humanoid $M = 2.32$, $SD = .81$, machine-like $M = 2.34$, $SD = .92$). There was an unexpected significant effect of the procedural justice manipulation on anthropomorphism ($B = .34$, $SE = .05$, $p < .001$), as well as a significant interaction ($B = .30$, $SE = .10$, $p = .003$).

Descriptive Statistics

Table 1 shows the distribution of participants, and the means and standard deviations for mediators and outcomes in the experimental groups. Table 2 shows the reliabilities and intercorrelations of all measured variables.

-------insert Tables 1 and 2 approximately here--------

Mediation of Procedural Justice Effects

We conducted regression analyses for parallel mediation models. The results showed significant parallel mediation models for job satisfaction ($R^2 =.72$, $p < .001$), commitment ($R^2 = .64$, $p < .001$), OCB ($R^2 = .28$, $p < .001$), and CWB ($R^2 = .09$, $p = .001$).

The total and specific indirect effects (Table 3) showed which mediators transmitted the effects of procedural justice on the outcomes. A confidence interval of the indirect effect that does not include zero denotes significance. The effect of procedural justice on job satisfaction was significantly mediated by all four mediators. The effect on commitment was significantly mediated by positive affect, trust, and identification, but not by negative affect. The effect on OCB was significantly mediated by negative affect and identification. The effect on CWB was only significantly mediated by negative affect.

-------insert Table 3 approximately here--------

To compare the indirect effects of the four mediators for each outcome and to answer our research question, we examined contrasts (i.e., differences between absolute values of two indirect effects: Hayes, 2018) and compared significant mediators. A confidence interval that does not include zero shows that the contrast ($C = |a_ib_i|-|a_jb_j|$) is significantly different from zero. For job satisfaction, trust was the strongest mediator; with significantly different indirect effects compared to negative affect ($C = -.30$, CI = [-.43, -.17]), positive affect ($C = -$

.32, CI = [-.45, -.20]), and identification (*C* = .29, CI = [.16, .41]). For commitment, indirect effects through trust were significantly higher than those through positive affect (*C* = -.09, CI = [-.17, -.003]), while identification did not differ from trust or positive affect. For OCB, the contrasts did not show any significant differences between significant mediators. For CWB, the only significant mediator was negative affect.

Moderated Mediation of Procedural Justice Effects

To test the moderation hypothesis, we conducted regression analyses for conditional indirect effects. In Hypothesis 1, we predicted that the type of decision agent moderates the indirect effect of procedural justice on employee attitudes and behavior through a) negative affect, b) trust, and c) identification. The comparisons showed that only the indirect effect through trust significantly differed between human and AI decision agents (for job satisfaction: *B* = .30, 95% CI = [.16, .46], and for commitment: *B* = .12, CI = [.05, .20]; indirect effects as well as all pairwise comparisons for the moderator groups can be found in Tables 4 and 5. The relationship between procedural justice and trust was weaker when AIs make the decision compared to humans. This confirms H_{1b}, yet cannot confirm H_{1a} and H_{1c}.

-------insert Tables 4 and 5 approximately here--------

Discussion

In this first of the two experiments, we could show that the strength of mediators is specific for certain procedural justice effects on attitudes and behaviors. With regard to effects on satisfaction, trust was the strongest mediator. With regard to commitment, trust was a stronger mediator than positive affect, yet indirect effects through identification were similarly high than those through trust. With regard to OCB, the two significant mediators (negative affect and identification) did not differ in strength. With regard to CWB, the only significant mediator was negative affect, and therefore was clearly the strongest mediator. In summary, the results show that trust was the best explaining mechanism for the effects of procedural justice on attitudes. With regard to effects on behaviors, negative affect was most influential.

Concerning the hypothesis on differences between human and AI decision agents, the results show differences concerning indirect effects through trust but not through identification or negative affect. In line with our assumptions, employees seem to have more difficulties trusting in AI decision agents than in human decision agents, even when fair decisions are made. Contrary to our assumptions, employees might identify with a group where an AI makes decisions in the same way as with a group where a human makes the decision. In addition, the perception of unfair procedures leads to negative affective reactions, whether the agent is human or AI. Because these results were surprising, we decided to conduct a second study. In this second study, we additionally examined whether judgements of responsibility could explain the differences between humans and AI as decision agents.

Study 2

In the second study, we aimed to replicate the main results from the first study. Consequently, we manipulated procedural justice and type of decision agent; we measured job satisfaction, commitment, OCB, and CWB as dependent variables; and we measured affect, trust, and identification as mediators. To extend the results from Study 1 and to find explanations for surprising results, we tested whether the two dimensions of judgements of responsibility (i.e., intentionality and controllability) moderated the indirect effects of procedural justice on the outcomes.

Method

Participants

The study was conducted in Germany in 2019. Target participants were employees, as in Study 1, recruited through social media. 145 participants finished this study. However, 13 participants had to be excluded because they indicated that they were not employed. The final sample consisted of 132 employees (55.3 % female; $M_{\text{age}} = 27.76$, $SD_{\text{age}} = 7.21$, $min_{\text{age}} = 18$, $max_{\text{age}} = 56$). The participants worked an average 35.15 hours per week ($SD = 11.81$, $min = 8$, $max = 60$) and had at least four months of work experience ($M = 5.35$, $SD = 6.02$, $max = 38$).

Design, Procedure, and Measures

The second study was designed and conducted in the same way as Study 1. We used the same procedure, vignettes, and measures, with the exception of adding measures for the dimensions of judgements of responsibility (approved by a university's ethics committee, approval no. 2019-090). We measured intentionality and controllability with two and four items, which we adapted from the corresponding subscales from (Wickens et al., 2011). Because no German translations are available, we translated the items ourselves; following Douglas and Craig (2007). Sample items are "The decision agent intended this decision.", and "The decision was not controllable for the decision agent."

Results

Preliminary Analyses

The manipulation check for perceived procedural justice ($R^2 = .52$) showed significant effects of the procedural justice manipulation ($B = .71$, $SE = .06$, $p < .001$). Participants in the unfair condition indicated less procedural justice ($M = 2.11$, $SD = .71$) than participants in the fair condition ($M = 3.53$, $SD = .70$). The analyses for perceived anthropomorphism of the decision agent ($R^2 = .41$) showed significant differences between human and AI ($B = 1.13$, $SE = .14$, $p < .001$), but not between robotic variations. Participants in the human condition indicated more anthropomorphism ($M = 3.34$, $SD = 0.93$) than participants in the AI conditions (humanoid $M = 2.27$, $SD = .78$, machine-like $M = 2.10$, $SD = .60$). There was a significant effect of the procedural justice manipulation on anthropomorphism ($B = .23$, $SE = .06$, $p = .001$), as well as a significant interaction ($B = .40$, $SE = .14$, $p = .005$).

As a pretest for Hypotheses 2 and 3, we conducted regression analyses with contrasts to test whether intentionality and controllability differed between decision agents. Both intentionality (B = .18, SE = .06, p = .002), and controllability (B = .26, SE = .06, p = .003) differ significantly depending on human or AI decision agents, but not between robots (intentionality: B = .10, SE = .09, p = .309; controllability: B = .05, SE = .09, p = .627).

Descriptive Statistics

Tables 1 and 2 show the distribution of participants, means, and standard deviations, reliabilities and correlations of all measured variables. Most values are comparable to Study 1, only the correlations with CWB are slightly higher in Study 2.

Mediation of Procedural Justice Effects

The results for the parallel mediation models could be replicated (job satisfaction: R^2 =.69, p < .001; commitment: R^2 = .70, p < .001; OCB: R^2 = .23, p < .001; CWB: R^2 = .21, p < .001). The results of the specific indirect effects (Table 3) could generally be replicated with the exception of the indirect effects for job satisfaction and OCB through identification.

The results of the contrasts to compare indirect effects of the four mediators for each outcome could partially be replicated. For job satisfaction and trust as mediator, we could replicate the findings from Study 1 for positive affect (C = -.25, CI = [-.47, -.01]) and identification (C = .29, CI = [.09, .49]). However, the comparison with negative affect was not significant. For commitment, we could not replicate the significant difference between trust and positive affect. The results for OCB and CWB could be replicated.

Moderated Mediation of Procedural Justice Effects

We could replicate the findings from Study 1 for the effects on job satisfaction (.17, CI = [.03, .35]), and for commitment (.11, CI = [.02, .23]) through trust. The difference test for the moderation of the indirect effect through identification was also significant (for commitment: .12, CI = [.02, .27]). This confirms H_{1b} and H_{1c}, but we cannot yet confirm H_{1a}.

Indirect effects and all pairwise comparisons for the moderator groups can be found in Tables 6 and 7.

-------insert Tables 6 and 7 approximately here--------

In Hypotheses 2 and 3, we predicted that intentionality and controllability moderate the mediation of procedural justice effects on employee attitudes and behavior, through negative affect, trust, and identification. To test these hypotheses, we conducted a regression analysis for conditional indirect effects for each outcome (Tables 8 and 9) and we interpreted the difference tests of the indirect effects (Table 10).

For intentionality, there were significant differences for all significant indirect effects, which confirms H_{2a}, H_{2b}, and H_{2c}. The conditional effects of procedural justice on the mediators were higher for high levels of perceived intentionality than for lower levels, for negative affect, for trust, and for identification (see Table 8).

For controllability, there were significant differences for all significant indirect effects through negative affect but not through trust and identification, which confirms H_{3a} but cannot confirm H_{3b}, and H_{3c}. The conditional effects were higher for high levels of perceived controllability than for lower levels, for negative affect, trust, and for identification (see Table 9).

----------insert Tables 8, 9, and 10 approximately here-----------

Discussion

In general, we were able to replicate the results from Study 1, with three exceptions. First, we could not replicate identification as a mediator of procedural justice effects on job satisfaction and OCB. Nevertheless, given that both experiments showed that other mediators were stronger than identification (trust and positive affect for attitudes and negative affect for behaviors), the overall indication stays the same. Second, differences between strength of trust and positive affect for commitment and trust and negative affect for job satisfaction could not be replicated. However, looking at the general picture, trust seems to be a highly important mediator for effects on attitudes. Finally, in addition to the significant differences between decision agents for the mediator trust in Study 1, we were also able to show differences for identification.

The additional investigation of the two dimensions of judgements of responsibility showed that intentionality moderated the indirect effects of procedural justice through all three assumed mediators. Controllability only moderated the indirect effects of procedural justice through negative affect. These results show that judgements of responsibility, especially intentionality, can be confirmed as one of the mechanisms that are responsible for differences between human and AI as decision agents.

General Discussion

In this two-experiment study, our aim was to examine mediating mechanisms of procedural justice effects in the context of AI and human decision agents. More specifically, we investigated differences between decision agents (human vs. AI) concerning justice effects and mediating mechanisms as well as differences in strength and specificity of mediators from prominent justice theories. In addition, we investigated judgements of responsibility as one possible explaining characteristic that might lead to differences between decision agents.

Both experiments showed that the effects of procedural justice on attitudes and behavior are explained by specific mediators. Overall, for attitudes, trust was the strongest mediator; whereas for behaviors, negative affect was the strongest mediator. Concerning differences between human and AI decision agents, trust is a weaker mediator for AI decision agents; whereas no differences occurred for negative affect as mediator. Our experiments yielded ambiguous results for identification because Study 2 confirmed differences while Study 1 did not. Finally, we could confirm the intentionality and controllability of a decision as underlying mechanisms for differences between human and AI decision agents. The employees perceive AIs as less intentional and in control of decisions. Therefore, effects of procedural justice on negative affect, identification, and trust are less pronounced.

Theoretical Implications

Our study shows that decision agents are an important factor for studying justice effects, which have not received enough attention. Our experiments show that it is not only important if the decision is fair but it is also relevant who made the decision. However, justice theories do not yet account for this in a sufficient manner. Researching the source of justice is underrepresented in empirical studies and theories do not incorporate characteristics of a decision agent that might be responsible for differences in how justice works (Rupp et al., 2014). In the present study, we investigated intentionality and controllability in a first attempt to find characteristics that explain differences between agents. Our results show that it is particularly important for a decision agent to be perceived as having made a decision intentionally. However, because this only partially explained the differences between human and AI decision agents, there are likely to be other characteristics of AI decision agents that can cause or compensate differences in how justice works.

Most justice theories propose explanations for justice effects on behaviors; for example, exchange behavior such as OCB in social exchange theory (Blau, 1964) or cooperation in the group engagement model (Tyler & Blader, 2003). However, they often simply assume that the same also holds true for effects on attitudes. Our results indicate that, depending on the outcome in question, mediators from different justice theories explain the effects of procedural justice and that a much more differentiated investigation of mediators and outcomes is needed. Consequently, future justice research needs not only to finally make an effort to compare and combine different theoretical approaches, as previous meta-analyses already demanded years ago (Colquitt et al., 2013), but it also needs more empirical and meta-analytical studies to investigate specific direct and indirect effects on attitudes.

Practical Implications

The results of the present study provide important information for organizations that aim to implement AI for decision making and for system designers who want to design accepted AIs. Organizations should make sure that any decision is made through just procedures, that are communicated via organizational policies and culture, and that decision agents (i.e., AIs just as humans) are trained or designed to decide in a just manner (Skarlicki & Latham, 2013; Whitman et al., 2012). In addition, communication of decision procedures has to clearly expose the responsibility for a decision. Our study showed that it is especially important for the decision agent to convey an impression of purposefully intending a decision. To achieve this, the communication of decisions needs to transparently and clearly describe how, why, and by whom the decision was made. Therefore, organizational representatives should carefully consider how decision procedures and communication are designed, especially before deciding which AI to implement as a decision agent.

System designers have to factor in justice rules early in the AI design process because implementing just procedures into AI decision making goes beyond the mere implementation of text modules to communicate a decision. Consequently, system designers need to find new ways to be able to consider employee opinions and preferences, and make sure that the decisions are unbiased, correctable, accurate, consistent, and explainable (Töniges et al., 2017). The need to design AIs that are perceived as having intent and control over the decision is an even more complex endeavor because research that investigates the factors that foster the attribution of intent and responsibility is rare. A possible explanation might be the attributed competence of the decision agent. AI decision agents might not be perceived as competent enough to make a certain decision (in the sense of professional, social, and moral competence) (Bartneck et al., 2009; Malle, 2016; Scheunemann et al., 2020). van der Woerdt and Haselager (2017), for example, showed that a perceived lack of ability of a robot reduced feelings of disappointment and blame towards it.

Limitations and Directions for Further Research

Experimental research is often criticized as lacking external validity, which ultimately compromises the generalizability of the results (Scandura & Williams, 2000). However, to investigate causal relationships and rule out alternative explanations, experimental research is essential. Therefore, we chose to use an experimental vignette methodology, which offers a unique way to retain experimental benefits while at the same time maximizing realism and external validity (Aguinis & Bradley, 2014). To provide further realism, we applied thoroughly designed vignettes that incorporated visual material, and which described vivid and detailed situations. We also chose a context that is likely to be part of any employee's life (i.e., a weekly routine meeting) and we only included employees, who are likely to have experienced such a situation. Nevertheless, more extensive research is needed to investigate real interactions over a longer time period in longitudinal field studies or Wizard-of-Oz experiments (see e.g., Dahlbäck et al., 1993; Riek, 2012), which will be able to go beyond a first impression.

A second limitation of our study is that there may be a potential common method bias because self-report measures were used to assess the majority of variables (Podsakoff et al., 2003). However, this cannot be applied to the investigated main and moderator effects because we experimentally manipulated the independent variables. Regarding the mediator effects and the moderation effect of judgements of responsibility, the experimental vignette methodology circumvented the use of different sources of information. Even though different sources are usually recommended to counter common method bias (Podsakoff et al., 2012), this method is not applicable when all questions concern the employees' perceptions in a hypothetical situation. In cases where the targeted information are perceptions rather than objective data, self-report measures are seen as appropriate (Conway & Lance, 2010).

Another promising direction for further research might be to examine the other justice dimensions (e.g., informational justice). Given that one of the major challenges in AI development is the explainability of decisions (Barredo Arrieta et al., 2020), informational justice might be hard to achieve (Töniges et al., 2017) and differences between decision agents might be more salient. Finally, to strengthen generalizability across different situations and recipients, other decision contexts should be explored (e.g., robotics in care, where both care takers and patients are decision recipients; or AI in service, where customers are recipients).

Conclusion

This study is one of the first to provide results that make the direct comparison of strength and specificity of mediators of procedural justice effects possible and which examines the context of AI decision agents in organizations. Our results show that a differentiated view on mediating mechanisms of procedural justice effects is needed. They also confirm that judgements of responsibility are an important factor to explain how and when procedural justice works. In particular, the perception of the intentionality of the decision agent is needed if procedural justice is to unfold its full benefits for positive employee attitudes and behavior. When researchers, practitioners, and AI designers have justice at the back of their mind, the emergence of AI as a decision agent might deliver a much more positive future than one might at first imagine.

References

Aguinis, H., & Bradley, K. J. (2014). Best practice recommendations for designing and implementing experimental vignette methodology studies. *Organizational Research Methods*, *17*(4), 351–371. https://doi.org/10.1177/1094428114547952

Aryee, S., Budhwar, P. S., & Chen, Z. X. (2002). Trust as a mediator of the relationship between organizational justice and work outcomes: Test of a social exchange model. *Journal of Organizational Behavior*, *23*(3), 267–285. https://doi.org/10.1002/job.138

Baccili, P. A. (2003). Effects of company manager and manager psychological contract violation on justice, negative affect and commitment. *Academy of Management Proceedings*, *2003*(1), D1-D6. https://doi.org/10.5465/ambpp.2003.13792579

Barredo Arrieta, A., Díaz-Rodríguez, N., Del Ser, J., Bennetot, A., Tabik, S., Barbado, A., Garcia, S., Gil-Lopez, S., Molina, D., Benjamins, R., Chatila, R., & Herrera, F. (2020). Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. *Information Fusion*, *58*, 82–115. https://doi.org/10.1016/j.inffus.2019.12.012

Bartneck, C., Kulić, D., Croft, E., & Zoghbi, S. (2009). Measurement instruments for the anthropomorphism, animacy, likeability, perceived intelligence, and perceived safety of robots. *International Journal of Social Robotics*, *1*(1), 71–81. https://doi.org/10.1007/s12369-008-0001-3

Blader, S. L., & Tyler, T. R. (2009). Testing and extending the group engagement model: Linkages between social identity, procedural justice, economic outcomes, and extrarole behavior. *Journal of Applied Psychology*, *94*(2), 445–464. https://doi.org/10.1037/a0013935

Blau, P. M. (1964). *Exchange and power in social life*. John Wiley & Sons Inc.

Britt, T. W., & Garrity, M. J. (2006). Attributions and personality as predictors of the road rage response. *British Journal of Social Psychology*, *45*(1), 127–147. https://doi.org/10.1348/014466605X41355

Brockner, J., & Wiesenfeld, B. M. (2020). Organizational justice is alive and well and living elsewhere (but not too far away). In E. A. Lind (Ed.), *Frontiers in social psychology. Social psychology and justice* (pp. 213–242). Routledge.

Brockner, J., Wiesenfeld, B. M., & Diekmann, K. A. (2009). Towards a "fairer" conception of process fairness: Why, when and how more may not always be better than less. *The Academy of Management Annals*, *3*(1), 183–216. https://doi.org/10.1080/19416520903047228

Caleo, S. (2016). Are organizational justice rules gendered? Reactions to men's and women's justice violations. *Journal of Applied Psychology*, *101*(10), 1422–1435. https://doi.org/10.1037/apl0000131

Chen, S.-Y., Wu, W.-C., Chang, C.-S., Lin, C.-T., Kung, J.-Y., Weng, H.-C., Lin, Y.-T., & Lee, S.-I. (2015). Organizational justice, trust, and identification and their effects on organizational commitment in hospital nursing staff. *BMC Health Services Research*, *15*, 363. https://doi.org/10.1186/s12913-015-1016-8

Cohen-Charash, Y., & Spector, P. E. (2001). The role of justice in organizations: A meta-analysis. *Organizational Behavior and Human Decision Processes*, *86*(2), 278–321. https://doi.org/10.1006/obhd.2001.2958

Colquitt, J. A. (2001). On the dimensionality of organizational justice: A construct validation of a measure. *Journal of Applied Psychology*, *86*(3), 386–400. https://doi.org/10.1037/0021-9010.86.3.386

Colquitt, J. A., Baer, M. D., Long, D. M., & Halvorsen-Ganepola, M. D. K. (2014). Scale indicators of social exchange relationships: A comparison of relative content validity. *Journal of Applied Psychology*, *99*(4), 599–618. https://doi.org/10.1037/a0036374

Colquitt, J. A., Scott, B. A., Rodell, J. B., Long, D. M., Zapata, C. P., Conlon, D. E., & Wesson, M. J. (2013). Justice at the millennium, a decade later: A meta-analytic test of social exchange and affect-based perspectives. *Journal of Applied Psychology*, *98*(2), 199–236. https://doi.org/10.1037/a0031757

Colquitt, J. A., & Zipay, K. P. (2015). Justice, fairness, and employee reactions. *Annual Review of Organizational Psychology and Organizational Behavior*, *2*(1), 75–99. https://doi.org/10.1146/annurev-orgpsych-032414-111457

Conway, J. M., & Lance, C. E. (2010). What reviewers should expect from authors regarding common method bias in organizational research. *Journal of Business and Psychology*, *25*(3), 325–334. https://doi.org/10.1007/s10869-010-9181-6

Cornelis, I., van Hiel, A., & Cremer, D. de (2011). Birds of a feather: Leader-follower similarity and procedural fairness effects on cooperation. *European Journal of Work and Organizational Psychology*, *20*(3), 388–415. https://doi.org/10.1080/13594321003630055

Cropanzano, R. S., Ambrose, M. L., & van Wagoner, P. (2020). Organizational justice and workplace emotion. In E. A. Lind (Ed.), *Frontiers in social psychology. Social psychology and justice.* Routledge.

Cropanzano, R. S., Weiss, H. M., Suckow, K. J., & Grandey, A. A. (2000). Doing justice to workplace emotion. In N. M. Ashkanasy, C. E. Härtel, & W. J. Zerbe (Eds.), *Emotions in the workplace: Research, theory, and practice* (pp. 49–62). Quorum Books/Greenwood Publishing Group.

Dahlbäck, N., Jönsson, A., & Ahrenberg, L. (1993). Wizard of Oz studies — why and how. *Knowledge-Based Systems*, *6*(4), 258–266. https://doi.org/10.1016/0950-7051(93)90017-N

Douglas, S. P., & Craig, C. S. (2007). Collaborative and iterative translation: An alternative approach to back translation. *Journal of International Marketing*, *15*(1), 30–43. https://doi.org/10.1509/jimk.15.1.030

Ekmekcioglu, E. B., & Aydogan, E. (2019). A moderated mediation model of the relationship between organizational justice and turnover intention. *International Journal of Organizational Analysis.* Advance online publication. https://doi.org/10.1108/IJOA-06-2018-1448

Felfe, J., & Franke, F. (2012). *COMMIT Commitment Skalen: Fragebogen zur Erfassung von Commitment gegenüber Organisation, Beruf/Tätigkeit, Team, Führungskraft und Beschäftigungsform* [COMMIT Commitment Scales: A questionaire to assess commitment with organization, job, team, supervisor, and form of employment]. Hans Huber.

Folger, R., & Cropanzano, R. S. (2001). Fairness theory: Justice as accountability. In J. Greenberg & R. S. Cropanzano (Eds.), *Advances in organizational justice* (pp. 1–55). Stanford University Press.

Franklin, S., Madl, T., D'Mello, S., & Snaider, J. (2014). LIDA: A systems-level architecture for cognition, emotion, and learning. *IEEE Transactions on Autonomous Mental Development*, *6*(1), 19–41. https://doi.org/10.1109/TAMD.2013.2277589

Goldman, B. M., & Cropanzano, R. S. (2015). "Justice" and "fairness" are not the same thing. *Journal of Organizational Behavior*, *36*(2), 313–318. https://doi.org/10.1002/job.1956

Greenberg, J. (1987). A taxonomy of organizational justice theories. *Academy of Management Review*, *12*(1), 9–22. https://doi.org/10.5465/amr.1987.4306437

Greenberg, J. (2011). Organizational justice: The dynamics of fairness in the workplace. In S. Zedeck (Ed.), *APA handbook of industrial and organizational psychology: Vol. 3. Maintaining, expanding, and contracting the organization* (pp. 271–327). American Psychological Association. https://doi.org/10.1037/12171-008

Häring, M., Kuchenbrandt, D., & André, E. (2014). Would you like to play with me? How robots' group membership and task features influence human-robot interaction. In G. Sagerer, M. Imai, T. Belpaeme, & A. Thomaz (Eds.), *HRI'14, HRI '14: Proceedings of the 9th 2014 ACM/IEEE International Conference on Human-Robot Interaction* (pp. 9–16). ACM Press.

Hayes, A. F. (2018). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach* (Second edition). *Methodology in the social sciences*. The Guilford Press. https://ebookcentral.proquest.com/lib/potsdamuni/detail.action?docID=5109647

Jiang, Z., Gollan, P. J., & Brooks, G. (2017). Relationships between organizational justice, organizational trust and organizational commitment: Across-cultural study of China, South Korea and Australia. *The International Journal of Human Resource Management*, *28*(7), 973–1004. https://doi.org/10.1080/09585192.2015.1128457

Johns, G. (2006). The essential impact of context on organizational behavior. *Academy of Management Review*, *31*(2), 386–408. https://doi.org/10.5465/amr.2006.20208687

Judge, T. A., Scott, B. A., & Ilies, R. (2006). Hostility, job attitudes, and workplace deviance: Test of a multilevel model. *Journal of Applied Psychology*, *91*(1), 126–138. https://doi.org/10.1037/0021-9010.91.1.126

Kepes, S., & McDaniel, M. A. (2013). How trustworthy is the scientific literature in industrial and organizational psychology? *Industrial and Organizational Psychology*, *6*(3), 252–268. https://doi.org/10.1111/iops.12045

Krohne, H. W., Egloff, B., Kohlmann, C.-W., & Tausch, A. (1996). Untersuchungen mit einer deutschen Version der "Positive and Negative Affect Schedule" (PANAS) [Investigation of a German version of the Positive and Negative Affect Schedule (PANAS)]. *Diagnostica*, *42*(2), 139–156.

Kuchenbrandt, D., Eyssel, F., Bobinger, S., & Neufeld, M. (2013). When a robot's group membership matters. *International Journal of Social Robotics*, *5*(3), 409–417. https://doi.org/10.1007/s12369-013-0197-8

Kulms, P., & Kopp, S. (2019). More human-likeness, more trust? In F. Alt, A. Bulling, & T. Döring (Eds.), *ICPS: ACM international conference proceeding series, Muc '19: Tagungsband Mensch und Computer 2019* (pp. 31–42). The Association for Computing Machinery, Inc. https://doi.org/10.1145/3340764.3340793

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*(4), 312–323. https://doi.org/10.1111/ijsa.12150

Larson, L., & DeChurch, L. A. (2020). Leading teams in the digital age: Four perspectives on technology and what they mean for leading teams. *The Leadership Quarterly*, *31*(1), 101377. https://doi.org/10.1016/j.leaqua.2019.101377

Lee, K., & Allen, N. J. (2002). Organizational citizenship behavior and workplace deviance: The role of affect and cognitions. *Journal of Applied Psychology*, *87*(1), 131–142. https://doi.org/10.1037//0021-9010.87.1.131

Lehmann-Willenbrock, N., & Kauffeld, S. (2010). Development and construct validation of the German Workplace Trust Survey (G-WTS). *European Journal of Psychological Assessment*, *26*(1), 3–10. https://doi.org/10.1027/1015-5759/a000002

Lin, X. (2015). How does procedural justice climate influence individual outcomes? An affective perspective. *Asia Pacific Journal of Management*, *32*(3), 771–800. https://doi.org/10.1007/s10490-015-9421-4

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

Maier, G. W., Streicher, B., Jonas, E., & Woschée, R.-M. (2007). Gerechtigkeitseinschätzungen in Organisationen [Assessment of justice in organizations]. *Diagnostica*, *53*, 97–108. https://doi.org/10.1026/0012-1924.53.2.97

Maier, G. W., & Woschée, R.-M. (2002). Die affektive Bindung an das Unternehmen: Psychometrische Überprüfung einer deutschprachigen Fassung des Organizational Commitment Questionnaire (OCQ) von Porter und Smith (1970) [Affective commtiment to the organization: Psychometic assessment of a German version of the Organizational Commitment Questionnaire (OCQ)]. *Zeitschrift Für Arbeits- Und Organisationspsychologie A&O*, *46*, 126–136. https://doi.org/10.1026//0932-4089.46.3.126

Malle, B. F., & Scheutz, M. (2014). Moral competence in social robots. In *ETHICS '14, ETHICS '14: Proceedings of the IEEE 2014 International Symposium on Ethics in Engineering, Science, and Technology* (8:1-8:6). IEEE Press. http://dl.acm.org/citation.cfm?id=2960587.2960597

Malle, B. F., Scheutz, M., Arnold, T., Voiklis, J., & Cusimano, C. (2015). Sacrifice one for the good of many? In J. A. Adams, W. Smart, B. Mutlu, & L. Takayama (Eds.), *HRI '15: Proceedings of the tenth annual ACM/IEEE International Conference on Human-Robot Interaction* (pp. 117–124). ACM Press. https://doi.org/10.1145/2696454.2696458

Marcinkowski, F., Kieslich, K., Starke, C., & Lünich, M. (2020). Implications of AI (un-)fairness in higher education admissions. In M. Hildebrandt, C. Castillo, E. Celis, S. Ruggieri, L. Taylor, & G. Zanfir-Fortuna (Eds.), *FAT '20: Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency* (pp. 122–130). ACM. https://doi.org/10.1145/3351095.3372867

Matta, F. K., Erol-Korkmaz, H. T., Johnson, R. E., & Biçaksiz, P. (2014). Significant work events and counterproductive work behavior: The role of fairness, emotions, and emotion regulation. *Journal of Organizational Behavior*, *35*(7), 920–944. https://doi.org/10.1002/job.1934

Naim, I., Tanveer, M. I., Gildea, D., & Hoque, E. (2016). Automated analysis and prediction of job interview performance. *IEEE Transactions on Affective Computing*, *9*(1), 191–204. https://doi.org/10.1109/TAFFC.2016.2614299

Nass, C. I., & Moon, Y. (2000). Machines and mindlessness: Social responses to computers. *Journal of Social Issues*, *56*(1), 81–103. https://doi.org/10.1111/0022-4537.00153

Neuberger, O., & Allerbeck, M. (1978). *Messung und Analyse von Arbeitszufriedenheit: Erfahrungen mit dem „Arbeitsbeschreibungsbogen (ABB)"* [Measurement and analysis of job satisfaction: Experiences with the „Arbeitsbeschreibungsbogen (ABB)"]. *Schriften zur Arbeitspsychologie: Nr. 26*. Huber.

Ngo, H.-y., Loi, R., Foley, S., Zheng, X., & Zhang, L. (2013). Perceptions of organizational context and job attitudes: The mediating effect of organizational identification. *Asia Pacific Journal of Management*, *30*(1), 149–168. https://doi.org/10.1007/s10490-012-9289-5

Obhi, S. S., & Hall, P. (2011). Sense of agency in joint action: Influence of human and computer co-actors. *Experimental Brain Research*, *211*(3-4), 663–670. https://doi.org/10.1007/s00221-011-2662-7

Olkkonen, M.-E., & Lipponen, J. (2006). Relationships between organizational justice, identification with organization and work unit, and group-related outcomes. *Organizational Behavior and Human Decision Processes*, *100*(2), 202–215. https://doi.org/10.1016/j.obhdp.2005.08.007

Organ, D. W. (1990). The motivational basis of organizational citizenship behavior. *Research in Organizational Behavior*, *12*(1), 43–72.

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

Peña, J., Ghaznavi, J., Brody, N., Prada, R., Martinho, C., Santos, P. A., Damas, H., & Dimas, J. (2019). Effects of human vs. computer-controlled characters and social identity cues on enjoyment. *Journal of Media Psychology*, *31*(1), 35–47. https://doi.org/10.1027/1864-1105/a000218

Podsakoff, P. M., MacKenzie, S. B., Lee, J.-Y., & Podsakoff, N. P. (2003). Common method biases in behavioral research: A critical review of the literature and recommended remedies. *Journal of Applied Psychology*, *88*(5), 879–903. https://doi.org/10.1037/0021-9010.88.5.879

Podsakoff, P. M., MacKenzie, S. B., & Podsakoff, N. P. (2012). Sources of method bias in social science research and recommendations on how to control it. *Annual Review of Psychology*, *63*, 539–569. https://doi.org/10.1146/annurev-psych-120710-100452

Porter, L. W., & Smith, F. J. (1970). *The etiology of organizational commitment*. University of California, Irvine, CA.

Riek, L. (2012). Wizard of Oz studies in HRI: A systematic review and new reporting guidelines. *Journal of Human-Robot Interaction*, 119–136. https://doi.org/10.5898/JHRI.1.1.Riek

Robert, L. P., Pierce, C., Marquis, L., Kim, S., & Alahmad, R. (2020). Designing fair AI for managing employees in organizations: A review, critique, and design agenda. *Human–Computer Interaction*, *17*(2), 1–31. https://doi.org/10.1080/07370024.2020.1735391

Robinson, S. L., & O'Leary-Kelly, A. M. (1998). Monkey see, monkey do: The influence of work groups on the antisocial behavior of employees. *Academy of Management Journal*, *41*(6), 658–672. https://doi.org/10.5465/256963

Rupp, D. E., Shao, R., Jones, K. S., & Liao, H. (2014). The utility of a multifoci approach to the study of organizational justice: A meta-analytic investigation into the consideration of normative rules, moral accountability, bandwidth-fidelity, and social exchange. *Organizational Behavior and Human Decision Processes*, *123*(2), 159–185. https://doi.org/10.1016/j.obhdp.2013.10.011

Scandura, T. A., & Williams, E. A. (2000). Research methodology in management: Current practices, trends, and implications for future research. *Academy of Management Journal*, *43*(6), 1248–1264. https://doi.org/10.5465/1556348

Scott, B. A., & Colquitt, J. A. (2007). Are organizational justice effects bounded by individual differences? An examination of equity sensitivity, exchange ideology, and the Big Five. *Group & Organization Management*, *32*(3), 290–325. https://doi.org/10.1177/1059601106286877

Skarlicki, D. P., & Latham, G. P. (2013). How can training be used to foster organizational justice? In J. Greenberg & J. A. Colquitt (Eds.), *Handbook of organizational justice* (pp. 499–522). Taylor and Francis.

Soenen, G., & Melkonian, T. (2017). Fairness and commitment to change in M&As: The mediating role of organizational identification. *European Management Journal*, *35*(4), 486–492. https://doi.org/10.1016/j.emj.2016.10.005

Staufenbiel, T., & Hartz, C. (2000). Organizational Citizenship Behavior: Entwicklung und erste Validierung eines Meßinstruments [Development and initial validation of a measurement instrument]. *Diagnostica*, *46*(2), 73–83. https://doi.org/10.1026//0012-1924.46.2.73

Steil, J. J., & Maier, G. W. (2017). Robots in the digitalized workplace. In G. Hertel, D. L. Stone, R. D. Johnson, & J. Passmore (Eds.), *Wiley Blackwell handbooks in organizational* *psychology: Vol. 7696. The Wiley Blackwell handbook of the psychology of the internet at work* (pp. 403–422). John Wiley & Sons Inc.

Steil, J. J., & Maier, G. W. (2020). Kollaborative Roboter: Universale Werkzeuge in der digitalisierten und vernetzten Arbeitswelt [Collaborative robots: universal tools in the digitalized and interconnected working environment]. In G. W. Maier, G. Engels, & E. Steffen (Eds.), *Springer reference. Handbuch Gestaltung digitaler und vernetzter Arbeitswelten* (pp. 323–346). Springer. https://doi.org/10.1007/978-3-662-52979-9_15

Töniges, T., Ötting, S. K., Wrede, B., Maier, G. W., & Sagerer, G. (2017). An emerging decision authority: Adaptive cyber-physical system design for fair human-machine interaction and decision processes. In H. Song, D. B. Rawat, S. Jeschke, & C. Brecher (Eds.), *Intelligent data-centric systems: Sensor collected intelligence. Cyber-physical systems: Foundations, principles and applications* (pp. 419–430). Elsevier Academic Press.

Trinkner, R., Mays, R. D., Cohn, E. S., van Gundy, K. T., & Rebellon, C. J. (2019). Turning the corner on procedural justice theory: Exploring reverse causality with an experimental vignette in a longitudinal survey. *Journal of Experimental Criminology*, *15*(4), 661–671. https://doi.org/10.1007/s11292-019-09358-1

Tyler, T. R., & Blader, S. L. (2003). The group engagement model: Procedural justice, social identity, and cooperative behavior. *Personality and Social Psychology Review*, *7*(4), 349–361. https://doi.org/10.1207/S15327957PSPR0704_07

Uhde, A., Schlicker, N., Wallach, D. P., & Hassenzahl, M. (2020). Fairness and decision-making in collaborative shift scheduling systems. In *CHI '20: Proceedings of the 2020 Conference on Human Factors in Computing Systems,* Honolulu, HI, USA. http://arxiv.org/pdf/2001.09755v2 van der Woerdt, S., & Haselager, P. (2017). Lack of effort or lack of ability? Robot failures and human perception of agency and responsibility. In T. Bosse & B. Bredeweg (Eds.), *Communications in Computer and Information Science: Vol. 765, BNAIC '16: Proceedings of the 28th Benelux Conference on Artificial Intelligence. Revised selected papers* (pp. 155–168). Springer. https://doi.org/10.1007/978-3-319-67468-1_11

van der Woerdt, S., & Haselager, P. (2019). When robots appear to have a mind: The human perception of machine agency and responsibility. *New Ideas in Psychology*, *54*, 93–100. https://doi.org/10.1016/j.newideapsych.2017.11.001

VanYperen, N. W., Hagedoorn, M., Zweers, M., & Postma, S. (2000). Injustice and employees' destructive responses: The mediating role of state negative affect. *Social Justice Research*, *13*, 291–312. https://doi.org/10.1023/A:1026411523466

Visser, E. J. de, Monfort, S. S., McKendrick, R., Smith, M. A. B., McKnight, P. E., Krueger, F., & Parasuraman, R. (2016). Almost human: Anthropomorphism increases trust resilience in cognitive agents. *Journal of Experimental Psychology: Applied*, *22*(3), 331–349. https://doi.org/10.1037/xap0000092

Voiklis, J., Kim, B., Cusimano, C., & Malle, B. F. (2016). Moral judgments of human vs. robot agents. In *RO-MAN '16: Proceedings of the 25th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)* (pp. 775–780). IEEE Press. https://doi.org/10.1109/ROMAN.2016.7745207

Walter, S., Wendt, C., Böhnke, J., Crawcour, S., Tan, J.-W., Chan, A., Limbrecht, K., Gruss, S., & Traue, H. C. (2014). Similarities and differences of emotions in human-machine and human-human interactions: What kind of emotions are relevant for future companion systems? *Ergonomics*, *57*(3), 374–386. https://doi.org/10.1080/00140139.2013.822566

Watson, D. (2000). *Mood and temperament*. *Emotions and social behavior*. Guilford Press.

Weiner, B. (1995). *Judgments of responsibility: A foundation for a theory of social conduct*. The Guilford Press.

Weiner, B. (2006). *Social motivation, justice, and the moral emotions: An attributional approach*. Psychology Press.

Weiss, H. M., & Cropanzano, R. S. (1996). Affective Events Theory: A theoretical discussion of the structure, causes and consequences of affective experiences at work. In B. M. Staw & L. L. Cummings (Eds.), *Research in organizational behavior: An annual series of analytical essays and critical reviews* (pp. 1–74). JAI Press.

Weiss, H. M., Suckow, K. J., & Cropanzano, R. S. (1999). Effects of justice conditions on discrete emotions. *Journal of Applied Psychology*, *84*(5), 786–794. https://doi.org/10.1037/0021-9010.84.5.786

Wesche, J. S., & Sonderegger, A. (2019). When computers take the lead: The automation of leadership. *Computers in Human Behavior*, *101*, 197–209. https://doi.org/10.1016/j.chb.2019.07.027

Whitman, D. S., Caleo, S., Carpenter, N. C., Horner, M. T., & Bernerth, J. B. (2012). Fairness at the collective level: A meta-analytic examination of the consequences and boundary conditions of organizational justice climate. *Journal of Applied Psychology*, *97*(4), 776–791. https://doi.org/10.1037/a0028021

Wickens, C. M., Wiesenthal, D. L., Flora, D. B., & Flett, G. L. (2011). Understanding driver anger and aggression: Attributional theory in the driving environment. *Journal of Experimental Psychology: Applied*, *17*(4), 354–370. https://doi.org/10.1037/a0025815

Xie, Y., Bodala, I. P., Ong, D. C., Hsu, D., & Soh, H. (2019). Robot capability and intention in trust-based decisions across tasks. In *HRI '19: Proceedings of the 14th ACM/IEEE International Conference on Human-Robot Interaction* (pp. 39–47). IEEE. https://doi.org/10.1109/HRI.2019.8673084

Yi, Y., & Gong, T. (2008). The effects of customer justice perception and affect on customer citizenship behavior and customer dysfunctional behavior. *Industrial Marketing Management*, *37*(7), 767–783. https://doi.org/10.1016/j.indmarman.2008.01.005

Yuan, G., Jia, L., & Zhao, J. (2016). Organizational identification moderates the impact of organizational justice on job satisfaction. *Work*, *54*(1), 189–195. https://doi.org/10.3233/WOR-162271

Table 1

Means and Standard Deviations by Experimental Groups

		N	Controll-ability	Intention-ality	Positive Affect	Negative Affect	Trust	Identification	Job Satis-faction[a]	Commit-ment	OCB	CWB
Study 1												
	Human	56	-	-	3.28 (0.47)	2.05 (0.69)	3.70 (0.54)	3.34 (0.73)	5.03 (1.04)	3.31 (0.69)	3.84 (0.47)	1.69 (0.56)
Fair	humanoid	31	-	-	3.07 (0.62)	2.29 (0.61)	3.20 (0.71)	3.21 (0.64)	4.31 (1.20)	3.14 (0.75)	3.72 (0.39)	1.81 (0.59)
	machine-like	30	-	-	3.05 (0.54)	2.68 (0.70)	3.21 (0.67)	3.09 (0.84)	4.01 (1.04)	2.82 (0.72)	3.49 (0.53)	2.24 (0.85)
	human	56	-	-	2.74 (0.65)	2.95 (0.66)	2.29 (0.71)	2.73 (1.01)	3.10 (1.12)	2.37 (0.72)	3.57 (0.67)	1.85 (0.63)
Unfair	humanoid	28	-	-	2.75 (0.64)	3.10 (0.71)	2.60 (0.82)	2.74 (0.87)	2.99 (1.23)	2.42 (0.71)	3.67 (0.47)	1.87 (0.57)
	machine-like	28	-	-	2.40 (0.61)	3.03 (0.64)	2.59 (0.74)	2.95 (0.89)	3.07 (1.06)	2.46 (0.90)	3.54 (0.72)	2.08 (0.65)
Study 2												
	human	20	3.59 (0.88)	3.68 (0.73)	3.84 (0.49)	1.92 (0.59)	3.97 (0.58)	3.94 (0.67)	5.68 (0.84)	3.81 (0.55)	4.27 (0.44)	1.41 (0.50)
Fair	humanoid	23	3.20 (0.89)	3.20 (1.09)	3.20 (0.57)	2.22 (0.70)	3.57 (0.71)	3.47 (0.70)	4.47 (1.28)	3.16 (0.83)	3.92 (0.45)	1.49 (0.60)
	machine-like	26	3.31 (0.92)	3.08 (1.18)	3.38 (0.61)	2.19 (0.68)	3.36 (0.49)	3.48 (0.73)	4.87 (1.13)	3.21 (0.64)	4.01 (0.41)	1.38 (0.37)
	human	18	4.28 (0.58)	4.03 (0.58)	2.36 (0.72)	3.29 (0.55)	2.27 (0.66)	2.58 (0.68)	3.21 (1.04)	2.33 (0.78)	3.56 (0.74)	1.52 (0.46)
Unfair	humanoid	24	3.67 (0.90)	3.58 (0.75)	2.49 (0.65)	3.29 (0.63)	2.39 (0.67)	2.69 (0.95)	3.04 (1.08)	2.21 (0.64)	3.82 (0.55)	1.81 (0.62)
	machine-like	21	3.39 (0.99)	3.36 (0.79)	2.91 (0.45)	3.38 (0.53)	2.31 (0.62)	2.94 (0.85)	2.99 (0.76)	2.70 (0.88)	3.70 (0.62)	1.92 (0.69)

Note: M (*SD*); $N_{Study1} = 229$, $N_{Study2} = 132$; [a] 7-point.

Table 2

Intercorrelations and Reliability of Study Variables

		1	2	3	4	5	6	7	8	9	10	11	12
1	Proc. justice[a]	*.92/.91*	.30**	.58***	-.68***	.78***	.55***	.78***	.60***	.30***	-.46***	-.04	-.13
2	Anthropom.[a]	.58***	*.82/.75*	.28**	-.31***	.36***	.35***	.42***	.34***	.24**	-.25**	.21*	.22**
3	Positive affect	.46***	.43***	*.83/.90*	-.63***	.71***	.67***	.71***	.75***	.45***	-.40***	-.06	-.22**
4	Negative affect	-.52***	-.40***	-.47***	*.87/.90*	-.66***	-.55***	-.70***	-.54***	-.34***	.45***	.12	.16
5	Trust	.80***	.61***	.46***	-.52***	*.90/.90*	.59***	.77***	.72***	.35***	-.39***	-.09	-.13
6	Identification	.44***	.37***	.48***	-.26***	.50***	*.90/.91*	.60***	.73***	.37***	-.37***	-.07	-.10
7	Job satisfaction	.77***	.68***	.54***	-.50***	.78***	.60***	*.91/.92*	.72***	.37***	-.46***	-.04	-.11
8	Commitment	.59***	.55***	.55***	-.39***	.66***	.69***	.77***	*.93/.94*	.46***	-.41***	-.15	-.16
9	OCB	.23**	.22**	.37***	-.33***	.29***	.42***	.29***	.36***	*.90/.92*	-.70***	-.01	.02
10	CWB	-.07	-.08	-.05	.25***	-.15*	-.17*	-.18**	-.17**	-.50***	*.88/.89*	-.02	.03
11	Intentionality	-	-	-	-	-	-	-	-	-	-	*-/.73*	.50***
12	Controllability	-	-	-	-	-	-	-	-	-	-	-	*-/.79*

Note: $N_{\text{Study1}} = 229$, $N_{\text{Study2}} = 132$. Diagonal represents Cronbach's alpha (Study 1/Study 2). Study 1 under diagonal, Study 2 above diagonal.

[a] Perceived measure.

* $p < .05$, ** $p < .01$, *** $p < .001$

Table 3

Indirect Effects for the Parallel Multiple Mediation Model

	Consequent											
	Job satisfaction			Commitment			OCB			CWB		
Mediator	*ab*	*SE*	CI	*ab*	*SE*	CI	*ab*	*SE*	CI	*ab*	*SE*	CI
Study 1												
Total	0.60	.07	[0.47, 0.75]	0.31	.04	[0.22, 0.39]	0.17	.04	[0.10, 0.24]	-0.07	.03	[-0.15, -0.01]
Pos. affect	0.05	.03	[0.01, 0.11]	0.06	.02	[0.02, 0.11]	0.03	.02	[-0.00, 0.06]	0.04	.02	[-0.00, 0.08]
Neg. affect	0.08	.03	[0.02, 0.14]	0.01	.02	[-0.03, 0.05]	0.07	.02	[0.03, 0.12]	-0.10	.03	[-0.16, -0.05]
Trust	0.38	.06	[0.27, 0.49]	0.15	.03	[0.09, 0.21]	0.02	.03	[-0.03, 0.08]	0.01	.04	[-0.07, 0.08]
Identification	0.09	.03	[0.04, 0.15]	0.09	.03	[0.04, 0.14]	0.05	.02	[0.02, 0.08]	-0.03	.01	[-0.06, 0.00]
Study 2												
Total	0.82	.10	[0.62, 1.02]	0.49	.07	[0.35, 0.64]	0.21	.06	[0.10, 0.34]	-0.21	.05	[-0.32, -0.11]
Pos. affect	0.13	.06	[0.01, 0.26]	0.14	.04	[0.07, 0.22]	0.08	.05	[-0.03, 0.19]	0.00	.04	[-0.09, 0.08]
Neg. affect	0.23	.08	[0.08, 0.40]	-0.04	.06	[-0.15, 0.08]	0.05	.04	[-0.02, 0.14]	-0.15	.05	[-0.26, -0.06]
Trust	0.38	.09	[0.21, 0.54]	0.24	.07	[0.10, 0.38]	0.05	.06	[-0.05, 0.17]	-0.02	.05	[-0.12, 0.09]
Identification	0.08	.05	[-0.02, 0.19]	0.15	.04	[0.08, 0.24]	0.02	.03	[-0.04, 0.09]	-0.04	.03	[-0.11, 0.02]

Note: $N_{\text{Study1}} = 229$, $N_{\text{Study2}} = 132$. *ab* = unstandardized indirect effect/indirect effect; CI = bias-corrected 95% bootstrap confidence interval, 10,000 bootstrap samples.

Table 4

Indirect Effects for the Moderated Mediation Models, Split Between Levels of the Moderator: Study 1

		Consequent											
	Level of	Job satisfaction			Commitment			OCB			CWB		
Mediator	Moderator	*ab*	*SE*	CI	*ab*	*SE*	CI	*ab*	*SE*	CI	*ab*	*SE*	CI
Pos. affect		**0.05**	**.03**	**[0.01, 0.11]**	**0.06**	**.02**	**[0.02, 0.11]**	0.03	.02	[-0.00, 0.06]	0.04	.02	[-0.00, 0.08]
	Human	**0.10**	**.04**	**[0.03, 0.18]**	0.01	.03	[-0.04, 0.06]	**0.09**	**.03**	**[0.04, 0.15]**	**-0.12**	**.04**	**[-0.20, -0.05]**
Neg. affect	Humanoid	**0.09**	**.04**	**[0.02, 0.17]**	0.01	.02	[-0.04, 0.06]	**0.08**	**.03**	**[0.03, 0.13]**	**-0.11**	**.04**	**[-0.19, -0.05]**
	Android	0.04	.02	[-0.00, 0.09]	0.00	.01	[-0.02, 0.03]	0.03	.02	[-0.00, 0.08]	-0.05	.03	[-0.11, 0.00]
	Human	**0.12**	**.04**	**[0.05, 0.21]**	**0.21**	**.05**	**[0.13, 0.30]**	0.03	.04	[-0.04, 0.11]	0.02	.05	[-0.09, 0.11]
Trust	Humanoid	**0.09**	**.04**	**[0.01, 0.19]**	**0.09**	**.03**	**[0.03, 0.16]**	0.01	.02	[-0.02, 0.05]	0.01	.02	[-0.04, 0.05]
	Android	0.03	.05	[-0.06, 0.12]	**0.09**	**.03**	**[0.04, 0.16]**	0.01	.02	[-0.02, 0.05]	0.01	.02	[-0.04, 0.06]
	Human	**0.53**	**.08**	**[0.39, 0.69]**	**0.12**	**.04**	**[0.05, 0.20]**	**0.06**	**.02**	**[0.02, 0.12]**	-0.04	.02	[-0.08, 0.00]
Identification	Humanoid	**0.23**	**.08**	**[0.08, 0.39]**	**0.09**	**.04**	**[0.01, 0.18]**	**0.05**	**.02**	**[0.01, 0.10]**	-0.03	.02	[-0.07, 0.00]
	Android	**0.23**	**.07**	**[0.09, 0.38]**	0.03	.04	[-0.06, 0.11]	0.01	.02	[-0.03, 0.07]	-0.01	.01	[-0.04, 0.02]

Note: $N = 229$. *ab* = unstandardized regression coefficient/indirect effect; CI = bias-corrected 95% bootstrap confidence interval, 10,000 bootstrap samples. CIs not including zero are printed in bold.

Table 5

Pairwise comparisons of the indirect effects of the moderator groups: Study 1

		Consequent											
		Job satisfaction			Commitment			OCB			CWB		
Mediator	Moderator	Diff.	*SE*	CI	Diff.	*SE*	CI	Diff.	*SE*	CI	Diff.	*SE*	CI
Negative affect	Contrast_{DA1}	0.04	.03	[-0.00, 0.09]	0.00	.01	[-0.02, 0.03]	0.03	.02	[-0.00, 0.08]	-0.04	.03	[-0.10, 0.00]
	Contrast_{DA2}	0.05	.04	[-0.00, 0.13]	0.01	.02	[-0.02, 0.04]	0.05	.03	[-0.00, 0.10]	-0.06	.04	[-0.14, 0.00]
Trust	Contrast_{DA1}	**0.30**	**.08**	**[0.16, 0.46]**	**0.12**	**.04**	**[0.05, 0.20]**	0.02	.02	[-0.02, 0.06]	0.01	.03	[-0.05, 0.07]
	Contrast_{DA2}	-0.01	.10	[-0.21, 0.20]	-0.00	.04	[-0.08, 0.08]	-0.00	.01	[-0.02, 0.02]	-0.00	.01	[-0.02, 0.02]
Identification	Contrast_{DA1}	0.06	.05	[-0.03, 0.17]	0.06	.05	[-0.03, 0.15]	0.03	.03	[-0.01, 0.09]	-0.02	.02	[-0.06, 0.01]
	Contrast_{DA2}	0.07	.06	[-0.05, 0.20]	0.07	.06	[-0.05, 0.19]	0.03	.03	[-0.03, 0.10]	-0.02	.02	[-0.07, 0.02]

Note: $N = 229$. DA = decision agent; Contrast_{DA1}: 2/3 = human, -1/3 = humanoid, -1/3 = machinelike; Contrast_{DA2}: 0 = human, 1/2 = humanoid, -1/2 = machinelike; Diff. = difference between indirect effects; CI = bias-corrected 95% bootstrap confidence interval, 10,000 bootstrap samples, CIs not including zero are printed in bold.

Table 6

Indirect Effects for the Moderated Mediation Models, Split Between Levels of the Moderator: Study 2

		Consequent											
		Job satisfaction			Commitment			OCB			CWB		
Mediator	Level of Moderator	*ab*	*SE*	CI	*ab*	*SE*	CI	*ab*	*SE*	CI	*ab*	*SE*	CI
Pos. affect		**0.13**	**.06**	**[0.00, 0.26]**	**0.14**	**.04**	**[0.07, 0.22]**	0.08	.05	[-0.03, 0.20]	-0.00	.04	[-0.09, 0.08]
	Human	**0.26**	**.10**	**[0.09, 0.47]**	-0.04	.07	[-0.17, 0.09]	0.06	.05	[-0.03, 0.17]	**-0.17**	**.06**	**[-0.29, -0.07]**
Neg. affect	Humanoid	**0.20**	**.08**	**[0.07, 0.38]**	-0.03	.05	[-0.13, 0.07]	0.05	.04	[-0.02, 0.13]	**-0.13**	**.05**	**[-0.25, -0.05]**
	Android	**0.22**	**.09**	**[0.07, 0.42]**	-0.04	.06	[-0.15, 0.08]	0.05	.04	[-0.02, 0.14]	**-0.15**	**.05**	**[-0.27, -0.05]**
	Human	**0.50**	**.12**	**[0.27, 0.75]**	**0.31**	**.10**	**[0.13, 0.51]**	0.07	.08	[-0.07, 0.23]	-0.02	.07	[-0.16, 0.11]
Trust	Humanoid	**0.35**	**.09**	**[0.18, 0.53]**	**0.22**	**.07**	**[0.09, 0.37]**	0.05	.05	[-0.05, 0.16]	-0.02	.05	[-0.11, 0.08]
	Android	**0.31**	**.09**	**[0.15, 0.49]**	**0.19**	**.06**	**[0.08, 0.33]**	0.04	-05	[-0.04, 0.15]	-0.01	.04	[-0.10, 0.07]
	Human	0.13	.08	[-0.03, 0.28]	**0.24**	**.07**	**[0.11, 0.39]**	0.04	.05	[-0.06, 0.14]	-0.06	.05	[-0.17, 0.04]
Identification	Humanoid	0.08	.05	[-0.01, 0.18]	**0.14**	**.05**	**[0.04, 0.25]**	0.02	.03	[-0.04, 0.08]	-0.03	.03	[-0.10, 0.02]
	Android	0.05	.04	[-0.01, 0.15]	**0.09**	**.04**	**[0.02, 0.19]**	0.01	.02	[-0.02, 0.07]	-0.02	.02	[-0.08, 0.01]

Note: $N = 132$. *ab* = unstandardized regression coefficient/indirect effect; CI = bias-corrected 95% bootstrap confidence interval, 10,000 bootstrap samples. CIs not including zero are printed in bold.

Table 7

Pairwise comparisons of the indirect effects of the moderator groups: Study 2

		Consequent											
		Job satisfaction			Commitment			OCB			CWB		
Mediator	Moderator	Diff.	*SE*	CI	Diff.	*SE*	CI	Diff.	*SE*	CI	Diff.	*SE*	CI
Negative affect	$Contrast_{DA1}$	0.05	.05	[-0.04, 0.15]	-0.01	.02	[-0.05, 0.02]	0.01	.02	[-0.01, 0.05]	-0.03	.03	[-0.09, 0.03]
	$Contrast_{DA2}$	-0.02	.05	[-0.14, 0.07]	0.00	.02	[-0.02, 0.05]	-0.01	.02	[-0.04, 0.02]	0.01	.04	[-0.05, 0.09]
Trust	$Contrast_{DA1}$	**0.17**	**.08**	**[0.03, 0.35]**	**0.11**	**.05**	**[0.02, 0.23]**	0.02	.03	[-0.02, 0.10]	-0.01	.03	[-0.07, 0.04]
	$Contrast_{DA2}$	0.04	.08	[-0.12, 0.19]	0.02	.05	[-0.07, 0.13]	0.01	.02	[-0.03, 0.04]	-0.00	.01	[-0.03, 0.02]
Identification	$Contrast_{DA1}$	0.07	.05	[-0.01, 0.17]	**0.12**	**.06**	**[0.02, 0.27]**	0.02	.03	[-0.03, 0.09]	-0.03	.03	[-0.10, 0.02]
	$Contrast_{DA2}$	0.02	.04	[-0.05, 0.11]	0.04	.06	[-0.07, 0.18]	0.01	.02	[-0.03, 0.04]	-0.01	.02	[-0.05, 0.03]

Note: N = 132. DA = decision agent; $Contrast_{DA1}$: 2/3 = human, -1/3 = humanoid, -1/3 = machinelike; $Contrast_{DA2}$: 0 = human, 1/2 = humanoid, -1/2 = machinelike; Diff. = difference between indirect effects; CI = bias-corrected 95% bootstrap confidence interval, 10,000 bootstrap samples, CIs not including zero are printed in bold.

Table 8

Indirect Effects for the Alternative Moderated Mediation Models, Split Between Levels of the Moderator: Intentionality

		Consequent											
	Level of	Job satisfaction			Commitment			OCB			CWB		
Mediator	Moderator	*ab*	*SE*	CI	*ab*	*SE*	CI	*ab*	*SE*	CI	*ab*	*SE*	CI
Pos. affect		**0.13**	**.06**	**[0.00, 0.26]**	**0.14**	**.04**	**[0.07, 0.22]**	0.08	.05	[-0.03, 0.19]	**-0.00**	**.04**	[-0.09, 0.08]
	-1*SD*	**0.15**	**.06**	**[0.05, 0.30]**	-0.02	.04	[-0.10, 0.05]	0.04	.03	[-0.02, 0.10]	**-0.10**	**.04**	**[-0.19, -0.04]**
Neg. affect	*M*	**0.23**	**.08**	**[0.08, 0.41]**	-0.04	.06	[-0.15, 0.08]	0.06	.04	[-0.03, 0.15]	**-0.15**	**.05**	**[-0.26, -0.06]**
	+1*SD*	**0.29**	**.10**	**[0.10, 0.50]**	-0.05	.07	[-0.19, 0.10]	0.07	.05	[-0.03, 0.18]	**-0.19**	**.06**	**[-0.32, -0.08]**
	-1*SD*	**0.26**	**.07**	**[0.12, 0.41]**	**0.16**	**.06**	**[0.06, 0.29]**	0.04	.04	[-0.04, 0.12]	-0.01	.04	[-0.08, 0.06]
Trust	*M*	**0.39**	**.09**	**[0.21, 0.56]**	**0.24**	**.07**	**[0.10, 0.39]**	0.06	.06	[-0.05, 0.17]	-0.02	.05	[-0.12, 0.09]
	+1*SD*	**0.47**	**.11**	**[0.26, 0.68]**	**0.30**	**.09**	**[0.13, 0.47]**	0.07	.07	[-0.07, 0.21]	-0.02	.07	[-0.15, 0.11]
	-1*SD*	0.04	.03	[-0.01, 0.12]	**0.08**	**.04**	**[0.01, 0.16]**	0.01	.02	[-0.02, 0.06]	-0.02	.02	[-0.07, 0.01]
Identification	*M*	0.09	.05	[-0.02, 0.19]	**0.16**	**.04**	**[0.08, 0.24]**	0.02	.03	[-0.04, 0.09]	-0.04	.03	[-0.11, 0.02]
	+1*SD*	0.12	.07	[-0.02, 0.26]	**0.21**	**.06**	**[0.10, 0.33]**	0.03	.04	[-0.05, 0.12]	-0.05	.04	[-0.14, 0.03]

Note: $N = 132$. *ab* = unstandardized regression coefficient/indirect effect; CI = bias-corrected 95% bootstrap confidence interval, 10,000 bootstrap samples. CIs not including zero are printed in bold.

Table 9

Indirect Effects for the Alternative Moderated Mediation Models, Split Between Levels of the Moderator: Controllability

		Consequent											
	Level of	Job satisfaction			Commitment			OCB			CWB		
Mediator	Moderator	*ab*	*SE*	CI	*ab*	*SE*	CI	*ab*	*SE*	CI	*ab*	*SE*	CI
Pos. affect		**0.13**	**.06**	**[0.00, 0.26]**	**0.14**	**.04**	**[0.07, 0.22]**	0.08	.05	[-0.03, 0.19]	-0.00	.04	[-0.09, 0.08]
	-1*SD*	**0.13**	**.06**	**[0.04, 0.28]**	-0.02	.04	[-0.09, 0.05]	0.03	.03	[-0.01, 0.10]	**-0.09**	**.04**	**[-0.18, -0.03]**
Neg. affect	*M*	**0.24**	**.09**	**[0.08, 0.42]**	-0.04	.06	[-0.15, 0.08]	0.06	.04	[-0.03, 0.14]	**-0.15**	**.05**	**[-0.26, -0.06]**
	+1*SD*	**0.31**	**.11**	**[0.11, 0.55]**	-0.05	.08	[-0.21, 0.10]	0.08	.06	[-0.04, 0.19]	**-0.21**	**.07**	**[-0.34, -0.08]**
	-1*SD*	**0.29**	**.09**	**[0.12, 0.49]**	**0.18**	**.07**	**[0.06, 0.34]**	0.04	.05	[-0.04, 0.15]	-0.01	.04	[-0.10, 0.07]
Trust	*M*	**0.38**	**.09**	**[0.21, 0.56]**	**0.24**	**.07**	**[0.10, 0.39]**	0.05	.06	[-0.05, 0.17]	-0.02	.05	[-0.12, 0.09]
	+1*SD*	**0.45**	**.11**	**[0.25, 0.67]**	**0.28**	**.08**	**[0.12, 0.45]**	0.06	.07	[-0.07, 0.20]	-0.02	.06	[-0.15, 0.10]
	-1*SD*	0.05	.04	[-0.01, 0.13]	**0.09**	**.04**	**[0.02, 0.19]**	0.01	.02	[-0.02, 0.07]	-0.02	.02	[-0.08, 0.01]
Identification	*M*	0.13	.05	[-0.02, 0.19]	**0.15**	**.04**	**[0.08, 0.24]**	0.02	.03	[-0.04, 0.09]	-0.04	.03	[-0.11, 0.02]
	+1*SD*	0.13	.07	[-0.02, 0.26]	**0.20**	**.06**	**[0.10, 0.32]**	0.03	.04	[-0.06, 0.11]	-0.05	.04	[-0.13, 0.03]

Note: $N = 132$. *ab* = unstandardized regression coefficient/indirect effect; CI = bias-corrected 95% bootstrap confidence interval, 10,000 bootstrap samples. CIs not including zero are printed in bold

Table 10

Difference Tests of the Indirect Effects of the Two Alternative Moderators

		Consequent											
		Job satisfaction			Commitment			OCB			CWB		
Mediator	Moderator	Index	*SE*	CI	Index	*SE*	CI	Index	*SE*	CI	Index	*SE*	CI
Negative affect	Intentionality	**0.07**	**.03**	**[0.02, 0.15]**	-0.01	.02	[-0.05, 0.02]	0.02	.01	[-0.01, 0.05]	**-0.05**	**.02**	**[-0.09, -0.01]**
	Control	**0.09**	**.04**	**[0.02, 0.18]**	-0.01	.02	[-0.07, 0.03]	0.02	.02	[-0.01, 0.06]	**-0.06**	**.02**	**[-0.11, -0.02]**
Trust	Intentionality	**0.12**	**.04**	**[0.04, 0.22]**	**0.08**	**.03**	**[0.03, 0.14]**	0.02	.02	[-0.02, 0.06]	-0.01	.02	[-0.04, 0.03]
	Control	0.08	.05	[-0.01, 0.19]	0.05	.03	[-0.00, 0.12]	0.01	.01	[-0.02, 0.04]	-0.00	.01	[-0.03, 0.02]
Identification	Intentionality	0.05	.04	[-0.01, 0.13]	**0.10**	**.04**	**[0.04, 0.18]**	0.02	.02	[-0.03, 0.05]	-0.03	.02	[-0.07, 0.02]
	Control	0.03	.03	[-0.01, 0.09]	0.06	.03	[-0.00, 0.12]	0.01	.01	[-0.02, 0.03]	-0.01	.01	[-0.04, 0.01]

Note: N = 132. Index = index of moderated mediation; CI = bias-corrected 95% bootstrap confidence interval, 10,000 bootstrap samples, CIs not including zero are printed in bold.

Appendix A

Figure A1

Original Vignettes Describing the Hypothetical Situation Used for the Manipulation

Im Nachfolgenden finden Sie die Beschreibung einer Situation, wie sie im Berufsalltag auftreten kann. Bitte lesen Sie diese aufmerksam durch und stellen Sie sich detailliert vor, diese Situation selbst zu erleben.

Als Mitarbeiter bzw. Mitarbeiterin in dem Unternehmen Car-Solutions, einem Produktionsunternehmen in der Automobilindustrie, sind Sie in einem Team im Fertigungsabschnitt „Leuchtanlage" unter anderem für die Montage der Scheinwerfer zuständig. Dieses Team ist ein gemischtes Team bestehend aus Ihnen, drei weiteren Mitarbeitenden, einem Teamleiter, dem Roboter Pepper und dem Roboter PX3000.

Für die Organisation Ihres Teams ist der **der Teamleiter (die Teamleiterin** oder **Roboter Pepper** oder **Roboter PX3000)** verantwortlich. Dies beinhaltet die Entscheidung über die Planung von Reihenfolge und Verteilung der Arbeitsaufträge. Bei dieser Entscheidung ist der **der Teamleiter (die Teamleiterin** oder **Roboter Pepper** oder **Roboter PX3000)** also der Entscheidungsträger.

-Foto-

Es ist Freitagmorgen und Sie sind gerade, gemeinsam mit den vier weiteren Mitarbeitenden Ihres Teams, bei der üblichen Teambesprechung.

Vor zwei Wochen teilte Ihnen der **der Teamleiter (die Teamleiterin** oder **Roboter Pepper** oder **Roboter PX3000)** mit, dass der Arbeitsprozess (aufgrund der Umstellung auf eine effizientere LED-Technik der Scheinwerfer) neu strukturiert wird. Diese Änderungen treten ab der nächsten Woche in Kraft und haben Auswirkungen auf Ihre Arbeit, d.h. Sie sind direkt betroffen.

Im Vorfeld der Entscheidung über diese neue Aufgabenverteilung bezog der **der Teamleiter (die Teamleiterin** oder **Roboter Pepper** oder **Roboter PX3000)** Ihre Präferenzen in der Aufgabenverteilung und Ihre Meinung *mit ein (nicht mit ein)* und *hat die Entscheidung somit unter Einbeziehung des Teams gefällt (hat die Entscheidung also allein gefällt).* Für das Vorgehen bei einer Neuverteilung von Aufgaben gibt es standardisierte und objektive Unternehmensrichtlinien, die besagen, dass die Argumente aller Teammitglieder bei der Entscheidung angehört werden und die Ergebnisse einiger (im Vorfeld durchgeführter) Tests besonnen und akkurat einbezogen werden müssen. Sie wissen, dass der **der Teamleiter (die Teamleiterin** oder **Roboter Pepper** oder **Roboter PX3000)** diese Richtlinien *vollständig befolgt hat (nicht vollständig befolgt hat).* Der **der Teamleiter (die Teamleiterin** oder **Roboter Pepper** oder **Roboter PX3000)** *nahm die Verteilung der Aufgaben so vor, dass kein Mitarbeiter für eine bestimmte Aufgabe bevorzugt ausgewählt wurde (bevorzugte bestimmte Personen bei der Aufgabenverteilung, sodass keine Gleichberechtigung bestand).* Insgesamt erfolgte die Entscheidung *nach demselben Vorgehen, wie bei vergangenen ähnlichen Situationen (im Vergleich zu vergangenen ähnlichen Situationen scheinbar willkürlich).* Nach einem Testdurchlauf gab es in der heutigen Teambesprechung *die Möglichkeit (keine Möglichkeit)* Änderungs- und Korrekturvorschläge zu machen bzw. sich zu melden, wenn Sie nicht mit der Entscheidung einverstanden sind.

Note: The manipulation of the type of decision agent is printed in bold, the manipulation of procedural justice is printed in italics. Human and fair condition are displayed in the main text, all other conditions are displayed in parentheses.

Figure A2

Translation of Vignettes Describing the Hypothetical Situation Used for the Manipulation

In the following, you will find a description of a situation that could occur in your everyday working life. Please be attentive while reading and imagine in detail experiencing this situation yourself.

Car-Solutions is a production company in the car industry. As an employee of the company, you are a member of the "lighting system" team, which is responsible for the assembly of headlights. This is a mixed team, consisting of yourself, three other employees, a team leader, the robot Pepper, and the robot PX3000.

The team leader (the robot Pepper or **the robot PX3000)** is responsible for the team's organization. This includes the decision on planning and scheduling of jobs. Hence, **the team leader (the robot Pepper** or **the robot PX3000)** makes this decision.

-picture-

It is Friday morning. You and the other four team members are at the usual team meeting.

Two weeks ago, **the team leader (the robot Pepper** or **the robot PX3000)** informed you that the work process will be restructured because of the conversion to a more efficient LED technology for the headlights. These changes become effective next week. They have an impact on your work and they concern you directly.

Before making the decision, **the team leader (the robot Pepper** or **the robot PX3000)** *did (did not) factor in* your preferences and opinions concerning the task allocation and therefore *made the decision along with the team (made the decision alone)*. Standardized and objective organizational guidelines are available for decisions about the reallocation of tasks. The guidelines state that the arguments of all of the team members have to be heard, and the results of some tests have to be included considerately and accurately. You know that **the team leader (the robot Pepper** or **the robot PX3000)** *entirely followed these guidelines (did not entirely follow these guidelines)*. **The team leader (the robot Pepper** or **the robot PX3000)** *made the allocation of tasks in such a way that no employee was favored with a certain task (favored certain persons, resulting in an unequal allocation of further tasks)*. The decision *was made in the same way as in similar situations (was made rather arbitrarily compared to similar situations)*. In a meeting after a trial phase, you *had the possibility (had no possibility)* to suggest changes and corrections, or to report that you disagree with the decision.

Note: The manipulation of the type of decision agent is printed in bold, the manipulation of procedural justice is printed in italics. Human and fair condition are displayed in the main text, all other conditions are displayed in parentheses.

Figure A3

Pictures of the Robotic Decision Agent With Captions

Two team members in interaction
with the robot Pepper

Two team members in interaction
with the robot PX3000

Two team members in interaction
with the team leader

Two team members in interaction
with the team leader

Printed on non-ageing paper °° ISO 9706