Technische Universität München
TUM School of Management

Leadership in the Digital Age: Utilizing Social Robots and Virtual Reality in Leadership

Jakub Edward Cichor

Vollständiger Abdruck der von der TUM School of Management der Technischen Universität München zur Erlangung eines

Doktors der Philosophie (Dr. phil.)

genehmigten Dissertation.

Vorsitz: Prof. Dr. Christoph Ungemach

Prüfer*innen der Dissertation:

1. Prof. Dr. Claudia Peus
2. Prof. Gudrun J. Klinker, Ph.D.

Die Dissertation wurde am 24.08.2023 bei der Technischen Universität München eingereicht und durch die TUM School of Management am 15.11.2023 angenommen.

Acknowledgments

I want to express my sincere gratitude to Prof. Dr. Claudia Peus, who supervised my dissertation and always provided support, guidance, and ample opportunities to grow both as a researcher and an individual.

I am deeply grateful to Dr. Franziska Emmerling, who mentored me throughout my Ph.D. She was not only the best mentor anyone could wish for; she became a cherished friend. Her unwavering support and encouragement were instrumental to my journey.

I am also thankful to Prof. Dr. Sylvia Hubner, who sparked my interest in robot leadership. Throughout our time together, she was both an important mentor and a great friend, whether we were celebrating successes or navigating (the luckily rare) challenges.

I also want to thank Dr. Maxim Egorov, who inspired me to switch my focus from computer science to leadership and to pursue a Ph.D. in this field. His encouragement and friendship served as a great motivator throughout my Ph.D.

I would also like to thank Prof. Dr. Christoph Ungemach for chairing my dissertation committee and Prof. Gudrun Klinker, Ph.D., for being my second examiner and offering invaluable insights into the technological facets of my work.

I am also very thankful for all the wonderful colleagues that I met throughout my Ph.D.: Anna Brzykcy, Leidy Cubillos-Pinilla, Regina Dutz, Enikö D’Errigo, Martin Fladerer, Alina Gales, Eleni Georganta, Elke Graf, Ilse Hagerer, Mira Kaut, Kristin Knipfer, Hannah Kunde, Kerstin Löwe, Fabrizio Palmas, Armin Pircher Verdorfer, David Plecher, Stephanie Rehbock, Ellen Schmid, Ulf Steinberg, Selina Stracke, Mojtaba Salem, Clarissa Zwarg, and the many others who I have met along the way.

I owe a debt of gratitude to Dr. Dominika Czernik, my aunt and role model. Taking me with her to her university classes when I was a child and sharing her own Ph.D. experiences with me profoundly influenced my interest in scientific pursuits.

Finally, I also want to thank all my family and friends for always supporting me and always believing in me.

"One thing I believe to the fullest is that if you think and achieve as a team, the individual accolades will take care of themselves. Talent wins games, but teamwork and intelligence win championships." – Michael Jordan

Table of Contents

Acknowledgments II

Table of Contents IV

List of Tables V

List of Figures VI

Abstract (English) VII

Abstract (German) IX

1. General Introduction 1

2. Robot Leadership – Investigating Human Perceptions and Reactions Toward Social Robots Showing Leadership Behaviors 9

3. I, Robot – or Leader? The Effects of a Robot's Transformational and Transactional Leadership Behaviors 32

4. Virtual Reality Environments for Assessing Leadership Behaviors 60

5. Overall Conclusion 90

6. References 97

List of Tables

Table 1 (Chapter 2): Questions Stated During Interviews and Group Discussions.

Table 2 (Chapter 4): Overview of Leadership Assessment Methods and Their Respective Validity

List of Figures

Figure 1 (Chapter 2):	Procedure Overview for Studies 1 and 2
Figure 2 (Chapter 2):	Gioia Methodology Based on Semi-Structured Interviews and Group Discussions
Figure 3 (Chapter 2):	Synthesized Overview of Participants' Preconceptions, Judgments, and Reflections
Figure 4 (Chapter 3):	Research Model with Dependent Variables and Hypotheses
Figure 5 (Chapter 3):	Overview of Experimental Conditions With Differences and Examples
Figure 6 (Chapter 3):	Mean Differences and Standard Deviations for Dependent Variables
Figure 7 (Chapter 4):	Overview of Considerations and Recommendations for VR Leadership Assessment

Abstract (English)

Digital transformation requires organizations to adapt rapidly. Leaders are crucial multiplicators during digital transformation, for their ability to maneuver an organization during periods of change substantially influences the organization's ability to come out ahead instead of being left behind. The approaches leaders can adopt to harness the benefits of novel technologies depend on the strengths, weaknesses, and implementation of those technologies. Two technologies that have shown promise in the context of leadership are social robots (i.e., by taking on leadership tasks) and Virtual Reality (VR) (i.e., by creating virtual environments for leadership assessment and training).

My dissertation focuses on leadership in the digital age and how technologies can be effectively leveraged. First, I investigated human perceptions and reactions to robots exhibiting leadership behaviors in a qualitative study (Chapter 2). Second, I examined the degree to which the effects of human leadership behaviors persist when robot leaders are employed, that is, the extent to which human followers can work on a task for a robot, in an experimental study (Chapter 3). Third, I explored the use of VR in leadership assessment and identified 11 concrete recommendations that scholars should consider when implementing leadership applications in VR (Chapter 4).

The results from this dissertation are essential to using technologies in leadership contexts. First, my work sheds light on perceptions of and reactions to robot leaders. Human followers tend to initially adopt strong opinions about and reactions toward robot leaders. However, reflecting on a technology's potential benefits and drawbacks allows followers to develop more balanced and nuanced views. Visions of the future also play an essential role; technological affinity nudges followers to think about utopian or dystopian futures that involve robot leaders and to articulate ethical concerns correspondingly. Second, my work examines the extent to which evidence from research on human leaders is relevant to robot

leaders. Robot leadership outcomes mirror human leadership in that transformational robot leaders influence affect, likability, perceived safety, and stress more positively than transactional robot leaders. The positive mediation effects on task engagement mirror prior findings on human leaders. Third, my dissertation explores the potential of VR technology in leadership assessment applications. I provide concrete recommendations for the implementation of avatars (leaders), characters (followers), and environments (contexts) in leadership assessments that are conducted in VR. I also propose a decision tree to guide researchers through critical points of the implementation process.

The practical and theoretical implications of the dissertation (Chapter 5) illustrate the importance of social robots and VR for leadership in the digital age. Both technologies show that robots can support leaders in their work and assessments.

Abstract (German)

Die digitale Transformation ist ein Prozess, der von Unternehmen schnelle Anpassungen erfordert, indem sie neue Technologien effektiv implementieren und agiler werden. Führungskräfte sind bei der digitalen Transformation von entscheidender Bedeutung, da ihre Fähigkeit, ein Unternehmen in Zeiten des Wandels zu manövrieren, einen wesentlichen Einfluss darauf hat, ob das Unternehmen kompetitiv bleibt und den Anschluss nicht verliert. Die spezifischen Ansätze, die Führungskräfte wählen können, um von neuen Technologien zu profitieren, hängen jedoch stark von den Stärken und Schwächen der Technologie und ihrer Implementierung ab. Zwei Technologien, die im Führungskontext vielversprechend sind, sind soziale Roboter (d. h. durch die Übernahme von Führungsaufgaben) und Virtual Reality (VR) (d. h. durch die Schaffung virtueller Trainingsumgebungen).

Meine Dissertation befasst sich daher mit dem Thema Führung im digitalen Zeitalter und der Frage, wie Technologien am effektivsten eingesetzt werden können. Zunächst habe ich die menschliche Wahrnehmung und Reaktion auf transformationale und transaktionale Führungsroboter untersucht (Kapitel 2). Zweitens untersuchte ich, inwieweit sich die Erkenntnisse über transformationale und transaktionale menschliche Führungskräfte auf Roboter-Führungskräfte übertragen lassen, indem ich Menschen in eine Position versetzte, in der sie eine Aufgabe "für den Roboter" bearbeiteten (Kapitel 3). Drittens untersuchte ich VR für die Beurteilung und Entwicklung von Führungskräften, indem ich 11 konkrete Empfehlungen für Wissenschaftler*innen vorgeschlagen habe, die bei der Implementierung von Führungsanwendungen in VR berücksichtigt werden sollten (Kapitel 4).

Die Ergebnisse dieser Dissertation sind für den Einsatz von Technologien im Führungskontext von Bedeutung. Erstens neigen Menschen zu Beginn zu starken Meinungen und Reaktionen gegenüber Roboter-Führungskräften, aber eine Reflexion über die

potenziellen Vor- und Nachteile der Technologie ermöglicht es den Menschen, ausgewogenere und nuanciertere Ansichten zu entwickeln. Zukunftsvisionen spielen ebenfalls eine wichtige Rolle, da Technologieaffinität dazu veranlasst, entweder über utopische oder dystopische Zukünfte mit Roboter-Führungskräften und ihren jeweiligen ethischen Bedenken nachzudenken. Zweitens habe ich zeigen können, dass die Ergebnisse von Roboter-Führungskräften die Literatur über menschliche Führungskräfte widerzuspiegeln scheinen, da transformationale Roboterführung im Vergleich zu transaktionaler Roboterführung Affekt, Sympathie, wahrgenommene Sicherheit und Stress positiv beeinflusst. Die positiven Mediationseffekte auf Engagement in Aufgaben spiegeln ebenfalls frühere Ergebnisse mit menschlichen Führungskräften wider. Drittens gebe ich konkrete Empfehlungen und Anleitungen für den Entscheidungsprozess bei der Implementierung von Führungsbewertungen in VR.

Die praktischen und theoretischen Implikationen meiner Dissertation (Kapitel 5) verdeutlichen die Bedeutung von sozialen Robotern und VR für die Führung im digitalen Zeitalter, da beide Technologien vielversprechend sind, indem sie Führungskräfte bei ihrer Arbeit und ihrer Beurteilung und Entwicklung unterstützen können.

1. General Introduction

Organizations face many complex challenges in their day-to-day and long-term operations. In particular, the digital transformation has pressured organizations to adopt technologies such as automated industrial systems (Leitao et al., 2016), big data (Sivarajah et al., 2017), and blockchain (Toufaily et al., 2021), which all entail specific opportunities and challenges. Researchers have pointed out that organizations must be adaptable to effectively navigate the digital transformation process (AlNuaimi et al., 2022; Menon & Suresh, 2020).

Digital transformation requires organizations to restructure themselves adequately to enable digital technology integration (Warner & Wager, 2019). An organization's preparedness for a digital transformation determines how much it can benefit from implementing technologies. In the case of Artificial Intelligence (AI), for example, a framework that evaluates an organization based on its technologies, activities, boundaries, and goals can be applied (Holmstrom, 2022). Implementing those technologies and the related organizational changes requires competent leadership (Saarikko et al., 2020).

According to Vial (2019), leadership is a vital structural factor affecting the value generated by digital technology. Leadership is at the center of the digital transformation process because leaders are tasked with navigating their organizations through demanding periods and thus helping organizations survive difficulties and transform them into opportunities (Hansen et al., 2011). The digital transformation process has resulted in the creation of the role of the Chief Digital Officer, which is dedicated to the coordination of organizational change in response to digital technologies (Kunisch et al., 2022). External crises such as the COVID-19 pandemic have also emphasized the importance of leaders in demanding and dynamic environments, and leaders emerged as critical drivers of the implementation of digital technologies when social contact had to be limited (Bartsch et al., 2021; Lin et al., 2021).

Yukl (2006) defined leadership as “the process of influencing others to understand and agree about what needs to be done and how it can be done effectively, and the process of facilitating individual and collective efforts to accomplish the shared objective.” How leaders enact their leadership style is crucial for maneuvering organizations in disruptive times. For instance, researchers have found that transformational leadership predicts leaders' intention to include technologies in their work (Gencer & Samur, 2015; Leng, 2008) and influences the handling of organizational change during digital transformations (AlNuaimi et al., 2022).

Given the central role of leadership during digital transformation processes, it is crucial to enhance the academic understanding of leadership processes and to determine how leadership can be improved effectively. One way in which leadership can be enhanced is through the utilization of novel technologies. Whereas technologies are frequently discussed by reference to their benefits and dangers in various societal and organizational contexts, their application to leadership still needs to be explored. This dissertation aims to contribute important insights into leadership in the digital age while investigating how technologies can fulfill concrete leadership tasks. The dissertation also inquires how digital tools can be used to assess leaders effectively. It is guided by three research questions.

Research Questions

A crucial component of the work of leaders during a digital transformation is the management of technology implementation (Pflaum & Gölzer, 2018). The use of robots almost doubled between 2013 and 2018 (IFR, 2019), a trend that mirrors those observed in the context of other technologies, and their use has recently been examined in organizational contexts (Wirtz et al., 2018; Zeng et al., 2020). Accordingly, robots are becoming an increasingly important element of the digitalization processes of organizations (Acemoglu et al., 2020; Cheng et al., 2019; Chiacchio et al., 2018; Graetz & Michaels, 2018). NetDragon Websoft recently even named a robot leader its new CEO (Bello, 2022). These developments exemplify the increasing relevance of robots in organizational contexts—they are assuming responsibility for complex tasks and social corporate roles, including leadership.

Scholars have pointed to the theoretical advantages of robot leaders, e.g., concerning repeatable implementation, multitasking, and advanced problem-solving skills (Samani et al., 2011). Robots are also known to have limitations, which have to do chiefly with the complexity of implementing leadership behaviors into the technology and lack of creativity (Samani et al., 2012). It has been posited that as robots become more capable and less distinct from humans, it will become more acceptable to use them for tasks executed initially by humans (Lewis et al., 2012; Lewis et al., 2013). Literature on collaboration between humans and robots also reports promising results, indicating that humans and robots can work together with varying forms of control (Lewis et al., 2010) and based on various internal, external, and technological factors (Simoes et al., 2020). However, whether leadership can be implemented in a robot must be clarified.

Since research on robot leadership is limited, understanding human perceptions of and reactions toward robot leaders is an essential first step to enhancing the current knowledge of the interaction between human followers and robot leaders. Human-robot interaction (HRI) researchers have already shown that humans are capable of accepting robots as teammates (Gombolay et al., 2015), that the formation of trust between humans and robots follows similar mechanisms to the building of trust between humans (Mota et al., 2016), and that the acceptance of robots in social interactions depends on context (Westlund et al., 2016). At the same time, whether humans can accept robots in positions of authority is a question that has yet to be explored in detail.

The desirable behavior of a robot leader merits investigation. Since a substantial amount of research has been conducted on different leadership styles and the behaviors and outcomes that are associated with them, leadership style research is a fitting theoretical framework for studies on robot leadership. Transformational and transactional leadership are, in particular, the two most extensively researched leadership styles, and they are associated with various effects on followers in different contexts (Bass, 1999; Eagly et al., 2003; Judge & Piccolo, 2004). My first research project aimed to explore human perceptions of and reactions toward robot leaders that exhibit transformational or transactional leadership behaviors. My first research question, therefore, is as follows:

RQ1: *How do humans perceive and react to transformational versus transactional robot leaders?*

Beyond perceptions and reactions, it is critical to consider the outcomes of using transformational and transactional robot leaders from the perspective of human followers. Transformational leadership has been linked to positive outcomes such as enhanced performance, creativity, and trust (Gumusluoglu & Ilsev, 2009; Judge & Piccolo, 2004; Podsakoff et al., 1990), whereas transactional leadership is associated with superior performance under time pressure (Bass, 1999). However, it is still being determined to what degree these findings apply to robot leaders.

In HRI research, the Uncanny Valley phenomenon has been found to affect how humans prefer to engage with and accept robots. Contrary to expectations, in the Uncanny Valley, increases in human likeness do not increase humans' familiarity with the robot linearly. Instead, a drop in familiarity occurs when a robot is distinctly humanlike but exhibits flaws that starkly contrast its human features. This contrast evokes feelings of eeriness in humans (Mori et al., 2012). Researchers have found that these perceptions can have pronounced adverse effects on humans interacting with robots (Walters et al., 2008). At the same time, leadership scholars have investigated various adverse outcomes that followers experience due to negative leader behaviors. Combining evidence from HRI and leadership research highlights that some outcomes of leadership styles and technology implementations can be found in interactions with humans and robots. For example, stress is generally considered to negatively affect humans (Ganster & Rosen, 2013; Garfin et al., 2018; Turner et al., 2020). Researchers have found that robots (Lu et al., 2022) and human leaders (Winston & Patterson, 2006) elicit stress reactions in various circumstances. Given the comparable foci of these disparate domains of research, my second research question is:

RQ2: *To what extent do findings related to outcomes of human leaders' behaviors on followers translate to robot leaders?*

Beyond robots taking on leadership roles and their supportive capabilities, leadership can be supported in the digital age through technology in leadership assessment. Current leadership assessment approaches frequently rely on self-report and introspective measurements, which lack internal and external validity as a consequence of the inaccuracy of follower ratings (Wang et al., 2019), the lack of context-specificity (Thoroughgood et al., 2016), and the practice of measuring perceptions of leadership instead of actual leader behaviors (Hansbrough et al., 2015). Leadership scholars have relied on leadership assessment methods, such as Situational Judgment Tests, which depend on a choice between different behavioral options (McDaniel & Nguyen, 2001; Whetzel & McDaniel, 2009). Although the Situational Judgment Test is a crucial tool for assessing leadership (Peus et al., 2013), its focus on measuring behavioral intentions rather than actual behaviors limits the amount of information that can be obtained from them.

The gap between intentions and actual behaviors has been bridged in the context of leadership development, which occurs primarily through the enactment of leadership, that is, through concrete behaviors and experiences (DeRue & Myers, 2014) that can be practiced during role-play exercises. Role-playing exercises are efficient for the development of leaders because they allow them to practice concrete behaviors in specifically designed leadership situations (Kark, 2011). This enactment of leadership behaviors can be utilized in leadership assessments by implementing prototypical leadership situations into immersive Virtual Reality (VR) and using them as frameworks in which leadership can be directly observed and assessed.

VR is an immersive technology that allows users to enter virtual environments that seem more realistic and believable than textual descriptions or traditional two-dimensional screen-based virtual worlds (Martirosov et al., 2021; Parong & Mayer, 2018). Since users are placed in immersive virtual environments and allowed to interact, they feel like genuine actors in a three-dimensional world (Schmid Mast et al., 2018). Presence (i.e., the feeling of being inside a virtual body), social presence (i.e., perceiving virtual characters as believable interaction partners), and telepresence (i.e., experiencing the virtual environment as a realistic representation of the real world; Bulu, 2012) enable situations to be experienced in forms that mimic the real world so closely that observed behavior can be expected to mirror real-world behavior.

Leadership must be correctly assessed to formulate the specific skills and behaviors trainees would benefit from most. Initial investigations in non-immersive VR already indicate that by combining non-immersive VR with measures like eye-tracking and machine learning, one can differentiate between the approaches of prospective leaders (Parra et al., 2021). However, the effectiveness of leadership assessments through immersive VR and the most effective means of implementing the technology remains to be determined. My third research question is thus:

RQ 3: *What specific recommendations can be made to support researchers in implementing leadership assessment applications in VR?*

Chapter Overview

I investigated my research questions in a qualitative study (Chapter 2, focused on robot leadership), a quantitative experiment (Chapter 3, focused on robot leadership), and a conceptual paper (Chapter 4, focused on VR).

To answer RQ1, I present two qualitative studies with 29 participants total (Chapter 2). The participants experienced a transformational or transactional robot leader and responded to questions during semi-structured interviews or group discussions. We manipulated the speech and the movements of the robot leader. The study yielded important insights into human perceptions of and reactions toward robot leaders and illuminated the factors that influence said perceptions and reactions. Those factors include humans' technological knowledge and their visions for the application of technology in the future.

To answer RQ2 and to shed light on the consequences of robots assuming leadership roles, I investigated the effects of robot leadership behaviors on human followers in Chapter 3. I conducted a quantitative between-participants experimental study (N = 218) with the robot leader described in Chapter 2. This way, I tested the differences between the participants who interacted with a transformational, transactional, or minimally leading robot leader. The participants were asked to work on a task described by the robot. In the experiment, the robot introduced itself as the leader of a marketing department and asked the participants to develop a marketing strategy for a specific product it described. I examined differences between participants experiencing the transformational, transactional, or minimal robot leaders on task engagement and mediating factors like affect or stress. With this study, I contribute novel findings into how human followers experience being led by a robot leader and working for such a leader for a brief period. Furthermore, I show that differences between robot leader implementations based on human leadership styles differentially affect human followers.

To investigate RQ3 (Chapter 4), I present a conceptual paper on leadership assessment in VR. I formulated 11 concrete recommendations for scholars who are interested in the application of VR in leadership assessments. Those recommendations are based on a study's intended setting, the technology being examined, user avatars, interactable characters, the requirements for the virtual environment, and other considerations. I also described current approaches to leadership assessment by reference to their internal and external validity and their potential drawbacks. Finally, I summarized the extensive evidence on immersive VR and learning tailored to the leadership context. Lastly, Chapter 5 summarizes the key findings and discusses the implications of my research for leadership in the digital age.

2. Robot Leadership – Investigating Human Perceptions and Reactions Toward Social Robots Showing Leadership Behaviors[1]

Introduction

As digitalization proceeds, novel technologies create organizational challenges and opportunities (Schwarzmüller et al., 2018). Robots represent one promising technology; their utilization across various domains and organizations doubled between 2013 and 2018. Robots have demonstrated their ability to complete menial tasks for humans and to help in dangerous situations, such as rescue missions. Due to their constantly improving social capabilities, robots are now also being introduced into complex social situations (Reich-Stiebert & Eyssel, 2015). Leadership is one complex social interaction that is relevant to organizations. Preliminary results have hinted at the potential of social robots to engage in leadership interactions, for instance, by motivating employees to continue working on mundane tasks (Young & Cormier, 2014).

Although human-robot interaction (HRI) research has provided initial insights into the boundary conditions of robot leadership, more is needed about the appearance and behavior that would enable robots to be accepted as leaders and effective in that role. Previous studies on robot leadership have only considered a small subset of the behaviors required by human leaders. To understand which behaviors are effective in robot leaders, knowledge about human perceptions of and reactions toward specific robot behaviors is necessary (Gombolay et al., 2015; Mota et al., 2016; Westlund et al., 2016).

In this explorative, qualitative study, we inquire how human followers perceive and react to specific leadership behaviors when they are adopted by a robot, that is when a robot

---

[1] Chapter 2 is based on a paper by Cichor, Hubner-Benz, Benz, Emmerling, and Peus (2023), published in PLoS ONE.

assumes a leadership role. Our research integrates the robot-based findings from HRI research (Kim & Hinds, 2006; Reich-Stiebert & Eyssel, 2015), which has investigated the capabilities of social robots and how they are perceived, and the findings from leadership and organizational behavior research (Avolio et al., 1999; Braun et al., 2013), which has examined follower reactions to leadership behaviors in humans. We answer the following question: "What perceptions and reactions does a robot leader elicit in human followers, and how do these perceptions differ from the leadership style of the robot?".

Our research contributes to the literature on HRI and leadership. Regarding the HRI literature, our study examines human perceptions of and reactions toward robots in a largely unexplored complex social situation, specifically robot leadership. Turning to leadership literature, we investigate perceptions of and reactions toward leadership behaviors in robots, which is a new perspective because the prior literature only covers human leadership.

We begin this paper with a section on the theoretical background of the current state of research on robot leaders. We then elaborate on our method and explain how we conducted and analyzed our interviews and group discussions. Then, we describe our results, including the assumptions, reactions, ethical concerns, and future visions that we observed. Finally, we discuss our results, the study's limitations, suggestions for future research, and then finish the paper with a conclusion.

Theoretical Background

HRI research has shown that robots can be utilized successfully for specific leadership-related behaviors. Initial research on robots engaging in motivational behavior indicates that robots can use algorithms to identify the most motivated employees (Canós-Darós, 2013) and motivate participants to work on mundane tasks, where otherwise they would be likely to quit (Young & Cormier, 2014). Thus, robots seem able to measure and increase motivation levels, an essential component of leadership (Kark & Van Dijk, 2007).

Humans who have worked with a robot for an extended period have even been found to prefer robot leadership to human leadership when they perceive it to increase efficiency (Gombolay et al., 2015). Thus, humans seem to accept social influences from robots (Kiesler et al., 2008; Nitsch & Glassen, 2015; Powers et al., 2007)—an essential prerequisite for robots to be taken seriously in leadership roles.

For robots to assume leadership roles effectively, acceptance and trust are crucial. The acceptance of robots has been shown to depend on their design and behavior. The specific features that are effective depend on the task context. In casual social settings, for instance, playful humanlike robots are preferred over serious, mechanical ones (Goetz et al., 2003). Robots in leadership roles need to behave in a way that makes them appear trustworthy. Trust in robots is vital for human-robot interaction (Mota et al., 2016), and trust in leaders is an important component of effective follower-leader relationships (Braun et al., 2013). Therefore, trust is essential for collaborative work relationships between human followers and robot leaders (Samani et al., 2012). Initial research on trust in robots suggests that the process through which trust between humans and robots develops is very similar to how humans begin to trust one another (Mota et al., 2016). In this process, mind attribution is crucial, defined as the robot's potential to be seen as capable of experience and agency. Mind attribution increases likability (Kozak et al., 2006) and trust (Dang & Liu, 2021). While trust in technologies can decrease after technical errors become apparent (Madhavan & Wiegmann, 2007), direct and prolonged technology interactions can increase acceptance, trust, and mind attribution (Ullman & Malle, 2018).

Their assumptions also influence what humans expect from robots and how they interact with robots (Branigan & Pearson, 2006). Specifically, human assumptions about robots (Andonova, 2006), e.g., whether the robot is seen as a tool or a social actor, influence how human-robot interactions unfold (Fischer, 2006). The importance of assumptions has been demonstrated in HRI research; a generally positive attitude with limited attribution of agency has been found to make the interaction more likely (Stafford et al., 2014). In the context of robot leaders, interactions might be influenced by human assumptions about the nature of robot leadership. A robot leader might be envisioned as an emotionless machine that cannot form personal connections or as a paragon of objectivity, fairness, and competence. Crucially, negative assumptions about robots will likely decrease after successful interactions (Andonova, 2006). Moreover, when human followers assume that the robot leader is programmed and optimized to support them, they could perceive increased organizational support, which has been linked to well-being (Rasool et al., 2021; Wang et al., 2020). The concrete content of these assumptions is relevant because it determines whether and how human followers engage in and react to interactions with robot leaders.

We suggest that human perceptions of and reactions toward robot leadership will likely depend on the robot's leadership style. Research has identified certain leadership styles as being particularly effective among humans. For example, transformational leadership, the most extensively studied leadership style, combines idealized influence, inspirational motivation, intellectual stimulation, and individualized consideration (Bass, 1999). Transformational leadership has been linked to numerous positive effects on, among others, organizational commitment, creativity, engagement, and trust (Avolio et al., 2004; Braun et al., 2013; Gumusluoglu & Ilsev, 2009; Khalili, 2016; Yasin Ghadi et al., 2013) in a variety of contexts (Braun et al., 2013; Schmid et al., 2019). Transactional leadership, in contrast, is characterized by management-by-exception (active, passive) and contingent reward. It has been suggested to be particularly effective under time pressure and for tasks that do not require creativity (Bass, 1999). To date, only one study has investigated transformational and transactional robot leadership. The researchers found that using a transformational robot leader leads to higher trust, whereas transactional robot leadership caused performance in manual tower-building tasks to improve (Lopes et al., 2021). However, the effects of a semi-humanoid and social robot leader in a realistic leadership task are yet to be studied.

We explore the potential of implementing leadership behaviors in semi-humanoid social robots and the hopes and fears of followers associated with robot leadership. We programmed a social robot to display leadership behaviors to investigate different reactions to transformational and transactional robot leadership. We expected that a transformational robot leader might have more positive effects on human followers when compared to a transactional robot leader, in line with the results on human leaders that have been reported in the literature (Avolio et al., 2004; Braun et al., 2013; Lopes et al., 2021).

Method

Research Approach

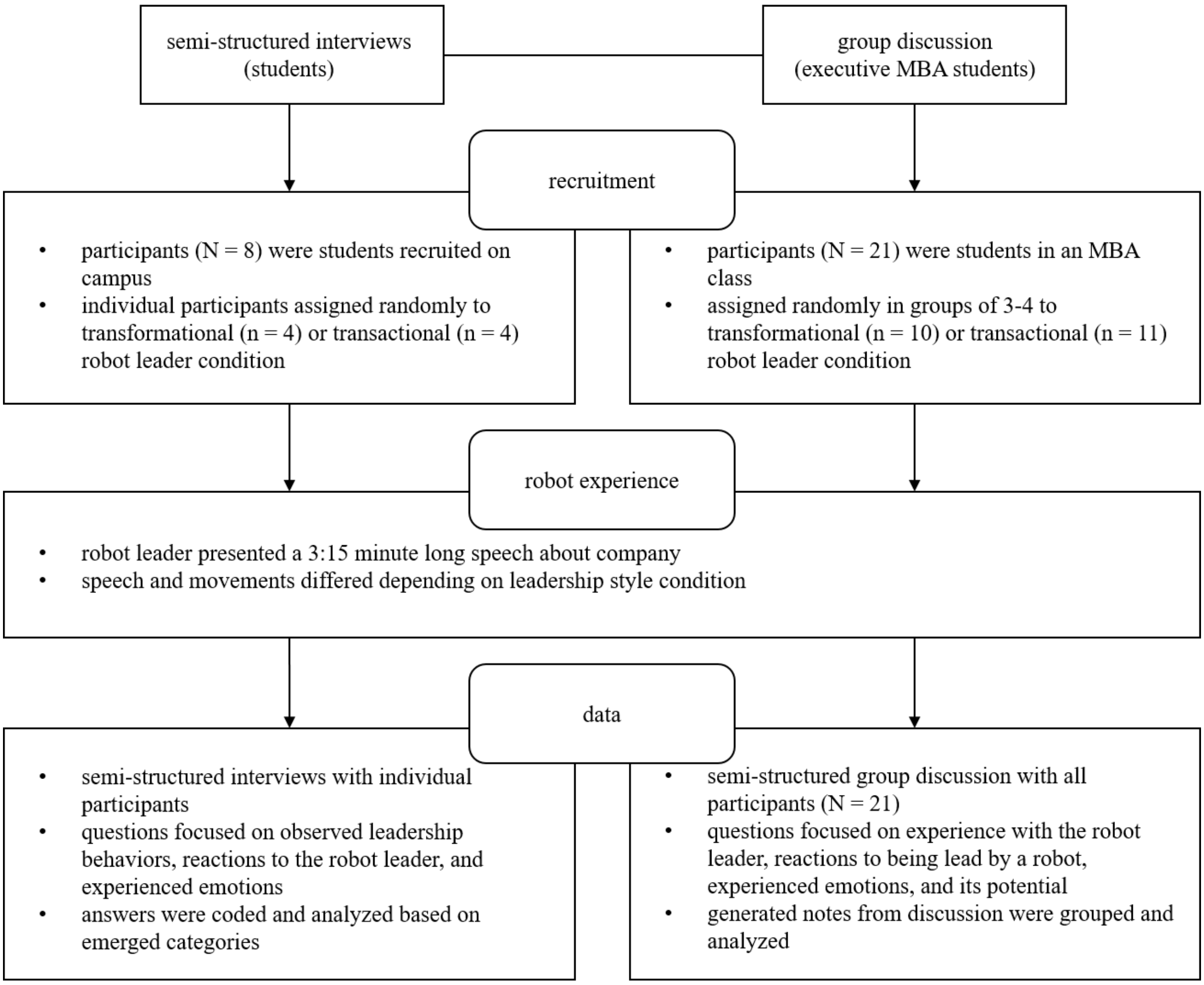
Since little research has been conducted on robot leadership behaviors, we chose a qualitative approach. Participants either saw a transformational or a transactional version of the robot. In both conditions, they received a task from the robot and subsequently worked on the task 'for the robot.' After that, we observed their perceptions and reactions. In the interview study, we informed the participants orally that their participation would be voluntary, that they could withdraw at any time, and that their answers would be recorded and evaluated to be used for scientific research. Written informed consent was obtained from the participants in the group discussions at the beginning of the study. They were also told that their participation would be voluntary, that they could withdraw at any point, and that notes from the group discussions would be used in research. Since no Ethics Committee regulates research in the behavioral experimental sciences at the institution where this study was conducted, we implemented the ethical principles for research on human participants from the Declaration of Helsinki and the German Society for Psychology (DGPS).

We integrated different data sources in two stages to obtain a differentiated representation of the participants' perceptions and reactions (see Figure 1). In the first stage, we conducted semi-structured interviews to obtain detailed insights into our interviewees' experiences of interactions with the robot leader. A male master's student conducted the interviews. That student was given clear instructions on conducting semi-structured interviews before the study and prepared the questions with the other researchers. Only the student and the interviewee were present during the interviews, which were recorded with an audio recorder. No relationship was established with the participants before the interviews, as the participants were approached on campus to participate in the study. The participants were informed that they would experience an interaction with a robot and would be interviewed afterward as part of research for a master's thesis. Each interview lasted approximately 45 minutes.

In the second stage, the first author oversaw semi-structured group discussions to observe controversial conversations between participants, which triggered in-depth reflections about experiences with the robot leader. The researcher had yet to have prior interactions with the participants in the group discussions. Still, the participants received their CVs before the group discussions and were told that the session would be on robot leadership. In addition, the participants knew that the study was being conducted as part of Ph.D. research. A research assistant and a lecturer were present during the discussions to take notes and keep track of time. The discussions lasted approximately 20 minutes.

Figure 1

Procedure Overview for Study 1 and Study 2



Our aim in the interviews and group discussions, in which we followed a phenomenological approach, was to understand what participants experience during live interactions with robot leaders (Alase, 2017). We asked questions and gave prompts related to working with a robot leader, the conditions under which a robot leader would be promising, and the features that made the robot most convincing in its role as leader. We asked the participants about their perception of the robot, whether they thought it could be a leader, how they felt while working for it, and what they thought working for a robot in a real-life setting would be like. We also compared the insights from the interviews with those from the group discussion to investigate similarities and contradictions.

Implementing Leadership Behaviors in a Robot

We programmed the social semi-humanoid robot Pepper, manufactured by SoftBank Robotics, to act as a transformational or transactional leader during a presentation. Each version of the presentation lasted approximately 3 minutes and 15 seconds. The difference between the leadership styles was emphasized by adjusting the robot's speech and movements. The transformational (Bass, 1999) version of the robot spoke of its vision of the company, its confidence in its followers' ability to complete the task, and its enthusiasm for its work while making emphatic projecting movements away from its body (e.g., throwing its hands up in the air when displaying enthusiasm). Those features are based on charismatic leadership tactics (Antonakis et al., 2011). In comparison, the transactional (Bass, 1999) robot leader focused on rewards for completing the task, setting specific goals, and specifying concrete requirements. Its movements were closer to its body, and it made more directive movements, such as shaking its head. Those movements were intended to be the opposite of charismatic leadership tactics (Antonakis et al., 2011). The transformational and transactional versions were tested with leadership experts before the first session.

In both conditions, the robot would state that it was the head of a marketing department of a botanical company that specializes in products that help customers who grow plants. After introducing the company and its product, the robot would instruct the participants to create a marketing strategy for the product based on a set of guiding questions. Developing the marketing strategy was intended to create a realistic situation in which the participants would feel like they were working 'for the robot.' We focused on the position of a leader giving a task to subordinates (i.e., we did not investigate the situation of subordinates who are presenting their work to a robot leader).

Data Collection

Interviews. We recruited eight interviewees from our networks on campus. The participants were employees (academic or educational staff) or students at the university. The participants were shown the robot's presentation. Four participants saw the transformational version, and four saw the transactional one. Then, the eight participants completed the task. Subsequently, we conducted in-depth interviews. Relying on a semi-structured approach, we started with several relevant questions (see Table 1 for a sample) and then delved into other topics whenever an interviewee made an interesting statement. We aimed to establish a natural conversation. No participants withdrew from the study. We did not conduct repeat interviews, and we did not provide recordings or transcripts to the participants. Data collection stopped when we started to identify recurring themes in the participants' responses.

Table 1

Questions asked during interviews and group discussions

Data-collection method	Question
Interviews	Transactional specific: How do you think Pepper would react if you were to make a mistake?
Interviews	Transactional specific: What did you think when Pepper said, "We do not tolerate any mistakes"?
Interviews	Transactional specific: Pepper said: "You should document what you do so I can follow your steps and identify mistakes. This will allow me to give you exact feedback on your work." Did you follow his steps exactly? Why?

Data-collection method	Question (continued)
Interviews	Transformational specific: What did you think when Pepper said: „My vision is that we, as the marketing department, can make people passionate about our product.”?
Interviews	Transformational specific: How well did Pepper define and present his vision for the future?
Interviews	Transformational specific: Pepper said: “As a leader, I always want to be a role model.” What do you think about that? To what degree do you believe that Pepper could be a role model for you as an employee?
Interviews	Transformational specific: If you were now in a team with Pepper as a leader: Would you aim to reach the goals that Pepper set for you?
Interviews	Transformational specific: To what degree do you feel that Pepper would have high expectations for you?
Interviews	Transformational specific: To what degree do you believe that Pepper would support you appropriately as a team member? Why?
Interviews	Transformational specific: To what degree do you have the feeling that Pepper adjusts himself to you and others, for example, by paying attention to your personal needs?
Interviews	To what degree did Pepper inspire you to produce new and innovative ideas?
Interviews	Do you think that Pepper would question the status quo? To what degree?
Interviews	Would you do more than what is minimally expected to reach the goal of Pepper’s department?
Interviews and group discussion	Which thoughts went through your head when Pepper was talking to you?

Data-collection method	Question (continued)
Interviews and group discussion	How did it feel when you were addressed by Pepper?
Interviews and group discussion	What was your overall impression of Pepper?
Interviews and group discussion	To what degree did you feel motivated by Pepper? Imagine that you are working in Pepper's department: To what degree would you feel motivated by Pepper?
Interviews and group discussion	Do you feel appreciated by Pepper? Why?
Interviews and group discussion	Do you think that Pepper is fair to his team members? Why?
Interviews and group discussion	Can you trust Pepper as a leader? Why?
Interviews and group discussion	Which of his behaviors did you like? Why?
Interviews and group discussion	What did you not like about his behavior?
Interviews and group discussion	How would you feel when Pepper was your leader? What would work particularly well or particularly badly?
Interviews and group discussion	If Pepper gave you strict feedback: How would you take it?
Interviews and group discussion	How well do you think can Pepper create plans and set goals for the future of his team?
Interviews and group discussion	What would you expect: How successful would a team led by Pepper be? Why?

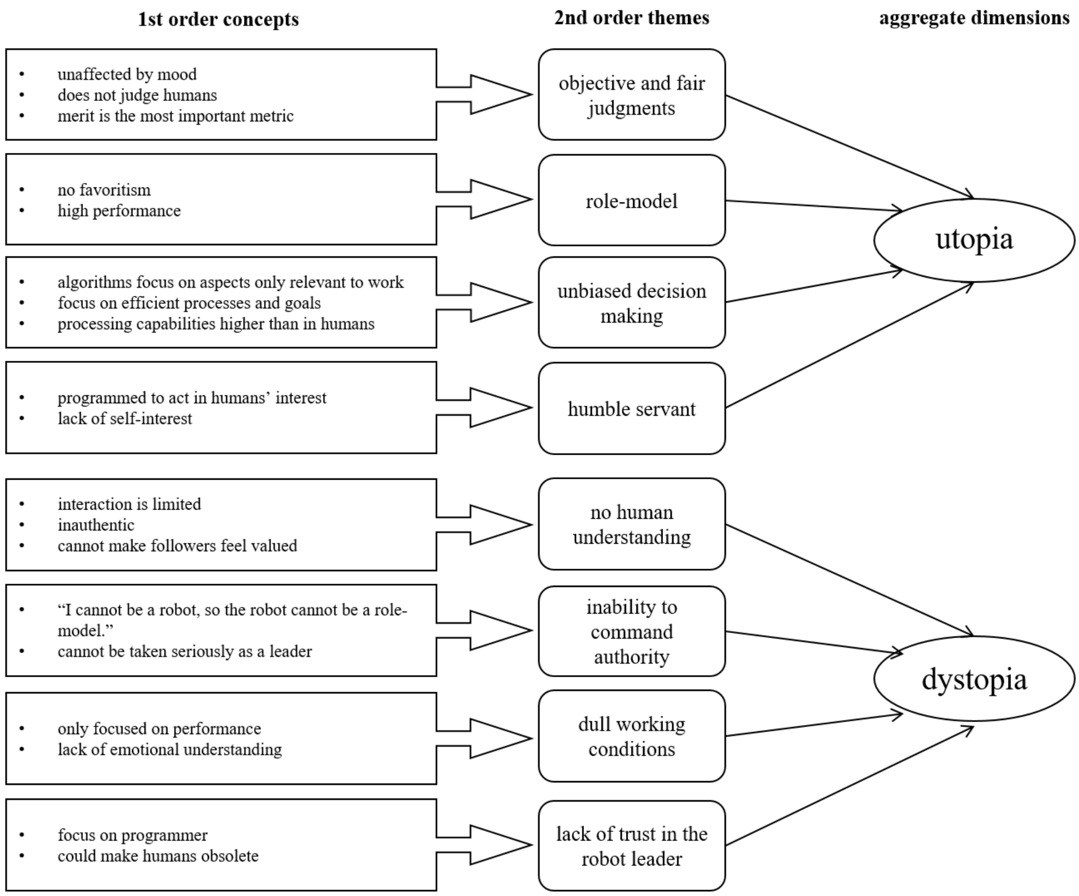
Group discussions. We recruited 21 company leaders or prospective leaders from an executive education program for the group discussions. Due to their comprehensive work experience, their perspectives on robot leadership were particularly interesting. They could compare their experiences with human leaders and their leadership approaches. The participants were presented with the robots' presentations in groups of three or four—10 were presented with the transformational version, and 11 were shown the transactional version. Then, they worked 'for the robot' individually. Subsequently, we conducted a focus-group discussion with all 21 participants. The discussions were based on a set of predefined questions that overlapped with our interview questions. We encouraged multiple participants to answer each question throughout the discussion and welcomed natural discussions between them. All the participants completed the entire procedure. We did not carry out repeat discussions, and the participants did not comment on the notes. The group discussion emphasized themes related to robot leadership and facilitating exchanges between participants. The research assistants took detailed notes pertaining to the points and themes participants made during the exchange. After we found that the group discussions did not introduce new themes to those we had already extracted from the interviews, we stopped collecting data.

Data Analysis

The first and the second author coded the transcripts of the interviews based on an iterative process using the MAXQDA software by VERBI GmbH. First, we coded statements about perceptions of robots. For example, we coded instances where the participants stated that they found the robot inspiring or intimidating. In this step, we also coded the reasoning behind answers to whether the participants thought the robot could be a leader. Second, we coded statements on participants' feelings during their work for the robot, such as their answers to the question of whether they had found themselves engaged with or bored by the task. Third, we coded the participants' thoughts about working for a robot in a real-life setting. We inquired whether the participants pictured similar situations and coded the positive and negative thoughts they mentioned. Fourth, we created codes for all emotions, fears, and envisioned scenarios the participants discussed. After the coding procedure, we structured our findings around the four categories that emerged: i) general assumptions regarding robot leadership, ii) reactions to specific robot leadership behavior, iii) emotional reactions to the robot leader, and iv) future visions and ethical concerns. By having two raters code the interview statements iteratively and assessing inter-rater agreement repeatedly, we were able to ensure the reliability of the coding process. The fit between the participants' ideas and the individual categories is visualized in Figure 2. Most perceptions and reactions appeared in the interviews as well as in the group discussions. The similarities indicate our approach's high ecological validity, especially since both data-collection methods are conducive to natural conversations and accommodate diverse perspectives. The similarities between the interview and group discussion findings also supply evidence of high test-retest reliability. The interviews were conducted initially and coded in German. We subsequently translated a selection of the statements into English to exemplify the participants' critical ideas in the paper.

Figure 2

Gioia Methodology Based on Semi-Structured Interviews and Group Discussions

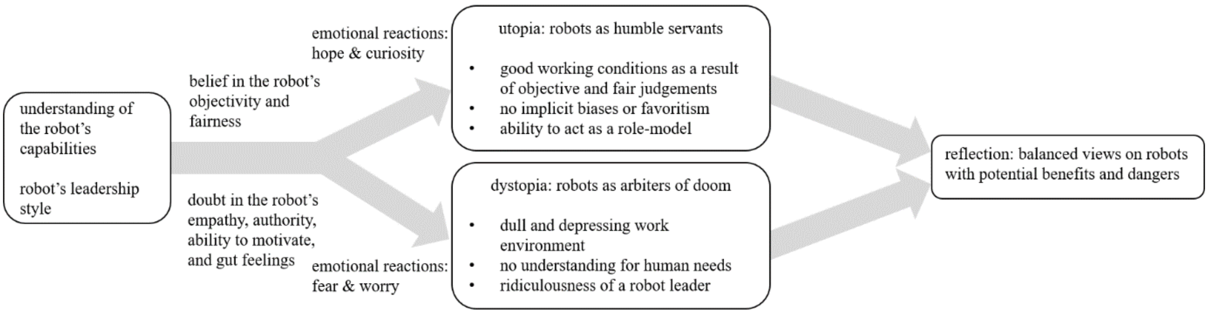


Results

The data analysis showed a combination of the robot's leadership style and the participants' assumptions that shaped subsequent perceptions. Depending on whether the participants had positive or negative beliefs, their emotional reactions were primarily characterized by either hope and curiosity or by fear and worry. These emotional reactions led the participants to imagine either a utopian fantasy in which robot leaders are humble servants to humans or a dystopian scenario in which robot leaders severely damage the contemporary work environment. The leadership style of the robot leader also affected the participants' reactions. The transformational robot leader was generally judged as a more suitable leader, and exposure to it frequently resulted in the participants adopting a more positive outlook. In Figure 3, we show the results of the effects of human assumptions about robot leadership and the leadership styles of robots on human perceptions, judgments, and reactions and the resultant fabrication of utopian or dystopian scenarios.

Figure 3

Synthesized Overview of Participants' Preconceptions, Judgments, and Reflections



General Assumptions Regarding Robot Leadership

The participants assumed that a robot leader would predominantly be used to communicate tasks and to evaluate human performance. Participants also expected a robot leader to be capable of formulating and monitoring goals based on quantifiable metrics. However, they thought personal interactions would be deficient and could not envision genuine interactions with a robot. One interviewee who interacted with the transactional robot, for example, said, "I believe that I would see Pepper as an instructor and evaluator who judges my performance fairly and relays the instructions to me. But not the interaction between [us]" (Participant #1). Assumptions of this kind were also prevalent in the group discussions, in which the participants saw the robot's potential as an instructor. They identified its primary function as relaying instructions to employees, but they could not imagine the robot fully assuming the leader role.

Interestingly, we observed that the participants' judgment of whether a robot could be a suitable leader depended on their technological knowledge and assumptions about robots. For instance, when the participants thought about the suitability of the robot for leadership roles (e.g., its empathy and capacity to demonstrate its authority or to motivate employees), they frequently doubted its ability to exhibit the necessary leadership abilities. This assumption was more frequent when the interviewees encountered the transactional robot. They were worried that robot leaders would negatively affect organizations in the future. If, however, the participants had prior technological knowledge about robots and their capabilities, they would see the potential of robot leaders. For instance, they expected that robot leadership would reduce favoritism, as one interviewee stated, After interacting with the transformational robot, one interviewee said, "Yes, I think that based on how he is programmed, he will not be unfair. And because he is not led by emotions but rather shows an objective approach, I believe that he is fair" (Participant #2).

Reactions to Specific Robot Leadership Behavior

The analysis of the reactions to working for the robot shed light on the hopes and fears that humans associate with robots in leadership positions. The participants hoped to be challenged to do their best work, especially by the transformational robot. The transformational robot's goals were taken seriously, and the participants imagined that the robot could have positive motivational effects on them and that they could use it as a de facto "role model" that guides their progress. One interviewee who saw the transformational robot stated, "Because [the robot] spoke with a certain emotion and also passion, I would say that he did it well and therefore also motivated me to develop good ideas" (Participant #3).

The transactional robot, conversely, was usually seen as creating a depressing and dull work environment. The participants found it unlikely that working for a transactional robot leader would enable them to learn and grow. They thought its leadership would be suited mainly to tasks that do not require creativity. We found that participants took the transformational robot leader more seriously than the transactional robot leader. The transformational robot leader inspired various hopes, including increased motivation, fairness, and gains in decision quality. In contrast, the transactional robot leader induced worries and fears about unfulfilled human needs, lack of creativity, and the unavailability of a margin for error.

Emotional Reactions to the Robot Leader

Irrespective of the robot leadership style that they encountered, the participants reported both hopes and worries. Their hopes were often related to the robot's lack of emotions, which some participants imagined would lead to unbiased decision-making and eliminate favoritism. Furthermore, the participants associated the robot with efficiency and machine learning capabilities due to the robot being mechanical. Many participants doubted that the robot could exhibit or detect emotions as well as a human leader. For instance, the participants frequently identified a lack of empathy as an essential problem and could not imagine how a robot could react empathically to their actions and needs. One interviewee who encountered the transformational robot, for instance, stated, "As I said, I would not be able to build an emotional connection [with the robot leader], so I would not feel appreciated" (Participant #4).

The participants generally did not trust the robot. They felt that trust can only be built through interactions on an emotional (and therefore human) level. This perception was more common after interactions with the transactional robot. The participants in the group discussions also expressed concern about the empathic capabilities of the robot and, subsequently, its suitability as a leader. The lack of empathy was related to clear communication and the neglect of personal development. For example, one interviewee, after encountering the transactional robot, commented, "And I would most likely be mad because I currently do not see how [the robot leader] could help me, for example, in my career development" (Participant #5). All in all, some participants saw the lack of emotions as fairer and more efficient. In contrast, others worried about not being understood or about the possibility of their needs not being met.

Future Visions and Ethical Concerns

We identified two basic cognitive scenarios that can emerge from encountering robot leaders: utopian fantasies and dystopian fears. If participants adopted a positive perspective, they would frequently explain how robot leaders would be more objective and fairer due to lacking biases and emotions. Consequently, these participants' emotional reactions were primarily characterized by hope for and curiosity about robot leaders. They idealized robot leaders and formed utopian visions. The participants believed that robot leaders would be humble servants that are highly objective, unbiased, capable of competent decision-making, and prepared to learn from their mistakes through machine learning. If a participant initially had a negative view of robots, they would immediately question the potential of robots as leaders. They referred to robots' lack of empathy, authority, and motivational capabilities.

We encountered ethical considerations that have to do with leadership responsibility among robots and worries about the future. The participants questioned the degree to which they interacted directly with the robot rather than as a vehicle for a programmer's instructions. They stated that they would need to know and trust the programmer to be able to trust the robot. One interviewee, who had been exposed to the transactional robot, said, "For that, I would like to get to know the programmers first. Can I trust him? No" (Participant #6). These concerns also extended to algorithms. The participants worried that they would only be led by a programmer and not by an actual leader, which, in turn, would make them doubt that any leadership was involved in the first place. For instance, one interviewee, after interacting with the transactional robot, observed, "I always find [robot leaders] difficult because it is a robot, and someone has programmed it. And then the question is, 'Can I trust the programmer and not [just the robot leader]?'" (Participant #7).

Some participants ascribed abilities and hope to robot leaders that are, to the best of our knowledge, far beyond the current horizon of technological possibility. Conversely, some participants who doubted the suitability of the robot for leadership also worried that robots could soon come to dominate humans in various areas of life. We found that, depending on the participants' assumptions and the leadership style of the robot, judgments would swiftly become polarizing. The resulting emotional reactions were permeated by fear and worry, which is why the participants in question were preoccupied with their visions of impending dystopia. In that dystopia, robot leaders would have no understanding of human needs, work would change to require no creativity or originality, and robot leaders would not be able to lead effectively due to not being taken seriously by their followers.

During the interviews, we noticed that the evaluations of many participants evolved. Participants who were initially skeptical often came to new conclusions during the interview. Sometimes, shifts were observable in answers to single questions. The participants' assumptions seemed to substantially affect their judgments of and reactions to robot leaders, causing them to quickly imagine the best or the worst scenario. However, thinking about robot leaders and reflecting on their benefits and dangers made many participants reconsider their initial judgments and develop a more nuanced perspective by considering both the positive and the negative aspects of robot leadership.

Discussion

In this study, we implemented leadership behaviors in social robots and investigated human perceptions of and reactions to those behaviors. Based on our findings, we conclude that assumptions and specific robot leadership styles influence judgments of and reactions toward robot leadership and determine future visions. Depending on their initial understanding of robots' capabilities and leadership styles, the participants evaluated robot leadership either positively or negatively. We identified two types of future visions concerning robot leadership. In utopian scenarios, robot leaders are imagined as humble servants to humans; in dystopian scenarios, robot leaders were seen as harbingers of doom.

The ethical challenges of using robot leaders were central to all evaluations. The worries that the dystopian scenarios reflect indicate that uncertainty about the programmer's role in the behavior of a robot leader and fears about lack of empathy resulted in some questions being left unanswered for our participants. Specifically, the possibility of robot leaders being incapable of understanding human needs incited resistance to the notion of robot leadership. Crucially, while initial ideas about robot leaders revolved around utopian or dystopian extremes, these tendencies were tempered by the reflective processes that the interviews and the group discussions induced. During those interviews and discussions, the participants re-evaluated their initial reactions and developed more balanced views about the advantages and perils of robot leadership. The participants frequently pointed out that they were hopeful because robots have the potential to be more objective and fairer than human leaders. However, they still thought that robot leaders might not solve all the problems traditionally attributed to human leaders.

Implications

Our study is a first step towards understanding human judgments of and reactions to robot leadership and contributes to the literature on leadership and HRI in three ways. First, we provided initial insights into the robot- and human-based factors that influence judgments of and reactions to the leadership behaviors of robots. The participants could imagine being motivated by a robot leader and were willing to trust it under appropriate conditions, which confirms findings from prior research (Canós-Darós, 2013; Mota et al., 2016). Our results emphasize the central role of assumptions in human-robot interactions, in line with previous HRI studies (Fischer, 2006). In the specific context of robot leadership, we found that initial positive or negative ideas about robots generally affect subsequent judgments of the potential of robot leadership and intuitive predictions about the future.

Second, our findings indicate that human perceptions of robot leaders mirror human perceptions of human leaders. The participants in our study frequently stated that the transformational robot leader could convey a compelling vision. They expected the robot to support their growth by posing appropriate challenges and saw it as a potential role model from which they could learn. Some other participants stated that the transactional robot might be unempathetic and dull. These differences indicate that the transformational robot induced more positive reactions than the transactional robot leader.

Third, we found that detailed reflections about the role of robot leaders can fundamentally affect the tendency of humans to adopt a balanced and nuanced perspective on robot leaders. This finding indicates that balanced discussions and reflections are needed to develop nuanced viewpoints on the issue, especially in the context of ethical concerns critical to robot leaders' perception. Reflecting on one's views and second-guessing one's judgments causes one to realize that robot leaders have certain advantages, such as objectivity and fairness, while also lacking empathy and being incapable of taking individual considerations into account, which makes it unlikely that they will replace human leaders altogether.

Limitations and Future Research

Our study is based on a small sample and only includes explorative analyses; future research needs to elaborate and replicate our results. Crucially, future research should aim to test leadership-style differences in robot leaders and differences that are due to human assumptions empirically to verify and extend the findings from our work, which is explorative and qualitative. Moreover, leadership styles other than transformational and transactional leadership (Bass, 1999), such as constructive and destructive leadership (Krasikova et al., 2013), and their effects on followers' task engagement and performance should also be investigated. Since we focused predominantly on perceptions of and reactions to robot leaders, further research should explore how specific perceptions and responses are formed and identify how personal characteristics are influenced. Additionally, investigating the reactions of the robot leader if participants hand in their results to it for evaluation is also a promising avenue for future research.

Future research should account for cultural differences and their influence on robot exposure. The same is true of prior experience with human leaders. Such research could indicate how robot leadership behaviors can be adapted to ensure that they fit the needs of employees, which may differ across individual preferences, personality types, and cultural backgrounds. Further studies should also examine the technological savviness of participants, for instance, by measuring traits such as affinity for technology interaction (Franke et al., 2019) and investigate its connection to assumptions and concerns about robots. This matter is beyond the scope of the present study.

Conclusion

We observed an interplay between the leadership style of a robot and its human followers' assumptions about robots in general. That interplay shapes the human followers' perceptions of and reactions toward robot leadership. Humans who interacted with the transformational robot leader or proceeded from favorable assumptions about robots adopted an optimistic view, and the potential of robot leaders to be fairer and more competent than humans was more salient to them. In contrast, those who encountered the transactional robot leader, who proceeded from unfavorable assumptions about robots or who were less confident of the robots' capabilities, worried about the dangers of robot leaders and their potentially harmful effect on society in the future. Nevertheless, subsequent reflection on the negative and positive aspects of robot leaders and their interactions yielded a more balanced view of robot leaders and their implications for the future.

Our research shows how humans perceive and react to robot leaders. It can thus open novel vistas on how organizations can integrate robot leaders effectively and safely into their work environments. Whether robot leaders will ultimately become humble servants of humans or harbingers of doom remains to be discovered. The available evidence indicates, however, that hopes and fears emerging around robot leadership rely less on the technology in question and much more on us—the human individual interacting with the technology.

3. I, Robot – or Leader? The Effects of a Robot's Transformational and Transactional Leadership Behaviors[2]

Introduction

Leadership research has identified several leadership styles and their effects on followers and organizations (Gandolfi & Stone, 2018). For example, transformational leadership has been shown to exert more positive effects on followers' work engagement relative to other forms of leadership in various contexts (Yasin Ghadi et al., 2013). However, whether leadership behaviors that have been studied extensively in humans have the same effects when implemented in robots remains unclear. Interactions between humans and robots follow unique mechanisms (Bartneck & Forlizzi, 2004). Consequently, findings from studies of human leaders are unlikely to be transferable to the context of robot leaders without adaptation. For instance, robot acceptance increases with human likeness, but this relationship is reversed at a certain point, a phenomenon known as "the Uncanny Valley" (Mori et al., 2012). Due to this constraint, leadership behaviors that seem inherently human, such as communicating enthusiasm for an inspiring vision and acting as a role model (Bass & Avolio, 1994), may not be suitable for robot leadership.

While the literature on human-robot interaction (HRI) comprehensively explains the factors influencing human perceptions and experiences during engagements with robots, knowledge about robot leadership and the effects of specific leadership behaviors in robots is almost nonexistent. This deficit is troublesome because robots are already assuming roles that exert social influences on humans (Canós-Darós, 2013; Samani et al., 2012). Scientific knowledge about robot leadership is needed urgently for the opportunities that recent

[2] Chapter 3 is based on a working paper by Cichor, Hubner-Benz, Emmerling, and Peus (2023), currently being prepared for submission.

technological advancements in robotics have generated to be leveraged responsibly and for the potentially adverse effects of utilizing robots in leadership roles to be avoided.

Integrating the leadership literature with the HRI literature, more specifically by drawing on full range of leadership theory (Antonakis et al., 2003) and the Uncanny Valley of the mind (Stein & Ohler, 2017), we develop a model of robot leadership that explains the effects of specific robot behaviors on human followers. We show how transformational and transactional leadership behaviors can be implemented in robots. We argue that transformational leadership behaviors in robots stimulate positive affect, perceptions of likability, and experiences of psychological safety in humans. In contrast, transactional behaviors enable negative affect, feelings of discomfort, and stress. These mechanisms explain how the nature of the robot's leadership behaviors determines the human followers' work engagement.

To empirically test the effects of robots' leadership behaviors on human followers, we conducted a between-participants experimental study in which human participants engaged with and worked for a robot leader, which was set up to exhibit different leadership behaviors. We implemented three leadership styles, namely transformational, transactional, and minimal leadership, into Pepper, a semi-humanoid robot that SoftBank Robotics manufactures. The 218 participants in our study engaged with the robot, which introduced itself as a leader, exhibited leadership behavior (which varied depending on the experimental condition) and assigned a task to the participants. After working for the robot, the participants completed a questionnaire so that we could investigate their perceptions and experiences.

Our study makes three important contributions to the literature on leadership. First, in shedding light on the effects of the specific leadership behaviors that robots can adopt, our study introduces a novel perspective on leadership behavior and opens a new avenue for research. While numerous studies have investigated the effects of transformational and transactional leadership in humans (Eagly et al., 2003; Judge & Piccolo, 2004), our study is among the first to analyze the effects of specific leadership behaviors in semi-humanoid social robots theoretically and empirically. Second, our model explains the mechanisms by which the leadership behaviors of robots affect their human followers. We show the effects of those behaviors on positive or negative affect, human perceptions of the robot as likable or as discomforting, and experiences of psychological safety or stress. The description of the positive and negative effects of robot leadership illuminates the specific opportunities and threats that digitization entails, which organizations must address but which have so far proven elusive. Third, by introducing findings from the HRI literature to the leadership literature, we arrive at a comprehensive understanding of robot leadership. Specifically, by considering Uncanny Valley effects in the context of robot leadership, our study illuminates leadership mechanisms that are unique and particularly relevant to leadership by robots.

Theoretical Background

Robots, Social Robots, and the Uncanny Valley

Robots and HRI. Robots are machines that are built and programmed to complete tasks autonomously. They can support humans in menial or dangerous tasks, such as tools that efficiently complement manufacturing procedures (Urhal et al., 2019) or increase the precision of medical operations (Simaan et al., 2018). Robots' expanding functions and features and increasing autonomy have amplified their potential in highly complex domains (Reich-Stiebert & Eyssel, 2015). As the level of automation increases, humans have less control over their robot peers and are thus required to accept and trust them (Beer et al., 2014; Glikson & Woolley, 2020; Koenig et al., 2010).

Social Robots. The relevance of human-robot social interactions has increased since the advent of social robots (Fong et al., 2003). Social robots are designed to interact with humans, frequently over long or recurrent periods (Leite et al., 2013). Their effectiveness in social situations depends on their ability to evoke emotions and response patterns like those observed in human-to-human interactions whose course otherwise depends on apparent moods and their intensity (Kirby et al., 2010). Extending the ability of a social robot to interact with humans in more complex social domains, for example, by introducing emotions (Kozima et al., 2008) or engaging in simple forms of humor (Garcia et al., 2017), can improve the outcomes of interactions. Social robots are often (semi-) humanoid because adding humanlike features has increased the acceptance of robots (Walters, 2008). In many circumstances, adding elements such as heads, torsos, or arms is beneficial because the resulting increase in familiarity makes acceptance more likely (Walters et al., 2009).

Social robots increasingly thrive in social environments (Leite et al., 2013). Shopping malls have repeatedly embraced social robots as components of customer experience and allowed visitors to engage with robots to obtain overviews of areas or to ask for directions to specific stores (Niculescu et al., 2013). While using social robots in public spaces can attract potential customers, the novelty of the experience might wear off and give rise to questions about the long-term meaningfulness of social interactions between humans and robots (Niemelä et al., 2017; Smedegaard, 2019). Furthermore, social robots have been used successfully in domains such as education (Belpaeme et al., 2018; van den Berghe et al., 2019) and elderly care (Broekens et al., 2009; Leite et al., 2013).

The Uncanny Valley. While human likeness has been shown to improve human-robot social interactions, this effect is reversed at a certain point, a phenomenon that is known as the Uncanny Valley. The term describes the experience of humans when an unnatural entity, i.e., a virtual character or a robot, suddenly appears to be eerie because it is too similar to a human (Mori et al., 2012). Many explanations for the Uncanny Valley exist, and the most prominent among them is based on the consistency of human realism (MacDorman & Chattopadhyay, 2016). Since human features are used to increase familiarity in human-robot interaction, the human-realism theory of the Uncanny Valley would require, for example, that all aspects of the face of the robot, while resembling those of a human face, be clearly recognizable as robot features. If, however, one feature is predominantly humanlike while other features are more typical for a robot, the human realism would be inconsistent, and, thus, the Uncanny Valley effect would be induced (MacDorman & Chattopadhyay, 2017). Robots should be designed to ensure consistency across all humanlike features for the Uncanny Valley to be avoided.

The Uncanny Valley is important because it affects the acceptance of a robot, which, in turn, can reduce the likelihood that the robot will be seen as a technology worth using (Davis, 1985). The questions of how social robots can be used in organizations and how their human peers relate to them have been investigated in various contexts. For instance, humans seem to categorize social robots into social groups, similarly to how they would their human peers (Westlund et al., 2016). When robots are perceived as autonomous, humans ascribe credit or blame to them in a way that goes beyond the ascription of blame for technological malfunctions; instead, the focus is on how humans would otherwise blame human teammates (Kim & Hinds, 2006). These findings indicate that humans consider robots as accountable partners in interactions. Thus, we raise the question of whether robots can be accepted not only as coworkers but also as leaders.

Leadership Behaviors in Robots

The initial studies on robots that exhibit leadership behaviors suggest that robots and algorithms can assume responsibility for leadership tasks, such as identifying the most motivated employees (Canós-Darós, 2013), and that they have enough authority to convince humans to continue performing a mundane task when they show intentions of quitting (Young & Cormier, 2014). When working with a robot, human participants report liking being led by a robot when that leadership increases efficiency (Gombolay et al., 2015).

Previous research has also identified crucial aspects of robot behavior implementation. Acceptability depends on whether the human likeness and the robot's demeanor match the task context. For example, playful humanlike robots are preferred over serious, mechanical robots in social settings (Goetz et al., 2003). Studies have also shown that humans develop trust in robots in the same way they form trust in humans (Mota et al., 2016).

Proceeding from the assumption that humans can trust robots and might accept their guidance, we investigate the potential of implementing leadership behaviors in social robots in organizational contexts. To that end, we draw on leadership literature analyzing human leadership behaviors and inquire which findings may be transferable to robot leadership and what adaptations are necessary.

Research has identified specific leadership styles, including transformational and transactional leadership, as influential in various domains (Eagly et al., 2003; Judge & Piccolo, 2004). Transformational leadership is a leadership style that has been studied extensively. It combines elements of several other positive leadership styles, and it has been linked to a multitude of desirable outcomes that affect, among others, organizational commitment, creativity, engagement, and trust (Avolio et al., 2004; Braun et al., 2013; Gumusluoglu & Ilsev, 2009; Khalili, 2016; Yasin Ghadi et al., 2013). Transformational leadership is characterized by idealized influences, inspirational motivation, intellectual stimulation, and individual consideration (Bass, 1999). Transactional leadership, conversely, entails management-by-exception (active, passive) and contingent reward (Bass, 1999) and—in contrast to transformational leadership—is suggested to be particularly efficient under time pressure and for tasks that do not require creativity. However, whether these differential effects of leadership styles are obtained when a robot is used is an open question. With this in mind, we studied how transformational and transactional robot leadership influences human followers.

The first study on adopting concrete leadership behaviors by robots indicates that transformational and transactional leadership have different influences on productivity and engagement in a tower-building task (Lopes et al., 2021). The researchers had participants work on a tower-building task as instructed by a social robot, which had a semi-humanoid head but no other humanoid features, such as arms or legs. The authors manipulated leadership styles by varying the speech of the robot. They explored the effects on team productivity, team engagement, role ambiguity, and trust measures. While this study yielded interesting insights into the impact of robot leadership behaviors when human followers work on a clearly defined task, organizational contexts often require leaders to harness the creativity of their employees such that they can come up with novel and innovative solutions to complex problems (Gumusluoglu & Ilsev, 2009; Hughes et al., 2018; Klijn & Tomic, 2010). Creative tasks are more likely to remain within the purview of humans because they are more difficult for machines to execute. To generate knowledge about robot leadership in organization-related tasks, we inquire how robot adoption of different leadership styles affects followers' task engagement in creative organizational tasks.

We investigated human followers' perceptions of and reactions to a semi-humanoid robot leader that can move its torso and arms to perform human gestures that accord with descriptions of specific leadership styles. The robot also spoke to the participants. Thus, our theory of robot leadership acknowledges recent developments in robotics and the large body of literature on the effects of specific leadership styles on innovation tasks, which are most relevant to organizational success. In the following pages, we elaborate on the mechanisms by which the leadership behaviors of robots influence human work, i.e., effects on human followers' task engagement via affective reactions and perceptions of the robot leader.

Hypothesis Development: The Effects of Robot Leadership Behaviors on Human Followers

One mechanism through which the effects of different leadership behaviors on human followers can be explained is their experienced ***positive and negative affect*** when the human follower is working for the robot leader. Positive and negative affect on leaders have been linked to follower mood because the leader's affect influences affective reactions among followers (Gooty et al., 2010). Similarly, a second theory focuses on the effects of contagion mechanisms. Affective displays on the part of the leader directly influence outcomes for followers (Gaddis et al., 2004). Crucially, transformational and transactional leadership have been shown to positively and negatively affect followers directly (Rowold & Rohmann, 2009). In studies with human followers, transactional and transformational leadership have been connected to positive affect (Lyons & Schneider, 2009). However, we argue that the effects will likely differ for robot leaders. Horstmann and Krämer (2020) found that a robot is perceived to be more sociable and competent when presented as an assistant rather than a competitor, indicating fewer positive reactions towards a robot that can potentially be a social threat. Therefore, we expect engagement with a transactional robot that exhibits directive behaviors to induce negative affect. Accordingly, while we expect the inspiring transformational robot to have the same positive effects on affect as a human transformational leader, the directive transactional robot should be associated with negative affect.

H1: Human followers experience (a) more positive affect when they engage with a transformational robot and (b) more negative affect when they engage with a transactional robot.

The influence of the robot leader's behaviors on human followers is likely to be shaped by the human followers' evaluation of the robot leader's traits, including its ***likability*** and the ***discomfort*** that it induces. Research indicates that the likability of human leaders is an important metric (Eichenauer et al., 2022; Rojahn & Willemsen, 1994), as it influences important leadership outcomes, such as performance (Dionne et al., 2002). Similarly, the perceived likability of social robots has been studied extensively (Cameron et al., 2021; Ling & Bjorling, 2020). If robots are seen as likable by their social interaction partners, the likelihood that they can be judged similarly to humans increases (Bartneck et al., 2009). Beyond the measures directly relevant to humans, HRI researchers have found evidence of a unique negative factor in evaluations of robots, namely discomfort. Discomfort seems to be related to applications of human stereotypes and standards, such as human likeness, to robots (Carpinella et al., 2017), and it could be related to the eeriness that is associated with the Uncanny Valley effect (Mori et al., 2012). We expect that, in robots, the transactional leadership style, with its directive components, could be seen as less likable and induce more discomfort than the transformational leadership style. Therefore, we expect likability to be at its highest and discomfort to be at its lowest when the transformational robot leader is employed.

H2: Human followers experience (a) the transformational robot as more likable and (b) more discomfort when they engage with the transactional robot.

Interactions with human leaders and with robots are capable of causing ***stress*** or, conversely, ***perceived safety***. In a meta-analysis, Harms et al. (2017) explored how human leader-follower relationships substantially affect the stress outcomes of followers. In collaborations between humans and industrial robots, features such as the appearance of the robot and its movements have been found to impact mental stress in humans (Arai et al., 2010; Lu et al., 2022). In interactions between humans and social robots, perceived stress has been connected to robot likability and intention to use a robot (Ling & Bjorling, 2020). In HRI research, a concrete construct has been developed to measure evaluations of robots by reference to stress. That construct is perceived safety (Bartneck et al., 2009; Lu et al., 2022). Since transformational leadership has been linked to lower stress levels in followers than transactional leadership (Lyons & Schneider, 2009; Rowold & Rohmann, 2009), we expect transformational leadership to lead to lower stress and higher perceived safety.

H3: Human followers experience (a) more perceived safety when they engage with the transformational robot and (b) more stress when they engage with the transactional robot.

In line with H1, H2, and H3 and the findings by Lopes et al. (2021), we expect the use of the transformational robot leader to be associated with higher task engagement. This effect is likely to be mediated by the dependent variables of this study. The whole model is displayed in Figure 4.

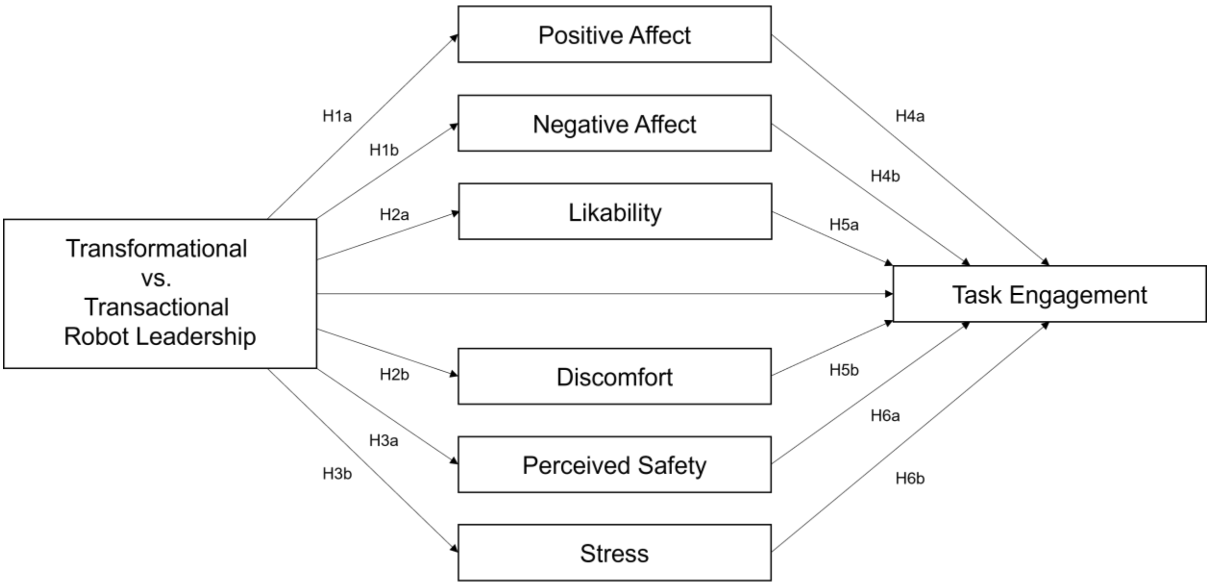
H4: Human followers experience higher task engagement when they engage with the transformational robot, mediated by (a) positive affect, and lower task engagement when they engage with the transactional robot, mediated by (b) negative affect.

H5: Human followers experience higher task engagement when they engage with the transformational robot, mediated by (a) likability, and lower task engagement when they engage with the transactional robot, mediated by (b) discomfort.

H6: Human followers experience higher task engagement when they engage with the transformational robot, mediated by (a) perceived safety, and lower task engagement when they engage with the transactional robot, mediated by (b) stress.

Figure 4

Research Model with Dependent Variables and Hypotheses



Method

To explore our research question, we implemented transformational, transactional, and minimal leadership behaviors based on modifications to speech and movements in a semi-humanoid social robot. We then conducted a between-participants experiment in which our participants worked on a creative task for the robot leader. We preregistered the study in the Open Science Framework (Cichor et al., 2019, August 29).

Design and Manipulation

We manipulated leadership behaviors by designing a text vignette implemented in the robot. Specifically, our robot was designed to act as the leader of the marketing department of a fictitious company called Pocket Gardener. In the vignette, the robot leader introduces the company and describes how Pocket Gardener created a product that is supposed to help customers grow plants without prior botanical knowledge. In the second half of the vignette, the robot leader describes a concrete task. The participants are asked to develop a marketing strategy for the botanical product based on five guiding questions.

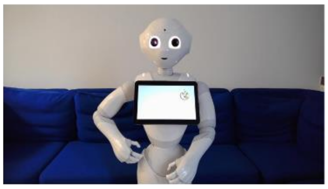
We developed three versions of the vignette, in which we controlled the information about the organization, the product, and the task. However, we altered each version's wording and tone to capture one of the three leadership styles we examined (i.e., transformational and transactional leadership, as well as minimal leadership, which serves as a control condition). We adjusted the vignettes based on the Leadership Style Assessment (Peus et al., 2013) and the definitions of the concrete dimensions of transformational leadership (i.e., idealized influence attributes, idealized influence behaviors, and inspirational motivation) and transactional leadership (i.e., contingent rewards and active management-by-exception active; (Avolio et al., 1999). We omitted other dimensions (i.e., intellectual stimulation, individualized consideration, and passive management-by-exception passive), as their inclusion was unfeasible to be shown in a brief experience with a robot leader. For our minimal leadership control condition, we focused primarily on factual information and providing suggestions without overemphasizing inaction. We aimed to avoid resemblance to the *laissez-faire* leadership style, which is experienced as a negative form of leadership (Skogstad et al., 2007).

We implemented the vignettes that we created in Pepper, a semi-humanoid robot that is manufactured by SoftBank Robotics. Compared to other commercially available social robots like Nao, which is made by the same company, Pepper is taller, which makes it more suitable for a leadership role because human participants can see it at eye level when seated. To have the robot leader pronounce the text of the vignette, we used a Python script and the Google Text-to-Speech API. The robot spoke with a male voice.

The second component of the robot leader's implementation included differential movement. Being a semi-humanoid robot, Pepper can be programmed to move its head, torso, and arms freely. As far as transformational leadership is concerned, we focused primarily on the resemblance between that style and charismatic leadership, and we drew on examples of charismatic-leadership video manipulations from the work of Antonakis et al. (2014). We identified the gestures from the charismatic leadership videos that appeared to be the most directive and implemented them in the transactional robot. In the minimal leadership condition, we included comparatively fewer movements. Moreover, all movements were closer to the body of the robot. An overview of the different conditions and the differences in speech and movement is displayed in Figure 5. We limited the duration of the robot presentations to approximately 3 minutes and 15 seconds per condition.

Figure 5

Overview of Experimental Conditions with Differences and Examples

	Transformational	**Transactional**	**Minimal (Control)**
Speech	vision, inspiration, motivation	not tolerating mistakes, asking to follow directions exactly	matter-of-fact, passive
Movement	large movements, moves hands high and outwards	points at participants, places fist in hand, props hands up on hips	close to the body, minimal
Sample Pose			

Prior to data collection, we conducted a pretest with six leadership scholars, each of whom saw one implementation (i.e., two transformational, two transactional, and two minimal). The experts' evaluations on the Multifactor Leadership Questionnaire (MLQ) and their explicit statements (Avolio et al., 1995) made it clear that all six had identified the leadership style that they had been shown correctly and that they were unsure and had failed to perceive direct leadership in the minimal leadership condition. Once the pretests were complete, we conducted a three-factorial (transformational, transactional, and minimal robot leader) between-participants experiment.

Sample

We conducted the experiment at distinct time points between July 2019 and November 2022. The first three sessions took place in 2019, and we conducted the last eight in late 2021 and 2022. The COVID-19 pandemic made experimental in-person sessions infeasible, necessitating an intermission in the data-collection procedure between March 2020 and August 2021. In total, we included 218 participants, whereas the participants in the first three sessions were Executive Master of Business Administration (EMBA) students (n = 60), and the remaining participants were predominantly business-oriented Master students (n = 158). Among the participants, 47.7% were aged between 18 and 24, 35.8% were aged between 25 and 34, 13.8% were aged between 35 and 44, and the remaining 2.8% were 45 or older. Regarding gender, 45% of the participants were female, 54.1% were male, and 0.9% identified with another gender or did not want to disclose such information.

Procedure and Materials

The data collection procedure was integrated into executive MBA and general university courses. We began each data collection session with a brief introduction to the topic of robots. After the introduction, each student would draw a card to be randomly assigned to groups of three to four individuals. We opted for a group setting to closely place the participants in a situation that resembles respective situations in an organizational context. The participants experienced one of the three implementations of the robot in different time slots. After the experiences, the participants were given a link to access the task and the questionnaire online, where each participant completed the task and the questionnaire individually.

The questionnaire consisted of a written consent form, a manipulation check, the marketing task, various quantitative scales, and questions on demographics. Completing the questionnaire took 38 min 53 s on average, with a standard deviation of 10 min 21 s. The task was implemented so that the participants had precisely 10 minutes to develop a product marketing strategy. The five guiding questions the robot gave the participants were also included in the questionnaire as a reminder.

Since we focused on transformational and transactional leadership by robots, we included several dimensions from the Multifactor Leadership Questionnaire (Avolio et al., 1995) as part of a manipulation check. For transformational leadership, we included the MLQ dimensions of idealized influence attributes (four items, e.g., "Instills pride in me for being associated with him"), idealized influence behaviors (four items, e.g., "Talks about their most important values and beliefs"), and inspirational motivation (four items, e.g., "Talks optimistically about the future"), with a Cronbach's alpha of .89. For transactional leadership; we included the following dimensions: contingent reward (four items, e.g., "Provides me with assistance in exchange for my efforts") and management-by-exception active (four items, e.g., "Keeps track of all mistakes"), with a Cronbach's alpha of .80.

Regarding the dependent variables, we included measures for task engagement, positive affect, negative affect, likability, discomfort, perceived safety, and stress. Unless explicitly stated otherwise, all scales were measured on a seven-point Likert scale, with the available answers ranging from "strongly disagree" to "strongly agree." We used a modified version of the Utrecht Work Engagement Scale–9 by Schaufeli et al. (2006) for task engagement. The items were adjusted to refer to the task instead of a job ("I was enthusiastic about the task"). We also removed one item before initiating data collection ("When I get up in the morning, I feel like going to work") because it did not fit the task context. The Cronbach's alpha of the modified scale, 0.82, is high.

We measured positive and negative affect with the short form of the Positive and Negative Affect Schedule (Thompson, 2007). This scale includes 10 adjectives that describe a participant’s emotional state. The sample items for positive affect include “alert” and “inspired," and the sample items for negative affect include "upset” and “afraid.” Cronbach’s alpha for positive affect is .66, and Cronbach’s alpha for negative affect is .75.

We used the semantic differentials from Bartneck et al. (2009) to measure the likability of the robot leader. Likability is a subdimension of the Godspeed indices and includes five adjective pairs, such as “awful-nice” and “unkind-kind.” At .88, Cronbach’s alpha for the likability measurement is high. For discomfort, we chose the discomfort subdimension of the Robot Social Attributes Scale by Pan et al. (2017). This scale measures the perceived discomfort that a robot induces based on six adjectives, such as “awkward” and “strange,” which are rated on a seven-point Likert scale. The responses range between “definitely not associated” and “definitely associated.” The Cronbach’s alpha for the discomfort scale is .80.

To measure perceived safety, we used another three-item dimension from the Godspeed indices (Bartneck et al., 2009). This scale consists of pairs of adjectives, such as “anxious-relaxed” and “agitated-calm.” The Cronbach’s alpha for this scale is low, at 0.55. However, removing the last item, “quiescent-surprised,” causes Cronbach’s alpha to improve substantially and to reach .89. One possible explanation for the problematic item could be that the courses from which the data was collected were being taught in Germany. Lack of familiarity with advanced English-language vocabulary may have prevented the participants from understanding the word “quiescent.” Finally, we measured stress using six items from the State-Trait Anxiety Inventory Short Form (Marteau & Bekker, 1992). The scale consists of items such as “I felt calm” or “I was tense.” The Cronbach’s alpha for this scale is .74.

Analysis

We calculated ANOVAs with Bonferroni *post hoc* tests to explore direct effects in our manipulation check and on our dependent variables. For the mediation analyses, we used the PROCESS macro for SPSS (Hayes, 2012).

Results

Manipulation Check

The ANOVAs show that the transactional robot leader ($F(2,217) = 5.83$, $p = 0.003$) was perceived as significantly more transactional ($M = 4.50$, $SD = 0.90$) than the transformational ($M = 4.06$, $SD = 1.00$) and the minimal ($M = 3.95$, $SD = 1.01$) one. The transformational robot leader ($F(2,217) = 6.45$, $p = 0.002$) did not appear to be significantly more transformational ($M = 4.97$, $SD = 0.93$) than the transactional ($M = 4.61$, $SD = 1.04$) one, but it was seen as significantly more transformational than the minimal robot leader ($M = 4.37$, $SD = 1.12$).

Direct Effects of Robot Leadership on Human Followers' Reactions and Perceptions (Hypotheses 1, 2, and 3)

First, we analyzed the post hoc comparisons by reference to all followers’ reactions and perceptions. Figure 6 displays an overview of the results.

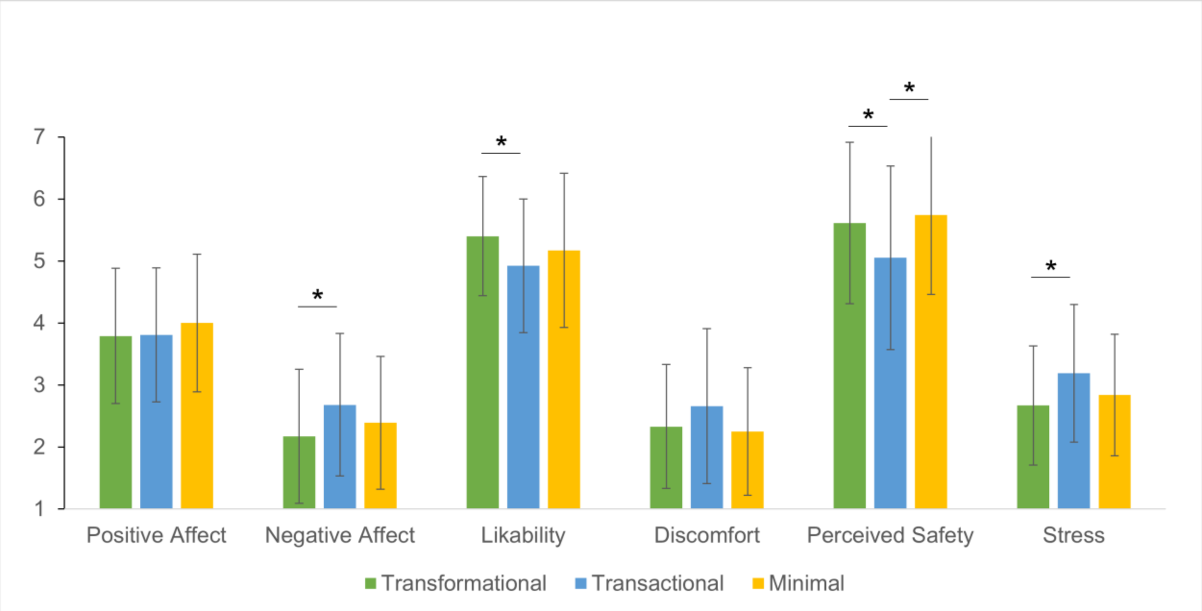
Affect. Positive affect did not differ significantly ($F(2,217) = 0.81$, $p = 0.45$) between transformational, transactional, and minimal robot leadership. Negative affect was significantly lower ($F(2,217) = 3.63$, $p = 0.028$) for transformational leadership ($M = 2.17$, $SD = 1.08$) than for transactional leadership ($M = 2.68$, $SD = 1.15$). No other effects on likability were significant. Therefore, H1a is not supported, and H1b is partially supported.

Likability and perceived discomfort. Our participants perceived the robot as significantly more likable ($F(2,217) = 3.33$, $p = 0.038$) when they engaged with the transformational version ($M = 5.40$, $SD = 0.96$), relative to the instances in which they engaged with the transactional version (*M* = 4.92, *SD* = 1.08). None of the other comparisons revealed significant differences in likability. Likewise, the differences in perceived discomfort failed to show significant differences. Therefore, we found partial support for H2a but not H2b.

Perceived safety and stress. The participants felt significantly less safe ($F(2,217) = 4.97$, $p = 0.008$) when they engaged with the transactional robot (*M* = 5.05, *SD* = 1.48) than when they engaged with the transformational one (*M* = 5.61, *SD* = 1.30) and the minimal one (*M* = 5.74, *SD* = 1.28). There was no significant difference in perceived safety between the transformational and the minimal robot leader. Stress was significantly higher ($F(2,217) = 4.76$, $p = 0.010$) among the participants who experienced the transactional robot (*M* = 3.19, *SD* = 1.11) than among the participants who experienced the transformational robot (*M* = 2.66, *SD* = 0.96). The other comparisons revealed no differences. H3a and H3b are, therefore, partially supported.

Figure 6

Mean Differences and Standard Deviations for the Dependent Variables



Mediated Effects of Robot Leadership on Human Followers' Work Engagement (Hypotheses 4, 5, and 6)

Affect. Hypothesis 4 concerns the mediating role of positive and negative affect on the relationship between the leadership style of a robot and task engagement. No significant effects were found. H4 is, therefore, not supported.

Likability and perceived discomfort. Hypothesis 5 focuses on the mediation effect of likability and discomfort in the relationship between leadership style and task engagement. The results reveal that likability has a significant indirect effect when one compares the influence of transformational and transactional leadership on task engagement (b = -0.09, 95% CI: -0.21, -0.02), which provides partial support for H5a. H5b is not supported.

Perceived safety and stress. Hypothesis 6 explores the degree to which perceived safety and stress mediate the relationship between robot leadership style and task engagement. We found stress to be a significant indirect mediator of that relationship (b = -0.10, 95% CI: -0.23, -0.001). We did not observe any statistically significant effects on perceived safety. Therefore, H6b is supported, and H6a is not supported.

Discussion

We aimed to determine how findings on transformational and transactional leadership behaviors in humans translate to robot leaders. For this purpose, we conducted an experiment in which the robot Pepper instructed participants to perform a task using transformational, transactional, or minimal leadership behaviors through speech and movement.

In line with Hypotheses 1, 2, and 3, we found that a transformational robot leader is significantly more likable and that it is perceived as being significantly safer. A transformational robot also causes significantly lower negative affect and less stress. Furthermore, a transformational robot leader is associated with lower discomfort. However, this result is not significant. As far as positive affect is concerned, leadership styles do not appear to differ. Hypotheses 5 and 6 are partially confirmed in that the likability of the robot and the participants' perceived stress indirectly mediate the effect of transformational versus transactional leadership on task engagement.

Theoretical Implications

Our study contributes initial insights into the effects of specific leadership behaviors in robot leaders. By placing human followers in a situation in which they are led by a robot leader, we found that transformational robots can have positive indirect effects on engagement in a creative task. Similar results have been observed in the context of human leaders (Yasin Ghadi et al., 2013) and in a prior study of robot leadership that focused on a simpler tower-building task (Lopes et al., 2021). These results indicate that, by mimicking a complex assemblage of robot behaviors, such as transformational leadership, followers can experience outcomes that are similar to the ones that they obtain when they are led by a human. Researchers can extend these findings to other vital constructs in leadership research by investigating different positive leadership styles, such as servant leadership (Eva et al., 2019) and ethical leadership (Bedi et al., 2016).

We also investigated the effects of robot leadership behaviors on human followers' positive and negative affect, their perceptions of the robot as likable or discomforting, and their perceptions of safety and stress. The results for Hypotheses 1, 2, and 3 align with our expectations. They show that the transformational robot leader induced lower negative affect, higher likability, higher perceived safety, and less stress than the transactional robot leader. Prior research has shown that robot behavior matters and that it is context-dependent (Gray & Wegner, 2012). With its idealized and inspirational behaviors, a transformational robot leader can be perceived as kinder and less disruptive than a robot leader that exhibits more conventional robotic and transactional behaviors.

Since interactions with social robots can be novel for humans (Smedegaard, 2019), experiencing the robot leader as less directive and strict could be less distressing, resulting in gains in likability and perceived safety and reductions in negative affect and stress. Our findings suggest that the transformational leadership of a robot intensifies feelings of safety and perceptions of likability while reducing negative affect and stress. Importantly, this finding aligns with the positive perception of transformational human leaders, indicating that the outcomes of the two may also be analogous. Current theories of the behavior of social robots primarily revolve around non-verbal cues (Huang et al., 2012; Kennedy et al., 2017) or behaviors broadly classified as social or asocial (Kennedy et al., 2015). Based on our results, those theories could be developed by including concrete behaviors from transformational leadership (e.g., speaking about visions, being inspiring, and stating values) to provide more comprehensive models of robot social behavior.

Turning to Hypotheses 4, 5, and 6, the roles of likability and stress warrant further examination. These dependent variables have a significant mediation effect on task engagement. Research has shown that perceptions of one's leader are important determinants of one's subsequent performance (Meslec et al., 2020). In particular, as the experiences of human leaders show, transformational leadership is connected to enhanced follower performance (Braun et al., 2013; Yasin Ghadi et al., 2013). It is, therefore, understandable that a robot leader that speaks about its vision for the company and the product, about wanting to support the work of an individual, and about believing in the follower's ability to perform well is perceived as more likable, which, in turn, affects task engagement positively. The halo effect, which captures the human tendency to evaluate one feature of another individual and to draw conclusions about other features from that evaluation, is essential to understanding the importance of likability (Nisbett & Wilson, 1977). Human followers might have had a much more positive image of the robot leader because of its higher likability, which may have increased their motivation to work for it.

Similarly, stress plays a crucial role in the ability to focus and perform (Berkun, 1964; Liao & Masters, 2002). Since the transactional robot leader mentioned ideas such as intolerance of errors, its strict and directive style could have influenced the followers' perceived stress. They might have felt pressured to perform exceptionally well in the task. The transformational robot leader, however, presented itself as more supportive of the participants and more interested in their ideas, creating a more relaxed atmosphere that positively influenced the followers' focus on the task.

While likability and stress indirectly mediate the effects of leadership styles on task engagement, the other variables do not. Negative affect and perceived safety, which are significantly influenced by the leadership style of the robot, do not affect task engagement. The same is true of positive affect and discomfort. One explanation for this finding is that the phenomena that the four variables capture are more transient than likability and stress. Whereas likability and stress could activate lasting impressions and effects in human followers after a brief interaction with the robot leader, positive and negative affect, discomfort, and perceived safety might lack salience after the end of an interaction to cause any differences in their subsequent work. Future research could inquire whether these effects change if the robot leader is present while the followers work on the task—it may transpire that the impressions that those effects produce persist.

Our study also transplants findings from the HRI literature into the domain of robot leadership. The notion of the Uncanny Valley captures a phenomenon whereby an entity is perceived as eerie because it exhibits behaviors that humans do not expect from it. This is the case when a robot expresses emotion believably (Gray & Wegner, 2012) or when humans perceive a virtual character as having a mind (Stein & Ohler, 2017). We placed a social robot in the role of a leader, which is different from a role that humans would generally expect a robot to perform. Consequently, we showed that the Uncanny Valley does not appear to raise concerns in the study context. Likability and discomfort are two measures that are commonly used to investigate the Uncanny Valley. Our results reveal that the transformational robot is the most likable. It can also be the most humanlike due to its emphasis on idealized influences and inspirational motivation. Furthermore, we did not find a significant difference in discomfort between conditions. It appears, therefore, that prior findings on human leaders are more applicable to robot leaders than findings on robots. The implication is that leadership style behaviors can have similar effects when enacted by robot and human leaders.

Practical Implications

The description of robot leadership's positive and negative effects illuminates specific digitization-related opportunities and threats that organizations need to address. Whenever an organization intends to employ a form of robot leadership, decisions on the design of the robot leader should be based on evidence from human leadership. The effects of the leadership styles that our robot leader adopted indicate that different robot leader implementations affect followers differently. If the followers are engaged in creative work, a transformational robot leader would likely produce more desirable outcomes than a transactional one. Practitioners would benefit from robot leader implementations that emphasize supportive behavior, which would put the robot leader in a position where it could strengthen the organization’s messaging regarding its vision, values, and the motivation of its employees.

Various digital entities that are not robots have also been created to aid humans. AI, in forms such as GPT, provides a salient example. Researchers have highlighted trust as the foundation of successful interactions between humans and computers (Glikson & Woolley, 2020). By investigating the application of leadership styles to HRI, we provided initial insights into factors that have traditionally been seen as exclusive to humans, which can be necessary for implementing virtual entities. As technologies such as GPT find their way into organizations, where they are intended to support humans, focusing on implementations that emphasize support by drawing inspiration from research on traditionally human interactions could pave the way for successful human-computer relationships in the future.

Our study was conducted with the semi-humanoid robot Pepper. Practitioners should consider the type of robot that they use for leadership carefully. Our robot leader did not induce the negative perceptions of eeriness to which the Uncanny Valley theory points, which is most likely a result of its design. It clearly includes many human features, its head, and upper body, but it is still clearly identifiable as a robot. Robots that appear machinelike might be too different from conventional human leaders. In contrast, overwhelmingly humanlike robots could be considered disturbing if their robotic features do not align with their humanlike appearance. Semi-humanoid robots could provide an optimal balance between machine-like and humanlike appearance.

Robot leaders can be replicated exactly. While the limits of their capabilities must still be explored, organizations would benefit from using robot leaders in supportive leadership positions, in which they would serve as benchmarks. For instance, poor or non-existent leadership could be avoided by having robot leaders serve as baseline leaders that repeatedly and reliably complete lower- and middle-management tasks (e.g., scheduling assignments, maintaining records, or analyzing financial data), while more complex leadership requirements remain within the human domain.

Ethical Implications

The ethics of interactions between humans and technologies are also important. Our results show that while positive influences can result from transformational leadership, negative effects such as lower perceived safety and increased stress can also be observed. The extensive public debates and studies on technologies such as self-driving cars (Holstein et al., 2018) and social media (Terrasse et al., 2019) show that responsibility can be problematic when technologies malfunction or create dangers. Can a technology bear responsibility, or is it always the designer of the technology, be it a corporation or an individual programmer, who must be held accountable? Questions such as these become more pertinent when one observes that certain robot leader implementations can cause issues. Organizations must be aware of the repercussions of their technologies' or their implementation' detrimental impacts. While promising findings can motivate and inspire openness to technology, downsides, and dangers must be considered to avoid the risk of ignorance or poor implementations, causing the potential to go unfulfilled.

One limitation of robots that are considered frequently is their lack of empathy and emotion. Managing emotions, precisely their expression, and interpretation, is essential to leadership (Antonakis et al., 2009; Van Kleef & Cote, 2022). Robots, such as the robot leader in our study, can be designed to mimic human behaviors. For example, they may appear as if they care about the successes of followers. However, technological limitations currently make it impossible for robots and AI to show and understand genuine emotion. Situating robot leaders in positions of authority and power while they remain unable to make decisions on an emotional level could therefore affect human beings negatively because their emotional needs would not be considered appropriately. The specific decision-making authority that robot leaders should possess must be evaluated to avoid dystopian scenarios in which the needs of robots are prioritized over those of humans.

Since robot leadership must be combined with AI to become fully functional in organizational contexts, the problems of AI bias are also relevant. To train AI, humans must rely on data generated from human contexts, which means that biases inherent to humans, for instance, against particular races or genders, are incorporated into the output of AI (Ahmed et al., 2021). While AI biases are already creating substantial issues in domains as varied as facial recognition and recruitment, a biased AI in a leadership position with decision-making authority could aggravate discrimination. It is paramount that organizations that intend to use robot leaders be open and transparent about the data on which their AI is trained and ensure that its quality is high (Sanclemente & Cardozo, 2022).

Limitations and Avenues for Future Research

Our study is limited in some respects. First, the participants' encounter with the robot leader was short and non-interactive. Leadership is usually a process that unfolds over long periods, with the relationship between leader and follower playing a central role (Lord et al., 1999)—in addition, making it challenging to convey the full spectrum of each leadership style in a brief presentation within the scope of an experiment. Additionally, interaction is crucial in determining how humans perceive and react to robots (Pan et al., 2017). Most experiments on leadership suffer from such limitations, and we call for further research on robot leadership. That research should focus on more extensive and genuine interactions between robot leaders and human followers. For example, having followers work on a task in front of the robot and enabling them to solicit feedback, which could be controlled through the Wizard-of-Oz method or take the form of general dummy feedback that is provided at specific points during the task, could shed light on the different mechanisms that drive experiences of working for a robot leader.

Second, we explored the topic of robot leadership with the robot Pepper. In the context of the Uncanny Valley, it is important to consider other robots as well. Pepper is one of the few commercially available social robots that are also relatively tall and whose gestures can be adjusted. These factors made Pepper the correct choice for the implementation of leadership behaviors. Differently designed robots might elicit slightly different effects, depending on their features and potential relationship to the Uncanny

Valley. As a result, scholars should consider investigating leadership by using other robots as they become available.

Third, we focused only on transformational and transactional leadership. Since robot leadership is at a very early stage of its development, focusing on leadership styles that have been studied extensively was the correct decision. At the same time, it would be interesting to see how human followers react to a robot acting as a servant leader. Depending on one's assumptions and technological affinity, one could see a robot leader as a tool that should serve them, which makes servant leadership an exciting option. Likewise, exploring human reactions to an exploitative robot leader would be of substantial scientific value. Destructive leadership styles are common among humans and might also be implemented, be it by choice or negligence, in robots and virtual agents. Inquiring whether the negative effects on human followers persist or are exacerbated when a robot leader is used would yield important insights into the full range of robot leadership behaviors.

Conclusion

Robot leaders remain a technology of the future. That future can be utopian or dystopian (Cichor et al., 2023). Given the rapid development of AI in domains such as art generation, voice assistance, and intelligent chatbots, the future might arrive sooner than expected. Despite its downsides, which include bias and lack of empathy, AI is increasingly being used in domains of high social and environmental importance, such as sustainability. At the same time, the focus has been on responsible implementations and increasing the resilience of the technology to bias and cybersecurity threats (Galaz et al., 2021). It is paramount that researchers study these fascinating technologies so that humanity is prepared for digital transformation and capable of leveraging the benefits of the technologies in question instead of succumbing to the fears they induce. Our study is a first step in this direction. While robot leaders still have much to learn before replacing human leaders, the effects of their behaviors are more like the outcomes of the actions of their human counterparts than one would expect.

4. Virtual Reality Environments for Assessing Leadership Behaviors[3]

Introduction

Effective leadership is crucial for the success of an organization, especially when it faces challenges such as digitalization and the need for remote work. Leader selection and development are thus two processes that are particularly important for organizations (Day, 2011). Accurate leadership assessment is the crucial prerequisite for selecting or developing capable leaders successfully. Leadership assessment is based primarily on questionnaires, which focus on introspective reports. This methodological monopoly has been criticized for having various limitations (Gottfredson et al., 2020; Hansbrough et al., 2015; Vigil-Colet et al., 2012). The status quo creates a need for novel approaches that can address current challenges.

In this paper, we recommend leadership assessments in Virtual Reality (VR). By implementing leadership situations in VR and evaluating the designs of the leader, the followers, and the situation, researchers can observe leader behavior directly. Considering crucial principles, leadership researchers can create virtual environments. Natural leadership is more likely to emerge from those environments than from the assessment strategies that are available at present. We describe the fundamentals to consider when implementing leadership assessment applications in VR. We also address the benefits of this approach and the problems that arise when assessing leadership in VR.

Leadership Assessment

Leadership assessment is crucial for organizational research because it allows complex leadership phenomena to be measured accurately. The dominant method for

[3] Chapter 4 is based on a working paper by Cichor, Emmerling, and Peus (2023), currently being prepared for submission.

leadership assessment is the follower report, an introspection-based measure typically collected via questionnaires or interviews (Peus et al., 2013). Follower ratings have been used extensively in leadership studies as essential data sources—it has become common practice to ask followers to rate their leader by reference to various metrics important to researchers.

The full range of leadership model has been the dominant framework for leadership research in the last few decades (Anderson & Sun, 2017; Peus et al., 2013). Transformational leadership has repeatedly been linked to positive organizational outcomes, such as creativity, work engagement, and trust (Braun et al., 2013; Gumusluoglu & Ilsev, 2009; Yasin Ghadi et al., 2013). A prominent example of a leadership assessment questionnaire that is used widely for the full range of leadership model is the Multifactor Leadership Questionnaire (MLQ). It measures a leader's style on the dimensions of transformational, transactional, and *laissez-faire* leadership (Avolio et al., 1995). By answering questions related to experiences with and perceptions of their leader, followers indicate their leader's dominant leadership style.

Scholars have criticized questionnaires such as the MLQ for being ambiguous, misleading, and unsuitable for measuring complex social phenomena (Alvesson, 2020; Gottfredson et al., 2020; Hansbrough et al., 2015). Furthermore, their popularity notwithstanding, questionnaires primarily do not account for differences in context; however, the efficacy of leadership is evaluated differently depending on the context in which it is measured (Antonakis et al., 2003). In the following pages, we focus on three major limitations of questionnaires based on introspective reports that have been repeatedly emphasized, namely i) inaccuracies of follower ratings, ii) lack of context-specificity, and iii) measuring perceptions of leadership instead of actual leader behavior.

Limitations of Current Methods

Inaccuracies of follower ratings. Most leadership studies rely on followers' ratings of their leaders. Although it is assumed that the leader's behavior determines how they are assessed, research has shown that follower characteristics explain as much of the variance in leadership assessments as leader behaviors (Wang et al., 2019). One considerable problem in using follower ratings for leadership measurement is that interindividual differences between followers' personality traits, prior experience, individual characteristics, or needs influence ratings. For instance, individuals high in trait agreeableness tend to view leaders more positively (Hansbrough et al., 2015). Another issue arises when followers are asked to rate their leaders from memory, which can be influenced by recall and mental ability (Gottfredson, 1997; Hansbrough et al., 2015). Introspective assessment, often used in questionnaires, has also been found to be inaccurate due to social desirability and individual biases (Vigil-Colet et al., 2012).

Lack of context-specificity. Questionnaires rarely account for the context of leadership (Antonakis et al., 2003; Thoroughgood et al., 2016; Vroom & Jago, 2007). Leadership situations vary considerably, affecting a leader's performance (Peus et al., 2013). The context in which leadership occurs additionally explains more of the variance in leadership-related outcome variables than the differences between leaders (Vroom & Jago, 2007). Given the criticism of this lack of context-specificity, researchers have called for a renewed focus on the context in which organizational behavior occurs (Jordan et al., 2010).

Measuring perceptions of leadership instead of actual leader behavior. Researchers have written extensively on the limitations of questionnaires based on introspective reports. Introspective data have been shown to be a poor predictor of explicit behavior (Friese et al., 2008), especially of behavior that is constrained by uncertainty, such as risk-taking (Ronay & Kim, 2006), which is of particular importance for leadership (Aronson et al., 2006). Consequently, direct observation of leadership behavior has been found to reveal more about a leader's approach to leadership than indirectly reported behavior (Bledow & Frese, 2009;

Hansbrough et al., 2015). Researchers have thus called for methods that measure actual behavior (Hansbrough et al., 2015; Wang et al., 2019).

Alternative Assessment Strategies

To address the aforementioned limitations, researchers have developed various other methods that purport to increase the accuracy of leadership assessment (see Table 2 for an overview), such as i) confederates, ii) (video) vignettes, and iii) situational judgment tests.

Table 2

Overview of Leadership Assessment Methods and Their Validity

Assessment Method	Internal Validity	External Validity	Note
Introspective questionnaire	Low	Low	Relies on inaccurate follower ratings, lacks context specificity, and only measures perceptions of leadership instead of actual leader behavior
Confederate	Low	High	Classified as deception; should only be used when no other option is available
Vignette	High	Low	Dependent on imagination and recall
Video vignette	High	Medium	Videos reduce inaccuracies that are caused by deficits in mental ability and recall
Situational judgment test	High	Medium	Behavioral options measure behavioral intention

Confederates. One controversial approach to leadership assessment entails the use of confederates in specifically prepared scenarios (Blascovich et al., 2002). Confederates are individuals, commonly actors, who are instructed to act out a particular scenario together with the participant whose behavior is being assessed.

Using confederates results in high external validity. Scenarios are performed in ways that closely resemble similar situations in the real world, meaning that participants are more likely to exhibit genuine behavior (Bombari et al., 2015). The usage of confederates has been described as an attempt to create “illusions” for participants, thereby improving data quality through the creation of “mundane realism” (Blascovich et al., 2002). Confederate-based assessment is also characterized by low internal validity because confederates adjust their behaviors to their interaction partners (Kuhlen & Brennan, 2013).

Using confederates in the social sciences has been the topic of various debates because researchers consider it unethical to employ confederates without informing the participants beforehand. Ethics guidelines for psychologists state that deception of any kind may only be used under circumstances where there is no other way to complete a study (American Psychological Association, 2017).

Vignettes. Vignettes also go beyond introspective questionnaires. Vignettes are descriptions (written or verbal) of specific situations. In leadership assessments, these vignettes include prototypical leadership situations that result from discussions of critical situations that leaders have experienced (Peus et al., 2013). The participants are asked to imagine the situations vividly and then to answer questions about them.

Vignette-based assessments have a distinct advantage over questionnaires and introspective reports, mainly due to their ability to address specific contexts. By describing a situation in detail and requiring participants to think of it when answering questions, vignettes can standardize the reference point of the subsequent answers (Aguinis & Bradley, 2014). Vignettes can therefore facilitate experimental control by specifying concrete situations for experimental conditions (Blascovich et al., 2002). Although vignettes are intended to assess a leader's actions in concrete situations, they can only capture behavioral intention, not actual behavior (Eifler & Petzold, 2019).

Accordingly, vignette-based measurements are characterized by high internal validity. Once developed, the specific scenarios can be shown to participants in the exact form the researchers originally envisioned. However, vignette-based measurements have low external validity (Eifler & Petzold, 2019) because they only contain descriptions of hypothetical situations. Vignettes thus depend on the participant's imagination, which can vary considerably even when the same situations are described. To increase the external validity of vignette-based methods, researchers have suggested restricting their use to situations that participants are likely to have experienced firsthand, which are easier to imagine vividly.

Researchers have reverted to using picture or video vignettes to increase the external validity of vignette-based approaches (Eifler & Petzold, 2019). One approach motivated by the desire to circumvent the lack of context specificity is to use video vignettes that display leadership situations. The remainder of the procedure resembles that employed in studies with non-video vignettes, meaning that participants answer questions in various forms based on the scenario they have seen. Video-vignette-based methods have been identified as a promising method for testing the effects of manipulations on variables because they do not require the participants to exercise their imagination, which standardizes the manipulation and emphasizes its effects (Podsakoff et al., 2013). Importantly, video-vignette-based approaches have also been used to measure the impact of certain styles of leadership, such as charismatic leadership, on participants (Antonakis et al., 2011; Meslec et al., 2020). Although video vignettes have many advantages over regular vignettes, the main one being their high level of standardization, their external validity is still limited because the displayed situations remain hypothetical, and the participants are mere observers (Bombari et al., 2015).

In summary, leadership assessments based on introspective reports have been the primary tool that leadership scholars have used over the recent decades, and they have enabled insightful leadership assessments in the past. These types of assessments have pitfalls. Inaccuracies of follower ratings, lack of context-specificity, and measuring perceptions of leadership instead of actual leader behavior. Attempts to solve these problems, such as using confederates or (video) vignettes, have shown promise but struggle with internal and external validity as outlined.

Situational Judgment Tests

Situational judgment tests (SJTs) take vignette-based assessment one step further. They focus on situations that describe the context that participants should keep in mind when answering questions (Whetzel & McDaniel, 2009). However, compared to vignettes, SJTs do not measure outcome variables using typical questionnaires but instead include behavioral options from which participants can choose (McDaniel & Nguyen, 2001). Behavioral options can then be used to approximate the participants' hypothetical behavior (Peus et al., 2013; Salter & Highhouse, 2009).

One SJT that has been found to assess the full range of leadership model accurately is the Leadership Style Assessment (LSA). The LSA was developed to measure leader behavior in a wide range of critical situations (Peus et al., 2013). Followers are asked to choose behavioral options based on prior experiences with their leader. Based on the selected behavioral reactions, a leadership style from the full range of leadership model is assigned to the leader. Each available behavioral option can then be attributed to the dimensions of various leadership styles, meaning that the score resulting from the participant's choices determines the leader's dominant leadership style. The LSA is an attempt to effectively address the aforementioned limitations of introspection-based questionnaire assessments.

Improving the accuracy of follower ratings. It has been suggested that measurement approaches focusing on critical incidents, incidents that have evoked emotions in the past, can increase measurement precision (Hansbrough et al., 2015). Since concrete situations serve as reference points for ratings, followers can remind themselves how their leader acted and thus provide more accurate answers.

Enhancing context-specificity. In describing situations precisely, the LSA focuses on a detailed representation of the context in which leadership is to be assessed. The situations are described in a way that provides sufficient contextual information to determine how the leader would act in similar situations. Crucially, the availability of multiple scenarios in the LSA allows for a thorough examination of leadership style.

Measuring concrete behavior. Although participants are not asked to act out their behavior directly in the LSA, their choices from several behavioral options allow for a precise evaluation of the participants' behavioral intentions. The LSA measures behavioral intentions across important dimensions since it includes behavioral options based on the full range of leadership model (Bass & Avolio, 1994).

While SJT-based approaches address the limitations of introspective questionnaires to some degree, they do not eliminate them. Since participants are forced to rely on their imagination instead of acting, SJTs have similar limitations to vignettes and make cognitive and emotional aspects less salient due to lower engagement with the situation. While SJTs can also use picture and video vignettes as their foundations, participants can still only "experience" the situations that are presented to them as observers (Aguinis & Bradley, 2014). Alvesson (2020) recommended studying leadership processes and practices directly and researching leadership as it occurs. For leadership style research, he advocated emphasizing the individual components of leadership, such as behavior, instead of attempting to measure multiple aspects of leadership at once. Researchers have also suggested using technologies instead of videos because they can provide a fully immersive experience to participants and, therefore, a framework in which behavior can be observed directly (Pierce & Aguinis, 1997). One such technology is VR.

Virtual Reality

VR has been investigated for decades, but the term's meaning has evolved, primarily because of the technology's changes (Blascovich, 2002). While many researchers use the term VR for any virtual representation, including on a two-dimensional screen, we focus on what has frequently been called "immersive virtual reality," that is, on technologies that allow for a three-dimensional representation of the environment and transport their users into that space as direct actors (Schmid Mast et al., 2018).

To clarify the differences between various classifications of VR, this paper provides an overview of the types of VR that are commonly used. The first important distinction that has to do with how the VR system accomplishes the user's transportation to the virtual environment is the level of immersion that the system creates. Immersion focuses on the technical aspect of a virtual experience (Mantovani & Castelnuovo, 2003), characterized by aspects such as graphical fidelity, screen space, and sounds. It lays the foundation for users to be surrounded by the virtual space. Researchers classify VR as low-immersive, semi-immersive, or fully immersive (Martirosov et al., 2021).

In low-immersive VR, the virtual experience is delivered through a regular screen, e.g., on a laptop or desktop PC. Since they can see the physical distance between their eyes and the virtual environment, users of low-immersive VR are aware that the application they are experiencing is external to them and, thus, that they are not present inside the virtual environment. Low-immersive VR is useful in experiences that would otherwise lead to motion sickness (Martirosov et al., 2021), and it costs substantially less than semi-immersive VR systems such as the Cave Automatic Virtual Environment (CAVE).

Semi-immersive VR only renders a part of the environment virtually, while the other part remains outside the virtual space (Kyriakou et al., 2017). Semi-immersive VR is most used in the educational sector, for example, in various assembly or space-flight simulators. The users are placed in a real environment, while screens show a virtually rendered representation of the educational content (Abidi et al., 2018; Museth et al., 2001). One advanced, and commonly employed semi-immersive VR system is the CAVE (Martirosov et al., 2021). CAVEs attempt to immerse the users in virtual environments by surrounding them with screens. Users can walk around freely while simultaneously engaging with the virtual content on all screens. As the user changes their position, the system calculates stereoscopic images in real-time to create believable and deep virtual representations of an environment (Mullins, 2006). Whereas a CAVE can be built so that the user is surrounded by screens, transforming it into a highly engaging and fully immersive experience, the cost of that technology is high. A specifically designed room with multiple high-fidelity screens is needed, as is a sound system of sufficient quality.

Fully immersive VR takes the virtual space one step further and attempts to immerse the user fully in the virtual environment so that no references to the real environment can be found (Martirosov et al., 2021; Schmid Mast et al., 2018). Fully immersive VR environments are frequently displayed via Head-Mounted Displays (HMD) and can be divided into two types. PC HMDs, such as the HTC Vive Pro 2, must be tethered to a high-performance computer to run, allowing for better performance and higher-fidelity visuals. Standalone HMDs, such as the Oculus Quest 2, do not require any external devices, but they are less powerful and therefore limited in their performance and graphical output.

There has been a substantial resurgence of interest in VR due to commercial HMDs becoming widely available to consumers. HMDs project virtual images on two screens positioned directly in front of the user’s eyes. At the same time, all other light sources are blocked, resulting in full immersion (Martirosov et al., 2021). Not only does the HMD experience have an immersive quality as a result of feeling placed in the center of a virtual environment, but technological advancement has facilitated a sharp decrease in costs. Nowadays, high-quality commercial HMDs can be purchased for between $300 and $1,500. In 1997, fully immersive VR systems required an investment of between $50,000 and $80,000. In 1992, the cost could reach $1,000,000 (Pierce & Aguinis, 1997).

As a fully immersive technology, VR places participants in concrete situations where they are direct actors. The gap between behavioral intention and actual behavior is bridged, and cognitive and affective engagement is higher. Accordingly, external validity increases.

Virtual Reality for Leadership Assessment

While research on leadership in VR is scarce, many scholars have investigated components of leadership or leadership-adjacent behavior in VR. As far as research on leadership-adjacent behavior is concerned, VR interventions have focused on public-speaking training. VR public-speaking research originates from research on VR-supported therapies for social anxiety, in which appropriately designed environments have consistently recreated feelings of anxiety in users (Pertaub et al., 2002). Complex mechanisms, such as a virtual audience reacting to a presenter and giving real-time feedback, have been implemented in various virtual public-speaking courses (Batrinca et al., 2013; Palmas et al., 2021; Poeschl & Doering, 2012). Since the ability to present oneself and one’s vision confidently has been found to be essential for leaders (Antonakis, 2012; Judge et al., 2006), public speaking-related VR applications are useful in the leadership context.

While most research on leadership and VR has so far focused on military leadership and only considered non-immersive VR (Gordon et al., 2004; Raya et al., 2018), recent papers have also investigated leadership based on organizational research and non-immersive VR. For instance, Parra et al. (2021) developed a serious game for leadership assessment that investigates task-oriented and relationship-oriented leadership on a sailboat. Eye-tracking and machine learning were employed to determine whether participants adopt task-oriented and relationship-oriented approaches to leadership. Research on fully immersive VR and leadership is currently limited.

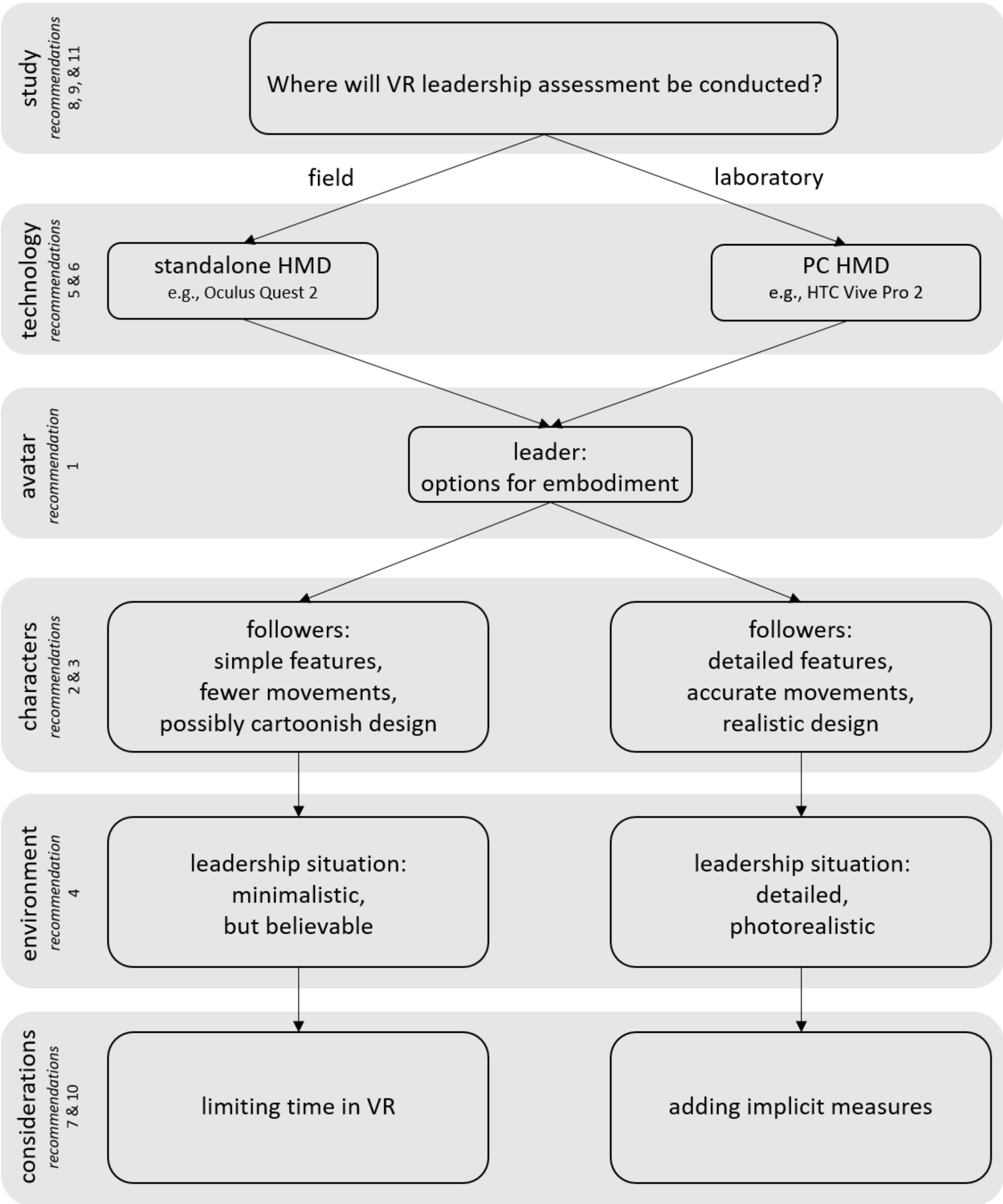
Immersion involves technical aspects that affect the degree to which the virtual environment will surround the user. The user's subjective perceptions ultimately determine an experience. In contrast to the technical determinants of immersion, presence represents the feelings that users experience when they find the virtual environment believable (Sanchez-Vives & Slater, 2005). Immersion has been identified as a precursor to presence because only an experience that is technically capable of creating a realistic environment can make users feel that the situation in which they find themselves is convincing (Bombari et al., 2015; Slater, 2003).

Presence is connected to how the user, their social interaction partners, and the environment are perceived (Bulu, 2012). Human-Computer Interaction (HCI) researchers are increasingly using fully immersive VR to investigate how social interactions unfold in virtual spaces and which aspects of design need to be considered in relation to avatars, characters, and environments (Bombari et al., 2015; Schmid Mast et al., 2018). In the leadership context, the VR social interaction partners are followers, and the VR environment represents a situation that reflects researchers' calls for leadership assessments that cover not only leaders but also followers and situations (Alvesson, 2020; Hansbrough et al., 2015).

VR applications require thorough implementation and consideration of all matters that affect user experience. The three aspects that primarily determine a user's experience in VR, particularly in leadership scenarios (i.e., virtually implemented complex social interactions), are embodiment, social presence, and telepresence. Avatar, characters, and environments must be implemented based on this triad. An overview of the considerations that must be considered before a VR leadership assessment application is implemented that reflect the context of the study is displayed in Figure 7 and elaborated in the following sections.

Figure 7

Overview of Considerations and Recommendations for VR Leadership Assessment



Avatar Implementation – Leaders in VR

Embodiment is defined as the feeling that users have ownership over their virtual body in virtual spaces (Kilteni et al., 2012) and is of particular importance to VR HMDs. Since screens are placed right before their eyes, and users cannot see their bodies, a believable virtual representation of their physical entity in VR is crucial (Slater, 2017). Much research has been conducted on embodiment and VR's most optimal body representations. One unresolved question is whether it is enough for users to see a virtual representation of their hands or if a representation of the entire body is required (Matamala-Gomez et al., 2019; Steed et al., 2016). Moreover, the haptic method of interacting with the virtual world can affect the degree to which the users feel embodied in their virtual representation (Spanlang et al., 2014). The profound effects of embodiment are evident from research on implicit bias. VR interventions decrease it effectively by placing participants in different virtual bodies (Hasler et al., 2017; Peck et al., 2013).

To facilitate virtual environments that invite leaders to make behavioral choices that mirror their reactions to similar real-world situations, leadership researchers must determine how leaders can be represented in virtual space. Whereas many VR applications do not include representations of the user's body at all (e.g., passive experiences that do not require complex interactions), the VR experiences that are most suitable for leadership assessment, that is, serious implementations with a focus on realism, must focus on those representations.

Once a whole body is implemented in VR, new questions arise about the type of body that should be presented. For instance, gender mismatch affects embodiment (Schwind et al., 2017). In the context of behavior, it is reasonable to expect that a sense of body ownership in VR would substantially affect the leader's ability to act in a manner that is congruent with their desired actual behavior. In fact, leadership scholars have drawn a link between leaders' enactment of behaviors and their subsequent perceived authenticity (Weischer et al., 2013).

Since preliminary research on embodiment suggests that including a virtual body leads to improved presence and engagement (Steed et al., 2016), participants can be given a choice between various body types, genders, and ethnicities. The choice would enable participants to select the avatar they feel most comfortable with, thus increasing the likelihood of feeling embodied and present in VR.

Recommendation 1: Participants in VR leadership assessments should be able to choose between diverse virtual bodies. The options should include different body types, genders, and ethnicities.

Character Implementation—Followers in VR

Whereas embodiment focuses on how the user is represented in VR, social presence is specifically related to the representation of virtual characters in a virtual environment and determines how believable social interactions with those characters are perceived to be (Bulu, 2012). Social presence is irreplaceable in VR applications that contain interactions with virtual agents. It captures the degree to which users feel a social interaction partner is "available" (Lee, 2004). Since any social interaction in the real world can be experienced as engaging, depending on the context, it is clear why social presence is crucial for engagement in VR (Oh et al., 2018).

As a result, the need for standardized interaction partners has been emphasized in VR research. Behavior in social situations is always a result of an interaction between two or more individuals (Bombari et al., 2015; Schmid Mast et al., 2018). Since leadership situations are defined by dyadic and reciprocal interactions between leader and follower (Yukl, 1989), social presence is essential for the realism of the representation of leadership in VR, and it can lead to higher engagement, deeper immersion, and a higher likelihood of the benefits of VR being realized. After all, as the perceived familiarity of virtual human characters increases, so does the possibility that users will exhibit natural behaviors in the virtual space

(Slater & Steed, 2002). These increases are crucial prerequisites for precisely assessing leader behavior in VR.

To maximize social presence, the acceptance of the virtual character or, in the case of VR leadership assessment, the virtual follower needs to be taken into account. The first aspect of their virtual counterparts that individuals experience is their appearance. In the context of realistic leadership situations in VR, one would expect that virtual followers will be designed as realistically as possible to increase familiarity. However, research has shown that this realism has diminishing effects due to a phenomenon that is called the Uncanny Valley.

Familiarity with virtual characters only increases linearly with human likeness up to a point. Thereafter, it decreases quickly, causing humans to view the characters as eerie (Mori et al., 2012). In other words, if a virtual character is too humanlike, it might not be perceived favorably or accepted. Researchers have given many explanations for the Uncanny Valley. The most prominent is the theory of category uncertainty, which asserts that if a character is intended to be fully humanlike, any element that deviates from this standard and thus from the user’s expectations detracts from the experience by making it more challenging to determine what level of human likeness the character is supposed to represent (MacDorman & Chattopadhyay, 2016, 2017). However, scholars have also found evidence contradicting this Uncanny Valley explanation. While the theory of category uncertainty is premised on the assumption that humans classify entities into distinct categories, the theory of perceptual mismatch posits that human likeness exists on a continuum. Singular features that appear artificial on otherwise humanlike entities can cause perceptions of eeriness (Kätsyri et al., 2015). Furthermore, disparities between a character’s “perfect” appearance and their unnatural movements can also elicit feelings of eeriness (Bombari et al., 2015).

The Uncanny Valley becomes particularly important whenever the design of virtual characters needs to be determined before they are placed in their intended context. While leadership assessment is considered a serious task, appearance problems make it necessary to design virtual characters in cartoonish styles. In this way, the complexity of their features is reduced, and issues of category uncertainty and eeriness are avoided (Kätsyri et al., 2017). This said, Uncanny Valley research is primarily based on work on two-dimensional environments and robots. It is reasonable to assume that important aspects of the Uncanny Valley can be found in VR. Due to VR's immersive and engaging nature, the triggers related to the Uncanny Valley are even more problematic in that context. In leadership situations, it is necessary to design characters before implementation because eerie characters would substantially affect the ability of leaders to feel socially present vis-à-vis their virtual followers, thereby diminishing the potential of VR to recreate scenarios for leadership assessment.

These findings have major implications for designing VR leadership assessment applications—characters must be designed to avoid the Uncanny Valley effect. Unless researchers have access to high-quality characters with high-quality animations and use a PC HMD like the HTC Vive Pro 2 to render all details fully, focusing on fewer features would make the Uncanny Valley less critical. Reducing the number of features means that category uncertainty and perceptual mismatch are less likely and that the impact of potentially unnatural movements is diminished.

Recommendation 2: For VR leadership assessments that are conducted with standalone or low-performance HMDs, the features of followers need to be simplified to prevent the occurrence of the Uncanny Valley effect and maximize social presence. For VR leadership assessments conducted on PC HMDs, highly detailed follower characters can be considered when high-quality movement animations are available.

The theory of the Uncanny Valley has been extended beyond the visual appearance of characters and to the mind. That version of the theory posits that factors that pertain to the experience as a whole and to the character's behavior can also elicit feelings of eeriness. The most prominent findings pertain to an agent's capacity to experience or feel (Gray & Wegner, 2012). When humans encounter virtual entities, they commonly expect them to act independently due to pre-programmed behaviors or artificial intelligence. However, suppose the agent acts in a way that creates the perception that they can feel emotions. In that case, default expectations are violated, which, in turn, triggers the Uncanny Valley of the Mind effect (Appel et al., 2020).

The design of an agent and its capacity to feel can be perceived directly through observation. A second perception can cause the Uncanny Valley of the Mind effect, namely mind attribution. Users expect agents to be pre-programmed or use artificial intelligence to choose between actions. However, if the agent's actions seem overly intelligent or unexpectedly complex, humans begin attributing a mind to it (Gray et al., 2007; Stein & Ohler, 2017). Mind attribution has been linked to similar effects as the Uncanny Valley of the mind because it is another violation of the users' expectations and therefore influences experiences negatively by creating distance between user and agent.

It is vital that virtual followers in VR leadership assessment applications react to the leader's actions realistically and logically. However, they should not appear emotional, intelligent, or as if they have minds. Researchers must ensure that the characters' reaction is congruent with the leader's expectation about the reactions of a human follower in a similar situation. This way, perceptions of characters deciding independently based on individual factors would not arise.

Recommendation 3: The reactions of virtual followers must be realistic and logical consequences of the leader's behavior. However, those followers should not be overly *emotional or intelligent because attributing feelings and minds to the followers will lead to the Uncanny Valley of the Mind effect.*

Since the LSA revolves around creating standardized leadership situations and VR enables researchers to create any situation in a highly immersive environment, a promising avenue for future research is the implementation of SJTs, such as the LSA, in VR. In addition, VR could be used to reverse the situation by having individuals play the role of followers interacting with virtual leaders. By placing the follower at the center of a situation and having a virtual leader exhibit a concrete behavior, researchers could investigate how leader behaviors affect followers in virtual environments. Moreover, leaders placed in follower roles would be able to experience the effects of leaders' actions and, thus, leverage the ability to see specific behaviors from the follower's perspective.

It is important to note that, in considering the VR-based implementation of SJTs such as the LSA, one must think of the interaction partners in VR being pre-programmed (agents) rather than represented by real users (avatars). Although simultaneously putting human leaders and followers into virtual environments may be interesting, the standardized situations would not align with the researchers' intentions, which is crucial for standardized leadership assessment.

Recommendation 4: For VR leadership assessments based on situational judgment tests, virtual followers' behavior must be pre-programmed to ensure high internal and external validity.

Environment Implementation—Leadership Context in VR

Telepresence is determined by the implementation of the VR environment around the user. The user should feel as if they are inside the virtual environment (Bulu, 2012; Draper et al., 1998). Telepresence describes the feeling of being physically present in an experience (Biocca et al., 2003), which, in turn, increases overall presence so that perceptions of space and time begin to blur and the user becomes fully engaged with the virtual world (Mantovani & Castelnuovo, 2003).

A major challenge that needs to be overcome for the potential of VR implementations to be leveraged is the unification of the virtual and the real. Although the VR experience can be highly immersive, there are no means of making the transition between reality and VR seamless— users will always know that they have put the HMD on at some point. Consequently, VR implementations must be coherent with how one would expect to experience the same or a similar situation in the real world. This proposition applies not only to the representation of the virtual environment but also to the behaviors that may result from interacting with the environment in VR.

Any deviations from expectations about objects or possible behaviors in the environment affect the user's experience negatively by creating a break from their expected and actual experience, creating a distracting effect on their presence. Furthermore, the environment around the user and the agent can directly influence perceptions of eeriness. If the virtual environment is designed in a way that violates the user's expectations, the Uncanny Valley of the Mind effect can occur (Howard, 2017).

Since leadership situations primarily unfold in office environments, the VR implementation must be designed so that the virtual environment is coherent with the user's expectations. Photorealism would be the first choice for serious applications. Still, the Uncanny Valley of the Mind can be problematic when photorealism and an exact representation of the environment cannot be achieved. Simpler and cartoonish implementations can be effective as well. If, for instance, a VR application runs on a comparatively cheap standalone HMD, aiming for photorealism would lead to graphical artifacts that not only generate breaks in immersion and perceptions of eeriness but also result in motion sickness when the HMD can no longer render enough frames per second to maintain the VR illusion for the user. It is, therefore, vital to determine, before implementation, what type of research the VR implementation is intended for and what the relevant portability and performance concerns are.

Recommendation 5: For VR leadership assessments that are conducted on standalone or low-performance HMDs, the virtual environment should be rendered in simple graphics and with few features to implementations that represent details inconsistently and to improve performance. VR leadership assessments conducted on PC HMDs can be more detailed and rely on superior graphics.

Considerations for Empirical Leadership Assessments

When combined, embodiment (how the user is represented), social presence (how believable social interactions are), and telepresence (how the environment is implemented) determine the presence that a user experiences in a VR application. Hence, to create realistic leadership situations for assessment in VR, the embodiment is determined by the leader's implementation, social presence by the followers' implementation, and telepresence by the leadership situation's representation.

In complex social situations such as leadership interactions, the feeling of being in a virtual world leads to performing actions that one would also perform in the real world. It has been shown that when virtual humans mirror the behavior of actual humans, users begin to experience the interaction similarly to how they experience analogous situations in the real world, for instance, in terms of social inhibition (Buck et al., 1992) and arousal (Slater, Guger, et al., 2006). Interestingly, research on virtual situations in medicine indicates that perceptions of authority, a relevant aspect of many leadership interactions, and the corresponding behavior adjustment unfold similarly in the virtual environment and the real world (Mast et al., 2008).

It has been suggested that critical incidents can be used for measurement to activate episodic memory more effectively and obtain more accurate follower ratings (Hansbrough et al., 2015). VR is an opportunity because followers do not need to remember critical incidents. Instead, leaders can directly act out critical leadership incidents in VR, and their behavior can be observed directly. Similarly, VR can be used effectively to increase the accuracy of follower ratings. Since episodic memory has been shown to be more active after exposure to emotional content (Allen et al., 2008) and since visual reminders of emotional events have been linked to higher rating accuracy (Naidoo et al., 2010), researchers can recreate emblematic leadership situations in VR to activate the followers' episodic memory and thus improve rating accuracy when they subsequently rate a leader. The implementation of specific leadership situations from the LSA in VR, thus, leverages all of the advantages of SJTs and creates a framework in which cognitive and affective engagement increase (Bulu, 2012; Slater, Guger, et al., 2006) and in which behavioral choices are more likely to represent actual behavior.

One crucial factor that needs to be considered when VR is used in leadership studies is the type of study being conducted. The key distinction is between experimental and non-experimental setups. The most important benefit of VR, which was noted previously, is that it can create standardized leadership situations that purport to replicate all significant contextual factors, which, in turn, allows concrete behavior to be observed in the virtual environment. Consequently, VR can be used in any study, whether experimental or non-experimental. The benefits of VR are more pronounced in experimental studies because various conditions can be set up in a standardized manner and because controls for contextual factors can be introduced. As a result, the experiments that are run are of high internal and external validity (Blascovich et al., 2002; Pierce & Aguinis, 1997), a standard that other experimental methods, such as the LSA, cannot attain to the same degree.

Recommendation 6: VR leadership assessments are particularly suitable for experimental leadership studies because multiple standardized conditions can be implemented and used to control for contextual factors. VR leadership assessments also have advantages over experimental studies based on situational judgment tests because of their higher internal and external validity.

Once a type of study has been selected, it is vital to consider the circumstances in which data is collected and, thus, the type of HMD that should be used. While some researchers have argued for field studies and criticized laboratory experiments for their lack of generalizability (Levitt & List, 2007), the recent literature contradicts these claims. It has been argued that laboratory studies are generalizable (Charness & Fehr, 2015) and enable more control over experimental variables. In addition, they are not affected by self-selection bias (Falk et al., 2013). Whatever a researcher's reasons for conducting a field study or a laboratory experiment, VR can be used in any context due to the availability of standalone and PC HMDs.

If researchers set up a complete laboratory experiment under controlled conditions, they can profit from the high performance of PC HMDs because portability is not a priority. As a result, graphical fidelity and performance can be emphasized. If a field study is being conducted at an organization, standalone HMDs are useful due to their portability, but concessions on visuals and performance are unavoidable. However, as explained previously, the photorealistic implementation does not necessarily lead to superior outcomes due to the operation of the Uncanny Valley effect, which is particularly problematic when high-fidelity characters are used and when breaks in immersion occur due to insufficiently detailed implementations.

Prolonged VR experiences can have a negative psychological impact on users. Since HMDs simulate fully three-dimensional worlds with just two small screens that are placed in front of the user's eyes, the technical setup and its implementation can cause motion sickness, nausea, and disorientation (Martirosov et al., 2021; Saredakis et al., 2020). The factors that determine whether an individual will experience such side effects from exposure to VR are still being investigated, but using state-of-the-art VR technology seems to reduce the negative influences of VR (Saredakis et al., 2020). Side effects can become more or less prevalent when certain situations are displayed. For example, a small number of slower movements reduces the likelihood of nausea and disorientation; the same is not true of a rollercoaster ride simulation. The technical aspects of the HMD can also have positive effects, including higher picture and movement clarity (Ray et al., 2018), higher frame and refresh rates (Kourtesis et al., 2019), and lower field of view (Fernandes & Feiner, 2016), which have all been linked to improved outcomes for VR users (Saredakis et al., 2020). These features are primarily available with PC HMDs.

Recommendation 7: If data is collected in the field, the portability of standalone HMDs can be leveraged. The side effects of VR, such as disorientation and nausea, mean that VR immersion needs to be relatively short. PC HMDs enhance laboratory setups through superior performance and graphical fidelity and have fewer side effects, but they are less portable.

Researchers who are interested in using VR for leadership assessment realize that although initial development costs and time requirements for VR implementations can be high, future-oriented development can create frameworks that allow for the modular expansion and adjustment of the implementation and can therefore be used for various research projects over a long time.

VR development is usually conducted via game engines, tools commonly used for video game development. Game engines deliver environments in which many fundamental aspects of two- and three-dimensional experiences, such as object interactions, game physics, and artificial intelligence, are implemented in their initial states, enabling developers to build on them easily. It is essential to consider the implementation process and whether the VR experience needs to focus on performance or portability. As explained previously, in the context of different VR HMDs, it is not only the device but also the engine and the experience of the developers that shape the experience. For instance, an engine such as Unity is simple to use and effective for mobile experiences, but it generally cannot attain the graphical fidelity of the Unreal engine. However, the latter requires more computational power and, thus, more advanced VR HMDs. Researchers should use commercially available game engines that have long-term support.

If all the benefits of VR are to be leveraged, the representation of a leader's actions in VR needs to be considered carefully. Freedom of action has implications for measurement. For instance, one situation in the LSA requires the leader to navigate a scenario in which an employee has prepared a subpar presentation for an important client. In VR, we can imagine that this situation and its context are implemented directly. For instance, the follower may walk into the office and apologize for having handed in such a poor presentation. The interaction between the leader and the follower would then unfold through direct speech and speech recognition. The same situation could also be presented as an email on a virtual computer in VR space. The leader would then read the email before the virtual follower walked into the office. If the interaction and the choice of behavior occur through an interface element whereby the user selects what they want to say, situations could be quickly combined with new behavioral options.

Both approaches have benefits. The first is more immersive because context is presented directly through an interaction; the second requires fewer modifications of the base material and represents the actual situation more directly. While interactions with virtual followers can also be conducted through speech and speech recognition, which increase immersion, adding more situations would require additional development resources.

Recommendation 8: Commercial and widely available game engines enhance implementation. Complex interactions through speech create a more immersive experience, so researchers should include these natural interaction options whenever possible. However, text-based descriptions of situations and behavioral decisions made through interface elements allow the application to be extended more quickly in the future.

VR has been used for decades in numerous contexts, and it is important to consider the ethical implications of the technology. Due to its nature as an immersive technology and as a subjectively real experience, researchers must consider the proposition that the presented content can affect participants substantially. While participants can be made aware that they are in a safe space and can stop the experience at any time, the effects it induces can still be profound. In psychology research, experiments that would be unethical by contemporary standards, such as the infamous Milgram experiment on direct physical aggression, have now been implemented in VR (Slater, Antley, et al., 2006) because the ethical restrictions are thought to be less stringent—the experience is presumed not to be real. Research on phobias (Garcia-Palacios et al., 2007; Safir et al., 2012) indicates that VR induces realistic fear responses. Given that the virtual world can be perceived so intensely that it triggers phobias, one needs to be cognizant of the fact that researchers have a responsibility to be careful with the content that they generate for participants.

The VR researchers who specialize in HCI realized early that the technology can be effective for investigating sensitive situations that would be difficult to study in the field. Scholars of organizational behavior, for instance, have suggested using VR for research into workplace romance and sexual harassment (Pierce & Aguinis, 1997). This suggestion is congruent with the clinical research in which VR has been used to simulate situations that would be impossible or unethical to recreate in the real world (Bombari et al., 2015). Consequently, VR leadership assessment can be employed to investigate sensitive leadership situations if this research adheres to ethical guidelines and as long as VR is not treated as a blank check. Participants must be informed about the content they will encounter and be allowed to abort the VR experience at all times.

Furthermore, VR provides a unique opportunity to vary the individual features of virtual followers, which is promising for leadership research because humans react differently to interaction partners with different appearances. For instance, competence is assessed differently depending on facial features (Todorov et al., 2005), and willingness to negotiate is judged more negatively in female candidates than in male candidates (Bowles et al., 2007). Given the different perceptions of various characteristics, researchers must frequently use multiple interaction partners to counteract the impact of specific traits on variables (Azmat & Petrongolo, 2014). In VR, it would be substantially more accessible for researchers to select from multiple characters with various facial features or genders. Consequently, VR is a promising means of investigating perceptions of individual traits and the way biases influence them in the leadership context.

Recommendation 9: VR leadership assessments can be used to investigate leader behavior safely in sensitive contexts that cannot be observed quickly otherwise. However, VR is not a blank check—ethical guidelines must be followed. Participants must be aware of the content presented to them, and the option to stop the VR experience must always be available because immersive VR experiences can induce similar reactions to experiences of sensitive situations in reality.

The evaluation of interactions through the assessment of markers relevant to automatic processing in real-time has been found to be necessary for the perception of leadership (Braun et al., 2018; Lord & Maher, 2002). This proposition is particularly true of VR because researchers have found that cognitive and affective responses can be measured due to VR experiences (Slater, Guger, et al., 2006). Methods that are used primarily in neuroscience or the adjacent fields, such as eye-tracking, Electroencephalography (EEG), or Galvanic Skin Response (GSR), are also increasingly being investigated in leadership research (Parra et al., 2021; Raya et al., 2018). It is difficult to use these methods effectively because they often require experimental setups to be created in laboratories, which limits researchers to controlled environments and excludes the possibility of field research, in which these methods would serve as promising new data sources. By creating VR leadership situations that can also be experienced in laboratories, researchers could combine situations that can otherwise only be observed in neuroscientific and physiological field experiments. The use of neuroscientific and physiological methods in VR is, therefore, capable of producing and enhancing the quality of behavioral data that is otherwise difficult to obtain.

EEG can measure a leader's cognitive load during a VR situation (Zhang et al., 2017), yielding insights into the determinations that need to be made during various critical leadership situations. GSR can generate high-quality data on stress responses common in critical leadership situations (Kurniawan et al., 2013; Perala & Sterling, 2007). Eye-tracking can be combined with either of those methods to produce additional insights into cognitive load and stress (Simonovic et al., 2018; Zagermann et al., 2016). Notably, various HMDs, such as the HTC Vive Pro Eye, have been developed specifically to include the state-of-the-art eye-tracking technology by the company Tobii, resulting in the easy integration of the VR stimulus and the eye-tracking measurement.

Recommendation 10: VR leadership assessments can be combined with measures such as eye-tracking, EEG, and GSR to assess the factors that underlie leader behavior, such as cognitive load and stress, in real-time. Such approaches allow neurophysiological *data to be collected in leadership situations where leaders exhibit their natural behaviors, which is much more challenging to achieve without VR.*

The implications of VR leadership assessment studies must be treated with caution. Crucially, in the past, much research focused on leader selection and leadership development because they are the most pertinent dimensions of leadership in practice (Day, 2011; Vardiman et al., 2006). Leadership assessment is the most promising focus for initial examinations of VR in leadership research. If leadership assessment is treated as a core concern, leadership in VR can be investigated in a way that allows one to understand how behavior and social interactions translate to VR in the context of leadership.

Leadership assessment is the middle ground that allows researchers to examine all important details of leadership interactions in virtual spaces because assessments serve as the foundation for determining how specific behaviors can be evaluated in VR. Once research on leadership in VR has been conducted and leadership can be measured effectively in virtual spaces, the next logical step would be to consider how leadership development would operate in VR. It is already known that complex social abilities such as public speaking and social interactions are practiced in VR. In addition, it is only if scholars are confident that leadership assessment in VR is effective that they can begin making claims about selecting leaders based on behaviors and abilities observed in these virtual spaces. It is also likely that VR-based leader selection will be used by organizations in selection processes, which might lead to issues that are similar to those observed in AI-conducted or AI-analyzed interviews, in which a lack of trust in the current technology leads to perceptions of unfairness (Suakanto et al., 2021).

Recommendation 11: Leadership assessment is a suitable starting point for leadership research on fully immersive VR. Once VR accurately represents and measures natural *leadership behaviors, leader selection, and leadership development can be addressed in VR, which will be particularly relevant to practitioners.*

Conclusion

We explained why VR is a promising technology for leadership assessment. Immersing participants into critical leadership situations in VR enables the leaders' behavior toward virtual followers to be observed and categorized based on accepted leadership assessment standards. The behavior of leaders in VR situations is likely to be representative of their behavior in the real world. By paying attention to the implementation of avatars, characters, and environments (see the 11 recommendations above), researchers can create situations in virtual spaces that invite leaders to behave naturally. VR can be used in field and laboratory experiments to investigate sensitive leadership situations, and it can be combined with neurophysiological measures of the cognitive and affective mechanisms that underlie leader behavior. VR leadership assessment is an appropriate starting point for investigations of the efficacy of VR in leadership research, which creates the foundation for further research into leader selection and leadership development. In summary, measuring leader behavior in VR is a promising extension of SJTs such as the LSA. That extension can be used by researchers and practitioners to assess leadership effectively and to create a novel method that goes beyond the ritualized use of questionnaires.

5. Overall Conclusion

Leadership in the digital age is a critical challenge for organizations because failure to adapt can lead to stagnation in the face of digital transformation (Menon & Suresh, 2020; Vial, 2019). Leaders are crucial in navigating organizations through changes and transformation processes (AlNuaimi et al., 2022; Bartsch et al., 2021; Kunisch et al., 2022). The ability of leaders to face the challenges of digital transformation is enhanced by technologies that can complete specific assignments and thus free up leadership resources for complex tasks.

Research on robot leadership has been primarily conceptual (Gladden, 2014; Samani et al., 2011; Samani et al., 2012). In Chapter 2, I examined human followers' initial perceptions of and reactions to robot leaders. My results indicate that human followers' ideas about robot leaders primarily depend on their general assumptions about the technology (i.e., they are based on their technological affinity), on their reactions to robot leadership styles (i.e., transformational leadership has more promise than transactional leadership), on their emotional reactions (i.e., lack of trust due to perceived lack of empathy), and their visions of the future and ethical concerns (i.e., utopian or dystopian views about robot leadership). Therefore, the study emphasized the importance of human assumptions about robot leaders, indicating that a shared understanding of what robot leaders should and should not be able to do would be vital to accepting robot leadership. Moreover, I found that reflecting on the advantages and disadvantages of robot leadership frequently prompted human followers to re-evaluate their initially rigid views and led them to develop a more balanced and nuanced perspective. In addition, continuous deliberation and discussions of the utility of robot leaders could reduce the impact of intuitive and strong assumptions about robot leadership, thus allowing robot leaders to prove themselves as valuable tools for organizations.

In Chapter 3, I built on the findings from Chapter 2 and focused on the outcomes of different styles of robot leadership. Past research not investigated whether the influence of human leaders on human followers in organizational tasks would be mirrored in robot leadership. I investigated the effects of transformational, transactional, and minimal leadership on task engagement while considering the mediating effects of affect, likability, discomfort, perceived safety, and stress. My findings show a high degree of overlap between the outcomes of using human and robot leaders. Transformational robot leadership was shown to influence the dependent variables positively, which was to be expected in the context of the creative task that the human followers completed. These findings illuminate important links between the literature on HRI and leadership. HRI research indicates that the Uncanny Valley causes humans to experience eeriness when they engage with a transformational robot leader because social robots are generally not expected to be motivational or to occupy positions of authority. In my study, however, the Uncanny Valley did not pose a problem. The implication is that a robot leader might not violate human expectations if its appearance and behavior accord with the leadership context.

Furthermore, leadership assessment is crucial to preparing organizations for the challenges of digital transformation. VR has the potential to be an essential assessment tool because its immersive capabilities increase the likelihood of behaviors in virtual environments being representative of behaviors in the real world. SJTs and role-playing exercises have been presented as important methods for leadership assessment and development (McDaniel & Nguyen, 2001; Peus et al., 2013). Both can be implemented in VR to assess leadership in immersive virtual environments. For the implementation to be effective, crucial factors that pertain to the implementation of avatars, characters, and environments must be considered (Bombari et al., 2015; Schmid Mast et al., 2018; Slater, 2003). These factors and their relationship to the leadership context, leaders, followers, and situations were yet to be explored.

Accordingly, in Chapter 4, I focused on assessing leaders in VR. Immersive VR has shown promise in various skill-development applications (Palmas et al., 2019; Palmas et al., 2021; Pertaub et al., 2002), which is why I investigated specific implementation factors in the context of leadership assessment VR applications. In this chapter, I argued that the context of a study must cause researchers to draw crucial distinctions when they intend to use VR. Laboratory experiments allow for higher-fidelity representations with dedicated VR equipment that can be combined with other devices to generate more accurate measurements. Mobile VR solutions are more suitable for field research. However, the subsequent implementation must be considered carefully. Mobile VR's potentially limited computing power might mean that highly photorealistic characters and virtual worlds cannot be rendered with sufficient detail, thereby violating users' expectations and leading to the eeriness typically associated with the Uncanny Valley. I provided 11 concrete recommendations and a decision tree for implementing VR leadership assessments to maximize the potential of VR for leadership scholars. My dissertation illuminated the potential applications of novel technologies (i.e., social robots and VR) in leadership during digital transformation. This said, the work presented here only yielded initial insights into these complex domains. I discussed each paper's contributions and implications in the corresponding chapters. In the following pages, I will focus on the overall theoretical contributions of the dissertation and avenues for future research, practical implications, and ethical considerations.

Theoretical Contributions and Future Research Directions

First, my findings enhance the academic understanding of leadership in a technology-driven world. Since digital transformation entails opportunities and challenges, the role that leaders play at the intersection of organizations and technologies has become less defined.

When robot leaders can support leaders in assisting, inspiring, motivating, and guiding followers, human leaders can focus on solving complex problems that technology cannot yet tackle. In addition, developing applications that assess leaders through situational behavior in VR shows how immersive technology can be used to evaluate and potentially develop leaders in a repeatable and safe way. Future scholars should consider these findings to develop comprehensive models that embrace the unique benefits and challenges of technology.

Second, I focused on the means of implementing social robots and VR applications. I provided insights showing how human leaders' research can be translated to robot leaders and leadership in VR. Other researchers can use my results as a starting point for studies on theories of human leadership in the context of novel technologies. Future research should focus on building on these initial findings by investigating interactions between robot leaders and human followers to ascertain how reciprocity influences robot leadership and human-robot collaboration. As far as VR is concerned, applications for the assessment and development of leadership must be implemented to test its efficacy against measurable learning outcomes. The complexity of leadership is such that research will need to engage in a process of iterative exploration to represent it correctly and comprehensively in the virtual world.

Third, the findings of this dissertation shed light on the potential of technology for the study of leadership. Implementing robot leaders and VR leadership situations enables researchers to examine leadership development through a technological lens. It is unclear who would benefit the most from using such technologies. Prior research has shown that individual differences in personality, gender, and culture affect technology acceptance (Nistor et al., 2011; Park et al., 2019; Svendsen et al., 2013; Yi et al., 2005). Since all individual differences are relevant to leaders and followers, their impact must be investigated further in experimental research to identify the instances in which these technologies are most effective.

Practical Implications

My work also has several important implications for practitioners. First, by using robot leaders and VR leadership applications, organizations can accumulate experiences that would create opportunities for the targeted use of technologies in leadership. If, for instance, a VR leadership assessment tool is used to find a suitable (e.g., transformational or transactional) human leader, and that leader is supported by a robot leader that reminds followers of the goals and values of the organization by adopting a transformational approach, the human leader’s performance and job satisfaction could be improved because they would be able to focus on the most important tasks that require human intervention.

Second, as organizations use and develop technologies, their openness to innovation can increase. By incorporating technologies into their structures and strategies, organizations can quickly learn the advantages and disadvantages of technology for their use cases and how its employment can be improved. Such a process would also reveal the shortcomings of the applied technologies, which, in turn, would generate opportunities for the targeted improvement of their implementation. Importantly, organizations could draw on their practical experience to identify the domains in which robots and VR can be deployed most effectively.

Third, regular interactions between leaders, followers, and novel technologies can increase the acceptance of those technologies. Technology acceptance is an essential factor that has been investigated extensively in research on human-computer interactions because it predicts the intention to adopt and use technology (Cheung & Vogel, 2013; Marangunic & Granic, 2015). If technologies are integral to day-to-day operations, humans could become accustomed to their existence and learn how to use them most effectively. Most importantly, potentially extreme views of technology, positive or negative, could be balanced through exposure to and experience of its benefits and drawbacks.

Ethical Considerations

The power of technology also creates ethical issues that must be discussed. First, as technologies improve, they may become so effective that they outperform humans in tasks that remain in the human domain. Suppose robot leaders or VR environments improve to a point where technology can easily replace human leaders. In that case, the role of leaders in society and the limits of technology will become unclear. From a philosophical perspective, utilizing technology to free up time for virtuous action could be considered ethical, but depriving humans of opportunities to take moral action might be unethical (Zhu, 2020). For example, if AI becomes so effective at providing leadership that humans can allocate more to sustainability interventions, human resources for moral action would be freed up. However, if AI acquires the ability to replace leaders at some point in the future, human leaders would miss opportunities to enact ethical leadership behaviors, potentially leading to the notion of ethical leadership being unlearned over time, evidently an unethical use of technology. The current discourse on ethics and technology already touches on dystopian scenarios in which technology causes destructive outcomes for humanity. Discussions and research on these topics must be supported to prepare humans for the possible scenarios.

Second, while the intentions behind the use of technology in leadership might be virtuous, it is impossible to predict how technologies will develop over time (Hagendorff, 2020). If, for instance, the AI of a robot leader decides that exploitation or abusive supervision are the most effective forms of leadership, the humans who are led by the robot would suffer. Similar effects would occur in VR if the virtual environment were used to assess leaders for destructive-leadership behaviors when the AI deems them most effective. The substantial public interest in AI further leads to conflicts between the rapid implementation of AI, spurred by economic interests and the AI race, and the ethical guidelines established to avoid the potentially catastrophic consequences of AI being developed irresponsibly (Hagendorff, 2020). The organizations that intend to use AI and the developers responsible for its implementation must ensure that appropriate guidelines are followed to avoid the negative outcomes of unethical AI use.

Third, data privacy and security become relevant when leadership interactions occur between humans and robots or between avatars and characters in VR environments. Since the information accumulated from interactions with technologies is the foundation of efforts to evaluate and improve digital tools, user data must be saved per data-privacy standards. Users must always be informed how their data will be used, and they must have the ability to withdraw their consent to the use of their data. Similarly, cybersecurity threats become increasingly relevant as the introduction of technologies into organizations creates the possibility of infiltration by actors with destructive intentions (Sanclemente & Cardozo, 2022). When they introduce AI to robots or virtual characters, organizations must ensure that technologies are not being compromised and misused for nefarious purposes by establishing robust security measures.

Conclusion

Technologies are increasingly gaining ground in the real world, and their capabilities are growing more powerful. For now, humans remain in charge of their effective and responsible utilization. My results highlight the importance of technologies during digital transformation processes and their potential for organizations. How these technologies can be employed effectively at organizations depends strongly on their specific implementations. Researchers must build on prior work to ensure optimal outcomes for society, organizations, and individuals.

6. References

Abidi, M. H., Al-Ahmari, A. M., Ahmad, A., Darmoul, S., & Ameen, W. (2018). Semi-immersive virtual turbine engine simulation system. *International Journal of Turbo and Jet Engines*, *35*(2), 149-160.

Acemoglu, D., Lelarge, C., & Restrepo, P. (2020). Competing with robots: Firm-level evidence from france. AEA Papers and Proceedings,

Aguinis, H., & Bradley, K. J. (2014). Best practice recommendations for designing and implementing experimental vignette methodology studies. *Organizational Research Methods*, *17*(4), 351-371.

Ahmed, S., Athyaab, S. A., Muqtadeer, S. A., & Ieee. (2021, Jan 20-22). Attenuation of Human Bias in Artificial Intelligence: An Exploratory Approach. [Proceedings of the 6th international conference on inventive computation technologies (icict 2021)]. 6th International Conference on Inventive Computation Technologies (ICICT), Coimbatore, INDIA.

Alase, A. (2017). The interpretative phenomenological analysis (IPA): A guide to a good qualitative research approach. *International Journal of Education and Literacy Studies*, *5*(2), 9-19.

Allen, P. A., Kaut, K. P., & Lord, R. R. (2008). Emotion and episodic memory. *Handbook of Behavioral Neuroscience*, *18*, 115-132.

AlNuaimi, B. K., Singh, S. K., Ren, S., Budhwar, P., & Vorobyev, D. (2022). Mastering digital transformation: The nexus between leadership, agility, and digital strategy. *Journal of Business Research*, *145*, 636-648. https://doi.org/10.1016/j.jbusres.2022.03.038

Alvesson, M. (2020). Upbeat leadership: A recipe for–or against–"successful" leadership studies. *The Leadership Quarterly*, *31*(6), 101439.

American Psychological Association. (2017). *Ethical principles of psychologists and code of conduct (2002, amended effective June 1, 2010, and January 1, 2017)*. http://www.apa.org/ethics/code/index.html

Anderson, M. H., & Sun, P. Y. (2017). Reviewing leadership styles: Overlaps and the need for a new 'full-range'theory. *International Journal of Management Reviews*, *19*(1), 76-96.

Andonova, E. (2006). On changing mental models of a wheelchair robot. Proceedings of the Workshop on'How People Talk to Computers, Robots, and Other Artificial Communication Partners', Hansewissenschaftskolleg, Delmenhorst,

Antonakis, J. (2012). Transformational and charismatic leadership. *The Nature of Leadership*, 256-288.

Antonakis, J., Ashkanasy, N. M., & Dasborough, M. T. (2009). Does leadership need emotional intelligence? *Leadership Quarterly*, *20*(2), 247-261. https://doi.org/10.1016/j.leaqua.2009.01.006

Antonakis, J., Avolio, B. J., & Sivasubramaniam, N. (2003). Context and leadership: an examination of the nine-factor full-range leadership theory using the Multifactor Leadership Questionnaire. *The Leadership Quarterly*, *14*(3), 261-295. https://doi.org/10.1016/s1048-9843(03)00030-4

Antonakis, J., d'Adda, G., Weber, R., & Zehnder, C. (2014). Just words? Just speeches? On the economic value of charismatic leadership. *NBER Rep. 4*.

Antonakis, J., Fenley, M., & Liechti, S. (2011). Can Charisma Be Taught? Tests of Two Interventions. *Academy of Management Learning & Education*, *10*(3), 374-396. https://doi.org/10.5465/amle.2010.0012

Appel, M., Izydorczyk, D., Weber, S., Mara, M., & Lischetzke, T. (2020). The uncanny of mind in a machine: Humanoid robots as tools, agents, and experiencers. *Computers in Human Behavior*, *102*, 274-286.

Arai, T., Kato, R., & Fujita, M. (2010). Assessment of operator stress induced by robot collaboration in assembly. *CIRP annals*, *59*(1), 5-8.

Aronson, Z. H., Reilly, R. R., & Lynn, G. S. (2006). The impact of leader personality on new product development teamwork and performance: The moderating role of uncertainty. *Journal of Engineering and Technology Management*, *23*(3), 221-247.

Avolio, B. J., Bass, B. M., & Jung, D. I. (1995). MLQ multifactor leadership questionnaire: Technical report. *Redwood City, CA: Mindgarden*.

Avolio, B. J., Bass, B. M., & Jung, D. I. (1999). Re-examining the components of transformational and transactional leadership using the Multifactor Leadership Questionnaire. *Journal of Occupational and Organizational Psychology*, *72*, 441-462.

Avolio, B. J., Zhu, W., Koh, W., & Bhatia, P. (2004). Transformational leadership and organizational commitment: Mediating role of psychological empowerment and moderating role of structural distance. *Journal of Organizational Behavior: The International Journal of Industrial, Occupational and Organizational Psychology and Behavior, 25*(8), 951-968.

Azmat, G., & Petrongolo, B. (2014). Gender and the labor market: What have we learned from field and lab experiments? *Labour Economics*, *30*, 32-40.

Bartneck, C., & Forlizzi, J. (2004). A design-centred framework for social human-robot interaction. RO-MAN 2004. 13th IEEE international workshop on robot and human interactive communication (IEEE Catalog No. 04TH8759), 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.

Bartsch, S., Weber, E., Buttgen, M., & Huber, A. (2021). Leadership matters in crisis-induced digital transformation: how to lead service employees effectively during the COVID-19 pandemic. *Journal of Service Management*, *32*(1), 71-85. https://doi.org/10.1108/josm-05-2020-0160

Bass, B. M. (1999). Two decades of research and development in transformational leadership. *European Journal of Work and Organizational Psychology*, *8*(1), 9-32.

Bass, B. M., & Avolio, B. J. (1994). *Improving organizational effectiveness through transformational leadership*. Sage.

Batrinca, L., Stratou, G., Shapiro, A., Morency, L.-P., & Scherer, S. (2013). Cicero-towards a multimodal virtual audience platform for public speaking training. International workshop on intelligent virtual agents,

Bedi, A., Alpaslan, C. M., & Green, S. (2016). A Meta-analytic Review of Ethical Leadership Outcomes and Moderators. *Journal of Business Ethics*, *139*(3), 517-536. https://doi.org/10.1007/s10551-015-2625-1

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.

Bello, C. (2022). *Would you want a robot as CEO? Chinese firm is first to try as it bets on 'metaverse workplace'*. euronews.next. https://www.euronews.com/next/2022/11/20/would-you-want-a-robot-as-ceo-chinese-firm-is-first-to-try-as-it-bets-on-metaverse-workpla

Belpaeme, T., Kennedy, J., Ramachandran, A., Scassellati, B., & Tanaka, F. (2018). Social robots for education: A review. *Science robotics*, *3*(21), eaat5954.

Berkun, M. M. (1964). Performance decrement under psychological stress. *Human Factors*, *6*(1), 21-30.

Biocca, F., Harms, C., & Burgoon, J. K. (2003). Toward a more robust theory and measure of social presence: Review and suggested criteria. *Presence: Teleoperators & Virtual Environments*, *12*(5), 456-480.

Blascovich, J. (2002). Social influence within immersive virtual environments. In *The Social Life of Avatars* (pp. 127-145). Springer.

Blascovich, J., Loomis, J., Beall, A. C., Swinth, K. R., Hoyt, C. L., & Bailenson, J. N. (2002). Immersive virtual environment technology as a methodological tool for social psychology. *Psychological Inquiry*, *13*(2), 103-124.

Bledow, R., & Frese, M. (2009). A situational judgment test of personal initiative and its relationship to performance. *Personnel Psychology*, *62*(2), 229-258.

Bombari, D., Schmid Mast, M., Canadas, E., & Bachmann, M. (2015). Studying social interactions through immersive virtual environment technology: virtues, pitfalls, and future challenges. *Frontiers in Psychology*, *6*, 869. https://doi.org/10.3389/fpsyg.2015.00869

Bowles, H. R., Babcock, L., & Lai, L. (2007). Social incentives for gender differences in the propensity to initiate negotiations: Sometimes it does hurt to ask. *Organizational Behavior and Human Decision Processes*, *103*(1), 84-103.

Branigan, H., & Pearson, J. (2006). Alignment in human-computer interaction. *How people talk to computers, robots, and other artificial communication partners*, 140-156.

Braun, S., Peus, C., & Frey, D. (2018). Connectionism in action: Exploring the links between leader prototypes, leader gender, and perceptions of authentic leadership. *Organizational Behavior and Human Decision Processes*, *149*, 129-144.

Braun, S., Peus, C., Weisweiler, S., & Frey, D. (2013). Transformational leadership, job satisfaction, and team performance: A multilevel mediation model of trust. *The Leadership Quarterly*, *24*(1), 270-283.

Broekens, J., Heerink, M., & Rosendal, H. (2009). Assistive social robots in elderly care: a review. *Gerontechnology*, *8*(2), 94-103.

Buck, R., Losow, J. I., Murphy, M. M., & Costanzo, P. (1992). Social facilitation and inhibition of emotional expression and communication. *Journal of Personality and Social Psychology*, *63*(6), 962.

Bulu, S. T. (2012). Place presence, social presence, co-presence, and satisfaction in virtual worlds. *Computers & Education*, *58*(1), 154-161.

Cameron, D., de Saille, S., Collins, E. C., Aitken, J. M., Cheung, H., Chua, A., Loh, E. J., & Law, J. (2021). The effect of social-cognitive recovery strategies on likability, capability and trust in social robots. *Computers in Human Behavior*, *114*, 106561.

Canós-Darós, L. (2013). An algorithm to identify the most motivated employees. *Management Decision*.

Carpinella, C. M., Wyman, A. B., Perez, M. A., & Stroessner, S. J. (2017). The robotic social attributes scale (RoSAS) development and validation. Proceedings of the 2017 ACM/IEEE International Conference on human-robot interaction,

Charness, G., & Fehr, E. (2015). From the lab to the real world. *Science*, *350*(6260), 512-513.

Cheng, H., Jia, R., Li, D., & Li, H. (2019). The rise of robots in China. *Journal of Economic Perspectives*, *33*(2), 71-88.

Cheung, R., & Vogel, D. (2013). Predicting user acceptance of collaborative technologies: An extension of the technology acceptance model for e-learning. *Computers & Education*, *63*, 160-175. https://doi.org/10.1016/j.compedu.2012.12.003

Chiacchio, F., Petropoulos, G., & Pichler, D. (2018). *The impact of industrial robots on EU employment and wages: A local labour market approach*.

Cichor, J. E., Emmerling, F., & Hubner-Benz, S. (2019, August 29). *Transformational and Transactional Robot Leadership*. osf.io/58pqe

Cichor, J. E., Hubner-Benz, S., Benz, T., Emmerling, F., & Peus, C. (2023). Robot leadership–Investigating human perceptions and reactions towards social robots showing leadership behaviors. *PLOS ONE*, *18*(2), e0281786.

Dang, J., & Liu, L. (2021). Robots are friends as well as foes: Ambivalent attitudes toward mindful and mindless AI robots in the United States and China. *Computers in Human Behavior, 115*, 106612.

Davis, F. D. (1985). *A technology acceptance model for empirically testing new end-user information systems: Theory and results* Massachusetts Institute of Technology].

Day. (2011). Leadership Development: A Review in Context. *The Leadership Quarterly*, *11*(4).

DeRue, D. S., & Myers, C. G. (2014). Leadership development: A review and Agenda for future research. In *The Oxford handbook of leadership and organizations.* Oxford University Press.

Dionne, S. D., Yammarino, F. J., Atwater, L. E., & James, L. R. (2002). Neutralizing substitutes for leadership theory: Leadership effects and common-source bias. *Journal of Applied Psychology*, *87*(3), 454.

Draper, J. V., Kaber, D. B., & Usher, J. M. (1998). Telepresence. *Human Factors*, *40*(3), 354-375.

Eagly, A. H., Johannesen-Schmidt, M. C., & Van Engen, M. L. (2003). Transformational, transactional, and laissez-faire leadership styles: a meta-analysis comparing women and men. *Psychological Bulletin*, *129*(4), 569.

Eichenauer, C. J., Ryan, A. M., & Alanis, J. M. (2022). Leadership during crisis: an examination of supervisory leadership behavior and gender during COVID-19. *Journal of Leadership & Organizational Studies*, *29*(2), 190-207.

Eifler, S., & Petzold, K. (2019). Validity Aspects of Vignette Experiments: Expected "What-If" Differences Between Reports of Behavioral Intentions and Actual Behavior. *Experimental Methods in Survey Research: Techniques that Combine Random Sampling with Random Assignment*, 393-416.

Eva, N., Robin, M., Sendjaya, S., van Dierendonck, D., & Liden, R. C. (2019). Servant Leadership: A systematic review and call for future research. *Leadership Quarterly*, *30*(1), 111-132. https://doi.org/10.1016/j.leaqua.2018.07.004

Falk, A., Meier, S., & Zehnder, C. (2013). Do lab experiments misrepresent social preferences? The case of self-selected student samples. *Journal of the European Economic Association*, *11*(4), 839-852.

Fernandes, A. S., & Feiner, S. K. (2016). Combating VR sickness through subtle dynamic field-of-view modification. 2016 IEEE symposium on 3D user interfaces (3DUI),

Fischer, K. (2006). The role of users' preconceptions in talking to computers and robots. Proceedings of the Workshop on How People Talk to Computers, Robots, and other Artificial Communication Partners,

Fong, T., Nourbakhsh, I., & Dautenhahn, K. (2003). A survey of socially interactive robots. *Robotics and Autonomous Systems*, *42*(3-4), 143-166.

Franke, T., Attig, C., & Wessel, D. (2019). A personal resource for technology interaction: development and validation of the affinity for technology interaction (ATI) scale. *International Journal of Human–Computer Interaction*, *35*(6), 456-467.

Friese, M., Hofmann, W., & Wänke, M. (2008). When impulses take over: Moderated predictive validity of explicit and implicit attitude measures in predicting food choice and consumption behaviour. *British Journal of Social Psychology*, *47*(3), 397-419.

Gaddis, B., Connelly, S., & Mumford, M. D. (2004). Failure feedback as an affective event: Influences of leader affect on subordinate attitudes and performance. *The Leadership Quarterly*, *15*(5), 663-686.

Galaz, V., Centeno, M. A., Callahan, P. W., Causevic, A., Patterson, T., Brass, I., Baum, S., Farber, D., Fischer, J., Garcia, D., McPhearson, T., Jimenez, D., King, B., Larcey, P., & Levy, K. (2021). Artificial intelligence, systemic risks, and sustainability. *Technology in Society*, *67*, Article 101741. https://doi.org/10.1016/j.techsoc.2021.101741

Gandolfi, F., & Stone, S. (2018). Leadership, leadership styles, and servant leadership. *Journal of Management Research*, *18*(4), 261-269.

Ganster, D. C., & Rosen, C. C. (2013). Work Stress and Employee Health: A Multidisciplinary Review. *Journal of Management*, *39*(5), 1085-1122. https://doi.org/10.1177/0149206313475815

Garcia-Palacios, A., Botella, C., Hoffman, H., & Fabregat, S. (2007). Comparing acceptance and refusal rates of virtual reality exposure vs. in vivo exposure by patients with specific phobias. *Cyberpsychology & Behavior*, *10*(5), 722-724.

Garcia, M., Béchade, L., Dubuisson-Duplessis, G., Pittaro, G., & Devillers, L. (2017). Towards metrics of Evaluation of Pepper robot as a Social Companion for Elderly

People. Proceedings of the 8th International Workshop on Spoken Dialog Systems, IWSDS,

Garfin, D. R., Thompson, R. R., & Holman, E. A. (2018). Acute stress and subsequent health outcomes: A systematic review. *Journal of Psychosomatic Research*, *112*, 107-113. https://doi.org/10.1016/j.jpsychores.2018.05.017

Gencer, M. S., & Samur, Y. (2015, Dec 10-12). Leadership Styles and Technology: Leadership Competency Level of Educational Leaders.*Procedia Social and Behavioral Sciences* [5th international conference on leadership, technology, innovation and business management 2015, icltibm 2015]. 5th International Conference on Leadership, Technology, Innovation and Business Management (ICLTIBM), Istanbul, TURKEY.

Gladden, M. E. (2014). The Social Robot as' Charismatic Leader': A Phenomenology of Human Submission to Nonhuman Power. Robophilosophy,

Glikson, E., & Woolley, A. W. (2020). Human Trust in Artificial Intelligence: Review of Empirical Research. *Academy of Management Annals*, *14*(2), 627-660. https://doi.org/10.5465/annals.2018.0057

Goetz, J., Kiesler, S., & Powers, A. (2003). Matching robot appearance and behavior to tasks to improve human-robot cooperation. The 12th IEEE International Workshop on Robot and Human Interactive Communication, 2003. Proceedings. ROMAN 2003.,

Gombolay, M. C., Gutierrez, R. A., Clarke, S. G., Sturla, G. F., & Shah, J. A. (2015). Decision-making authority, team efficiency and human worker satisfaction in mixed human–robot teams. *Autonomous Robots*, *39*(3), 293-312.

Gooty, J., Connelly, S., Griffith, J., & Gupta, A. (2010). Leadership, affect and emotions: A state of the science review. *The Leadership Quarterly*, *21*(6), 979-1004.

Gordon, A., van Lent, M., Van Velsen, M., Carpenter, P., & Jhala, A. (2004). Branching storylines in virtual reality environments for leadership development. Proceedings of the national conference on Artificial Intelligence,

Gottfredson, L. S. (1997). Why g matters: The complexity of everyday life. *Intelligence*, *24*(1), 79-132.

Gottfredson, R. K., Wright, S. L., & Heaphy, E. D. (2020). A critique of the Leader-Member Exchange construct: Back to square one. *The Leadership Quarterly*, *31*(6), 101385.

Graetz, G., & Michaels, G. (2018). Robots at work. *Review of Economics and Statistics*, *100*(5), 753-768.

Gray, H. M., Gray, K., & Wegner, D. M. (2007). Dimensions of mind perception. *Science*, *315*(5812), 619-619.

Gray, K., & Wegner, D. M. (2012). Feeling robots and human zombies: mind perception and the uncanny valley. *Cognition*, *125*(1), 125-130. https://doi.org/10.1016/j.cognition.2012.06.007

Gumusluoglu, L., & Ilsev, A. (2009). Transformational leadership, creativity, and organizational innovation. *Journal of Business Research*, *62*(4), 461-473.

Hagendorff, T. (2020). The Ethics of AI Ethics: An Evaluation of Guidelines. *Minds and Machines*, *30*(1), 99-120. https://doi.org/10.1007/s11023-020-09517-8

Hansbrough, T. K., Lord, R. G., & Schyns, B. (2015). Reconsidering the accuracy of follower leadership ratings. *The Leadership Quarterly*, *26*(2), 220-237. https://doi.org/10.1016/j.leaqua.2014.11.006

Hansen, A. M., Kraemmergaard, P., & Mathiassen, L. (2011). RAPID ADAPTATION IN DIGITAL TRANSFORMATION: A PARTICIPATORY PROCESS FOR ENGAGING IS AND BUSINESS LEADERS. *Mis Quarterly Executive*, *10*(4), 175-185. <Go to ISI>://WOS:000298165600003

Harms, P., Credé, M., Tynan, M., Leon, M., & Jeung, W. (2017). Leadership and stress: A meta-analytic review. *The Leadership Quarterly*, *28*(1), 178-194.

Hasler, B. S., Spanlang, B., & Slater, M. (2017). Virtual race transformation reverses racial in-group bias. *PLOS ONE*, *12*(4), e0174965.

Hayes, A. F. (2012). *PROCESS: A versatile computational tool for observed variable mediation, moderation, and conditional process modeling [White paper]*. http://www.afhayes.com/public/process2012.pdf

Holmstrom, J. (2022). From AI to digital transformation: The AI readiness framework. *Business Horizons*, *65*(3), 329-339. https://doi.org/10.1016/j.bushor.2021.03.006

Holstein, T., Dodig-Crnkovic, G., & Pelliccione, P. (2018). Ethical and social aspects of self-driving cars. *arXiv preprint arXiv:1802.04103*.

Horstmann, A. C., & Krämer, N. C. (2020). When a Robot Violates Expectations: The Influence of Reward Valence and Expectancy Violation on People's Evaluation of a Social Robot. Companion of the 2020 ACM/IEEE International Conference on Human-Robot Interaction,

Howard, M. C. (2017). Investigating the simulation elements of environment and control: Extending the Uncanny Valley Theory to simulations. *Computers & Education*, *109*, 216-232. https://doi.org/10.1016/j.compedu.2017.03.005

Huang, C. M., Mutlu, B., & Assoc Comp, M. (2012, Mar 05-08). Robot Behavior Toolkit: Generating Effective Social Behaviors for Robots.*ACM IEEE International Conference on Human-Robot Interaction* [Hri'12: Proceedings of the seventh annual acm/ieee international conference on human-robot interaction]. 7th ACM/IEEE International Conference on Human-Robot Interaction (HRI), Boston, MA.

Hughes, D. J., Lee, A., Tian, A. W., Newman, A., & Legood, A. (2018). Leadership, creativity, and innovation: A critical review and practical recommendations. *The Leadership Quarterly*, *29*(5), 549-569.

IFR. (2019). *World Robotics 2019*. Retrieved 01.12.2019 from https://ifr.org/img/office/Sales_Flyer_World_Robotics_2019_web.pdf

Jordan, P. J., Dasborough, M. T., Daus, C. S., & Ashkanasy, N. M. (2010). A call to context. *Industrial and Organizational Psychology*, *3*(2), 145-148.

Judge, T. A., Fluegge Woolf, E., Hurst, C., & Livingston, B. (2006). Charismatic and transformational leadership: A review and an agenda for future research. *Zeitschrift für Arbeits-und Organisationspsychologie A&O*, *50*(4), 203-214.

Judge, T. A., & Piccolo, R. F. (2004). Transformational and transactional leadership: a meta-analytic test of their relative validity. *Journal of Applied Psychology*, *89*(5), 755.

Kark, R. (2011). Games Managers Play: Play as a Form of Leadership Development. *Academy of Management Learning & Education*, *10*(3), 507-527. https://doi.org/10.5465/amle.2010.0048

Kark, R., & Van Dijk, D. (2007). Motivation to lead, motivation to follow: The role of the self-regulatory focus in leadership processes. *Academy of Management Review*, *32*(2), 500-528.

Kätsyri, J., Förger, K., Mäkäräinen, M., & Takala, T. (2015). A review of empirical evidence on different uncanny valley hypotheses: support for perceptual mismatch as one road to the valley of eeriness. *Frontiers in Psychology*, *6*, 390.

Kätsyri, J., Mäkäräinen, M., & Takala, T. (2017). Testing the ‘uncanny valley’ hypothesis in semirealistic computer-animated film characters: An empirical evaluation of natural film stimuli. *International Journal of Human-Computer Studies*, *97*, 149-161. https://doi.org/10.1016/j.ijhcs.2016.09.010

Kennedy, J., Baxter, P., & Belpaeme, T. (2017). Nonverbal Immediacy as a Characterisation of Social Behaviour for Human-Robot Interaction. *International journal of Social Robotics*, *9*(1), 109-128. https://doi.org/10.1007/s12369-016-0378-3

Kennedy, J., Baxter, P., Belpaeme, T., & Acm. (2015, Mar 02-05). The Robot Who Tried Too Hard: Social Behaviour of a Robot Tutor Can Negatively Affect Child Learning.*ACM IEEE International Conference on Human-Robot Interaction* [Proceedings of the 2015 acm/ieee international conference on human-robot interaction (hri'15)]. 10th Annual ACM/IEEE International Conference on Human-Robot Interaction (HRI), Portland, OR.

Khalili, A. (2016). Linking transformational leadership, creativity, innovation, and innovation-supportive climate. *Management Decision*, *54*(9), 2277-2293.

Kiesler, S., Powers, A., Fussell, S. R., & Torrey, C. (2008). Anthropomorphic interactions with a robot and robot–like agent. *Social Cognition*, *26*(2), 169-181.

Kilteni, K., Groten, R., & Slater, M. (2012). The sense of embodiment in virtual reality. *Presence: Teleoperators and Virtual Environments*, *21*(4), 373-387.

Kim, T., & Hinds, P. (2006). Who should I blame? Effects of autonomy and transparency on attributions in human-robot interaction. ROMAN 2006-The 15th IEEE International Symposium on Robot and Human Interactive Communication,

Kirby, R., Forlizzi, J., & Simmons, R. (2010). Affective social robots. *Robotics and Autonomous Systems*, *58*(3), 322-332.

Klijn, M., & Tomic, W. (2010). A review of creativity within organizations from a psychological perspective. *Journal of Management Development*.

Koenig, N., Takayama, L., & Matarić, M. (2010). Communication and knowledge sharing in human–robot interaction and learning from demonstration. *Neural Networks*, *23*(8-9), 1104-1112.

Kourtesis, P., Collina, S., Doumas, L. A., & MacPherson, S. E. (2019). Technological competence is a pre-condition for effective implementation of virtual reality head mounted displays in human neuroscience: a technological review and meta-analysis. *Frontiers in Human Neuroscience*, *13*, 342.

Kozak, M. N., Marsh, A. A., & Wegner, D. M. (2006). What do I think you're doing? Action identification and mind attribution. *Journal of Personality and Social Psychology*, *90*(4), 543.

Kozima, H., Michalowski, M., Nakagawa, C., Kozima, H., Nakagawa, C., Kozima, H., & Michalowski, M. (2008). A Playful Robot for Research. *Therapy, and Entertainment*.

Krasikova, D. V., Green, S. G., & LeBreton, J. M. (2013). Destructive Leadership. *Journal of Management*, *39*(5), 1308-1338. https://doi.org/10.1177/0149206312471388

Kuhlen, A. K., & Brennan, S. E. (2013). Language in dialogue: When confederates might be hazardous to your data. *Psychonomic Bulletin & Review*, *20*(1), 54-72.

Kunisch, S., Menz, M., & Langan, R. (2022). Chief digital officers: An exploratory analysis of their emergence, nature, and determinants. *Long Range Planning*, *55*(2), Article 101999. https://doi.org/10.1016/j.lrp.2020.101999

Kurniawan, H., Maslov, A. V., & Pechenizkiy, M. (2013). Stress detection from speech and galvanic skin response signals. Proceedings of the 26th IEEE International Symposium on Computer-Based Medical Systems,

Kyriakou, M., Pan, X., & Chrysanthou, Y. (2017). Interaction with virtual crowd in Immersive and semi-Immersive Virtual Reality systems. *Computer Animation and Virtual Worlds*, *28*(5), e1729.

Lee, K. M. (2004). Presence, explicated. *Communication Theory*, *14*(1), 27-50.

Leitao, P., Colombo, A. W., & Karnouskos, S. (2016). Industrial automation based on cyber-physical systems technologies: Prototype implementations and challenges. *Computers in Industry*, *81*, 11-25. https://doi.org/10.1016/j.compind.2015.08.004

Leite, I., Martinho, C., & Paiva, A. (2013). Social robots for long-term interaction: a survey. *International journal of Social Robotics*, *5*(2), 291-308.

Leng, N. W. (2008). Transformational leadership and the integration of information and communications technology into teaching. *Asia-Pacific Education Researcher*, *17*(1), 1-14. <Go to ISI>://WOS:000257955500001

Levitt, S. D., & List, J. A. (2007). What do laboratory experiments measuring social preferences reveal about the real world? *Journal of Economic Perspectives*, *21*(2), 153-174.

Lewis, J., Matson, E. T., Wei, S., & Assoc Comp, M. (2012, Sep 05-08). Using Indistinguishability in Ubiquitous Robot Organizations. [Ubicomp'12: Proceedings of the 2012 acm international conference on ubiquitous computing]. 14th ACM International Conference on Ubiquitous Computing (UbiComp), Carnegie Mellon Univ, Pittsburgh, PA.

Lewis, J., Matson, E. T., Wei, S., & Min, B. C. (2013). Implementing HARMS-based indistinguishability in ubiquitous robot organizations. *Robotics and Autonomous Systems*, *61*(11), 1186-1192. https://doi.org/10.1016/j.robot.2013.04.001

Lewis, M., Wang, H. D., Chien, S. Y., Scerri, P., Velagapudi, P., Sycara, K., Kane, B., & Ieee. (2010, Oct 10-13). Teams Organization and Performance in Multi-Human/Multi-Robot Teams.*IEEE International Conference on Systems Man and Cybernetics Conference Proceedings* [Ieee international conference on systems, man and cybernetics (smc 2010)]. IEEE International Conference on Systems, Man and Cybernetics, Istanbul, TURKEY.

Liao, C.-M., & Masters, R. S. (2002). Self-focused attention and performance failure under psychological stress. *Journal of Sport and Exercise Psychology*, *24*(3), 289-305.

Lin, W. P., Shao, Y. D., Li, G. Q., Guo, Y. R., & Zhan, X. J. (2021). The Psychological Implications of COVID-19 on Employee Job Insecurity and its Consequences: The Mitigating Role of Organization Adaptive Practices. *Journal of Applied Psychology, 106*(3), 317-329. https://doi.org/10.1037/apl0000896

Ling, H. Y., & Bjorling, E. A. (2020). Sharing stress with a robot: what would a robot say? *Human-Machine Communication*, *1*, 133-159.

Lopes, S. L., Rocha, J. B., Ferreira, A. I., & Prada, R. (2021). Social robots as leaders: leadership styles in human-robot teams. 2021 30th IEEE International Conference on Robot & Human Interactive Communication (RO-MAN),

Lord, R. G., Brown, D. J., & Freiberg, S. J. (1999). Understanding the dynamics of leadership: The role of follower self-concepts in the leader/follower relationship. *Organizational Behavior and Human Decision Processes*, *78*(3), 167-203.

Lord, R. G., & Maher, K. J. (2002). *Leadership and information processing: Linking perceptions and performance*. Routledge.

Lu, L., Xie, Z., Wang, H., Li, L., & Xu, X. (2022). Mental stress and safety awareness during human-robot collaboration-Review. *Applied Ergonomics*, *105*, 103832.

Lyons, J. B., & Schneider, T. R. (2009). The effects of leadership style on stress outcomes. *The Leadership Quarterly*, *20*(5), 737-748.

MacDorman, K. F., & Chattopadhyay, D. (2016). Reducing consistency in human realism increases the uncanny valley effect; increasing category uncertainty does not. *Cognition*, *146*, 190-205. https://doi.org/10.1016/j.cognition.2015.09.019

MacDorman, K. F., & Chattopadhyay, D. (2017). Categorization-based stranger avoidance does not explain the uncanny valley effect. *Cognition*, *161*, 132-135. https://doi.org/10.1016/j.cognition.2017.01.009

Madhavan, P., & Wiegmann, D. A. (2007). Similarities and differences between human–human and human–automation trust: an integrative review. *Theoretical Issues in Ergonomics Science*, *8*(4), 277-301.

Mantovani, F., & Castelnuovo, G. (2003). The sense of presence in virtual training: enhancing skills acquisition and transfer of knowledge through learning experience in virtual environments.

Marangunic, N., & Granic, A. (2015). Technology acceptance model: a literature review from 1986 to 2013. *Universal Access in the Information Society*, *14*(1), 81-95. https://doi.org/10.1007/s10209-014-0348-1

Marteau, T. M., & Bekker, H. (1992). The development of a six-item short-form of the state scale of the Spielberger State—Trait Anxiety Inventory (STAI). *British Journal of Clinical Psychology*, *31*(3), 301-306.

Martirosov, S., Bureš, M., & Zítka, T. (2021). Cyber sickness in low-immersive, semi-immersive, and fully immersive virtual reality. *Virtual Reality*, 1-18.

Mast, M. S., Hall, J. A., & Roter, D. L. (2008). Caring and dominance affect participants' perceptions and behaviors during a virtual medical visit. *Journal of General Internal Medicine*, *23*(5), 523-527.

Matamala-Gomez, M., Donegan, T., Bottiroli, S., Sandrini, G., Sanchez-Vives, M. V., & Tassorelli, C. (2019). Immersive virtual reality and virtual embodiment for pain relief. *Frontiers in Human Neuroscience*, *13*, 279.

McDaniel, M. A., & Nguyen, N. T. (2001). Situational judgment tests: A review of practice and constructs assessed. *International Journal of Selection and Assessment*, *9*(1-2), 103-113.

Menon, S., & Suresh, M. (2020). Factors influencing organizational agility in higher education. *Benchmarking: An International Journal*.

Meslec, N., Curseu, P. L., Fodor, O. C., & Kenda, R. (2020). Effects of charismatic leadership and rewards on individual performance. *The Leadership Quarterly, 31*(6), 101423.

Mori, M., MacDorman, K. F., & Kageki, N. (2012). The uncanny valley [from the field]. *IEEE Robotics & Automation Magazine, 19*(2), 98-100.

Mota, R. C. R., Rea, D. J., Le Tran, A., Young, J. E., Sharlin, E., & Sousa, M. C. (2016). Playing the 'trust game'with robots: Social strategies and experiences. 2016 25th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN),

Mullins, M. (2006). Interpretation of simulations in interactive VR environments: Depth perception in cave and panorama. *Journal of Architectural and Planning Research*, 328-340.

Museth, K., Barr, A., & Lo, M. W. (2001). Semi-immersive space mission design and visualization: case study of the" Terrestrial Planet Finder" mission. Proceedings Visualization, 2001. VIS'01.,

Naidoo, L. J., Kohari, N. E., Lord, R. G., & DuBois, D. A. (2010). "Seeing" is retrieving: Recovering emotional content in leadership ratings through visualization. *The Leadership Quarterly*, *21*(5), 886-900.

Niculescu, A., van Dijk, B., Nijholt, A., Li, H., & See, S. L. (2013). Making social robots more attractive: the effects of voice pitch, humor and empathy. *International journal of Social Robotics*, *5*(2), 171-191.

Niemelä, M., Arvola, A., & Aaltonen, I. (2017). Monitoring the acceptance of a social service robot in a shopping mall: first results. Proceedings of the Companion of the 2017 ACM/IEEE International Conference on Human-robot Interaction,

Nisbett, R. E., & Wilson, T. D. (1977). The halo effect: evidence for unconscious alteration of judgments. *Journal of Personality and Social Psychology*, *35*(4), 250.

Nistor, N., Weinberger, A., Ceobanu, C., & Heymann, J. O. (2011, Sep 20-23). Educational Technology and Culture: The Influence of Ethnic and Professional Culture on Learners' Technology Acceptance.*Lecture Notes in Computer Science* [Towards ubiquitous learning, ec-tel 2011]. 6th European Conference on Technology-Enhanced Learning (EC-TEL), Palermo, ITALY.

Nitsch, V., & Glassen, T. (2015). Investigating the effects of robot behavior and attitude towards technology on social human-robot interactions. 2015 24th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN),

Oh, C. S., Bailenson, J. N., & Welch, G. F. (2018). A systematic review of social presence: Definition, antecedents, and implications. *Frontiers in Robotics and AI, 5*, 114.

Palmas, F., Cichor, J., Plecher, D. A., & Klinker, G. (2019). Acceptance and effectiveness of a virtual reality public speaking training. 2019 IEEE international symposium on mixed and augmented reality (ISMAR),

Palmas, F., Reinelt, R., Cichor, J. E., Plecher, D. A., & Klinker, G. (2021). Virtual Reality Public Speaking Training: Experimental Evaluation of Direct Feedback Technology Acceptance. 2021 IEEE Virtual Reality and 3D User Interfaces (VR), Pan, M., Croft, E. A., & Niemeyer, G. (2017). Validation of the Robot Social Attributes Scale (RoSAS) for human-robot interaction through a human-to-robot handover use case. Proceedings of the IROS 2017 Workshop on Human-Robot Interaction in Collaborative Manufacturing Environments, Vancouver, BC, Canada,

Park, C., Kim, D. G., Cho, S., & Han, H. J. (2019). Adoption of multimedia technology for learning and gender difference. *Computers in Human Behavior*, *92*, 288-296. https://doi.org/10.1016/j.chb.2018.11.029

Parong, J., & Mayer, R. E. (2018). Learning Science in Immersive Virtual Reality. *Journal of Educational Psychology*. https://doi.org/10.1037/edu0000241

Parra, E., Chicchi Giglioli, I. A., Philip, J., Carrasco-Ribelles, L. A., Marín-Morales, J., & Alcañiz Raya, M. (2021). Combining Virtual Reality and Organizational Neuroscience for Leadership Assessment. *Applied Sciences*, *11*(13). https://doi.org/10.3390/app11135956

Peck, T. C., Seinfeld, S., Aglioti, S. M., & Slater, M. (2013). Putting yourself in the skin of a black avatar reduces implicit racial bias. *Consciousness and Cognition*, *22*(3), 779-787.

Perala, C. H., & Sterling, B. S. (2007). *Galvanic skin response as a measure of soldier stress*. A. R. Laboratory.

Pertaub, D.-P., Slater, M., & Barker, C. (2002). An experiment on public speaking anxiety in response to three different types of virtual audience. *Presence*, *11*(1), 68-78.

Peus, C., Braun, S., & Frey, D. (2013). Situation-based measurement of the full range of leadership model — Development and validation of a situational judgment test. *The Leadership Quarterly*, *24*(5), 777-795. https://doi.org/10.1016/j.leaqua.2013.07.006

Pflaum, A. A., & Gölzer, P. (2018). The IoT and digital transformation: toward the data-driven enterprise. *IEEE pervasive computing*, *17*(1), 87-91.

Pierce, C. A., & Aguinis, H. (1997). Using virtual reality technology in organizational behavior research. *Journal of Organizational Behavior: The International Journal of Industrial, Occupational and Organizational Psychology and Behavior*, *18*(5), 407-410.

Podsakoff, N. P., Podsakoff, P. M., MacKenzie, S. B., & Klinger, R. L. (2013). Are we really measuring what we say we're measuring? Using video techniques to supplement traditional construct validation procedures. *Journal of Applied Psychology*, *98*(1), 99.

Podsakoff, P. M., MacKenzie, S. B., Moorman, R. H., & Fetter, R. (1990). Transformational leader behaviors and their effects on followers' trust in leader, satisfaction, and organizational citizenship behaviors. *The Leadership Quarterly*, *1*(2), 107-142.

Poeschl, S., & Doering, N. (2012). Designing virtual audiences for fear of public speaking training–an observation study on realistic nonverbal behavior. *Annual Review of Cybertherapy in Telemedicine. Advanced Technologies in the Behavioral, Social and Neurosciences*, *181*, 218.

Powers, A., Kiesler, S., Fussell, S., & Torrey, C. (2007). Comparing a computer agent with a humanoid robot. Proceedings of the ACM/IEEE international conference on Human-robot interaction,

Rasool, S. F., Wang, M., Tang, M., Saeed, A., & Iqbal, J. (2021). How toxic workplace environment effects the employee engagement: the mediating role of organizational support and employee wellbeing. *International journal of Environmental Research and Public Health*, *18*(5), 2294.

Ray, B., Jung, J., & Larabi, M.-C. (2018). On the possibility to achieve 6-DoF for 360 video using divergent multi-view content. 2018 26th European Signal Processing Conference (EUSIPCO), Raya, A., Luis, M., Chicchi Giglioli, I. A., & Parra Vargas, E. (2018). Virtual reality as an emerging methodology for leadership assessment and training. *Frontiers in Psychology, 9,* 1658.

Reich-Stiebert, N., & Eyssel, F. (2015). Learning with educational companion robots? Toward attitudes on education robots, predictors of attitudes, and application potentials for education robots. *International journal of Social Robotics, 7*(5), 875-888.

Rojahn, K., & Willemsen, T. M. (1994). The evaluation of effectiveness and likability of gender-role congruent and gender-role incongruent leaders. *Sex Roles, 30*(1), 109-119.

Ronay, R., & Kim, D. Y. (2006). Gender differences in explicit and implicit risk attitudes: A socially facilitated phenomenon. *British Journal of Social Psychology, 45*(2), 397-419.

Rowold, J., & Rohmann, A. (2009). Transformational and transactional leadership styles, followers' positive and negative emotions, and performance in German nonprofit orchestras. *Nonprofit management and leadership, 20*(1), 41-59.

Saarikko, T., Westergren, W. H., & Blomquist, T. (2020). Digital transformation: Five recommendations for the digitally conscious firm. *Business Horizons, 63*(6), 825-839. https://doi.org/10.1016/j.bushor.2020.07.005

Safir, M. P., Wallach, H. S., & Bar-Zvi, M. (2012). Virtual reality cognitive-behavior therapy for public speaking anxiety: one-year follow-up. *Behavior Modification, 36*(2), 235-246.

Salter, N. P., & Highhouse, S. (2009). Assessing managers' common sense using situational judgment tests. *Management Decision*.

Samani, H. A., Cheok, A. D., & Ieee. (2011, May 19-21). From Human-Robot Relationship to Robot-Based Leadership.*Conference on Human System Interaction* [4th international conference on human system interaction (hsi 2011)]. 4th International Conference on Human System Interaction (HSI), Keio Univ, Yokohama, JAPAN.

Samani, H. A., Koh, J. T. K. V., Saadatian, E., & Polydorou, D. (2012). Towards robotics leadership: An analysis of leadership characteristics and the roles robots will inherit in future human society. Asian conference on intelligent information and database systems,

Sanchez-Vives, M. V., & Slater, M. (2005). From presence to consciousness through virtual reality. *Nature Reviews Neuroscience*, *6*(4), 332-339.

Sanclemente, G. L., & Cardozo, B. N. (2022). Reliability: understanding cognitive human bias in artificial intelligence for national security and intelligence analysis. *Security Journal*, *35*(4), 1328-1348. https://doi.org/10.1057/s41284-021-00321-2

Saredakis, D., Szpak, A., Birckhead, B., Keage, H. A., Rizzo, A., & Loetscher, T. (2020). Factors associated with virtual reality sickness in head-mounted displays: a systematic review and meta-analysis. *Frontiers in Human Neuroscience*, *14*, 96.

Schaufeli, W. B., Bakker, A. B., & Salanova, M. (2006). The measurement of work engagement with a short questionnaire: A cross-national study. *Educational and Psychological Measurement*, *66*(4), 701-716.

Schmid, E. A., Pircher Verdorfer, A., & Peus, C. (2019). Shedding light on leaders' self-interest: theory and measurement of exploitative leadership. *Journal of Management*, *45*(4), 1401-1433.

Schmid Mast, M., Kleinlogel, E. P., Tur, B., & Bachmann, M. (2018). The future of interpersonal skills development: Immersive virtual reality training with virtual humans. *Human Resource Development Quarterly*, *29*(2), 125-141. https://doi.org/10.1002/hrdq.21307

Schwarzmüller, T., Brosi, P., Duman, D., & Welpe, I. M. (2018). How does the digital transformation affect organizations? Key themes of change in work design and leadership. *mrev management revue*, *29*(2), 114-138.

Schwind, V., Knierim, P., Tasci, C., Franczak, P., Haas, N., & Henze, N. (2017). " These are not my hands!" Effect of Gender on the Perception of Avatar Hands in Virtual Reality. Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems,

Simaan, N., Yasin, R. M., & Wang, L. (2018). Medical technologies and challenges of robot-assisted minimally invasive intervention and diagnostics. *Annual Review of Control, Robotics, and Autonomous Systems, 1*, 465-490.

Simoes, A. C., Soares, A. L., & Barros, A. C. (2020). Factors influencing the intention of managers to adopt collaborative robots (cobots) in manufacturing organizations. *Journal of Engineering and Technology Management*, *57*, Article 101574. https://doi.org/10.1016/j.jengtecman.2020.101574

Simonovic, B., Stupple, E. J., Gale, M., & Sheffield, D. (2018). Performance under stress: An eye-tracking investigation of the Iowa Gambling Task (IGT). *Frontiers in Behavioral Neuroscience*, *12*, 217.

Sivarajah, U., Kamal, M. M., Irani, Z., & Weerakkody, V. (2017). Critical analysis of Big Data challenges and analytical methods. *Journal of Business Research*, *70*, 263-286. https://doi.org/10.1016/j.jbusres.2016.08.001

Skogstad, A., Einarsen, S., Torsheim, T., Aasland, M. S., & Hetland, H. (2007). The destructiveness of laissez-faire leadership behavior. *Journal of occupational health psychology*, *12*(1), 80.

Slater, M. (2003). A note on presence terminology. *Presence Connect*, *3*(3), 1-5.

Slater, M. (2017). Implicit learning through embodiment in immersive virtual reality. In *Virtual, augmented, and mixed realities in education* (pp. 19-33). Springer.

Slater, M., Antley, A., Davison, A., Swapp, D., Guger, C., Barker, C., Pistrang, N., & Sanchez-Vives, M. V. (2006). A virtual reprise of the Stanley Milgram obedience experiments. *PLOS ONE*, *1*(1), e39.

Slater, M., Guger, C., Edlinger, G., Leeb, R., Pfurtscheller, G., Antley, A., Garau, M., Brogni, A., & Friedman, D. (2006). Analysis of Physiological Responses to a Social Situation in an Immersive Virtual Environment. *Presence: Teleoperators and Virtual Environments*, *15*(5), 553-569. https://doi.org/10.1162/pres.15.5.553

Slater, M., & Steed, A. (2002). Meeting people virtually: Experiments in shared virtual environments. In *The social life of avatars* (pp. 146-171). Springer.

Smedegaard, C. V. (2019). Reframing the role of novelty within social HRI: from noise to information. 2019 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI),

Spanlang, B., Normand, J.-M., Borland, D., Kilteni, K., Giannopoulos, E., Pomés, A., González-Franco, M., Perez-Marcos, D., Arroyo-Palacios, J., & Muncunill, X. N. (2014). How to build an embodiment lab: achieving body representation illusions in virtual reality. *Frontiers in Robotics and AI*, *1*, 9.

Stafford, R. Q., MacDonald, B. A., Jayawardena, C., Wegner, D. M., & Broadbent, E. (2014). Does the robot have a mind? Mind perception and attitudes towards robots predict use of an eldercare robot. *International journal of Social Robotics*, *6*(1), 17-32.

Steed, A., Frlston, S., Lopez, M. M., Drummond, J., Pan, Y., & Swapp, D. (2016). An ‘in the wild’experiment on presence and embodiment using consumer virtual reality equipment. *IEEE Transactions on Visualization and Computer Graphics*, *22*(4), 1406-1414.

Stein, J. P., & Ohler, P. (2017). Venturing into the uncanny valley of mind-The influence of mind attribution on the acceptance of human-like characters in a virtual reality setting. *Cognition*, *160*, 43-50. https://doi.org/10.1016/j.cognition.2016.12.010

Suakanto, S., Siswanto, J., Kusumasari, T. F., Prasetyo, I. R., & Hardiyanti, M. (2021). Interview Bot for Improving Human Resource Management. 2021 International Conference on ICT for Smart Society (ICISS),

Svendsen, G. B., Johnsen, J. A. K., Almas-Sorensen, L., & Vitterso, J. (2013). Personality and technology acceptance: the influence of personality factors on the core constructs of the Technology Acceptance Model. *Behaviour & Information Technology*, *32*(4), 323-334. https://doi.org/10.1080/0144929x.2011.553740

Terrasse, M., Gorin, M., & Sisti, D. (2019). Social media, e-health, and medical ethics. *Hastings Center Report*, *49*(1), 24-33.

Thompson, E. R. (2007). Development and validation of an internationally reliable short-form of the positive and negative affect schedule (PANAS). *Journal of Cross-Cultural Psychology*, *38*(2), 227-242.

Thoroughgood, C. N., Sawyer, K. B., Padilla, A., & Lunsford, L. (2016). Destructive Leadership: A Critique of Leader-Centric Perspectives and Toward a More Holistic Definition. *Journal of Business Ethics*, *151*(3), 627-649. https://doi.org/10.1007/s10551-016-3257-9

Todorov, A., Mandisodza, A. N., Goren, A., & Hall, C. C. (2005). Inferences of competence from faces predict election outcomes. *Science*, *308*(5728), 1623-1626.

Toufaily, E., Zalan, T., & Ben Dhaou, S. (2021). A framework of blockchain technology adoption: An investigation of challenges and expected value. *Information & Management*, *58*(3), Article 103444. https://doi.org/10.1016/j.im.2021.103444

Turner, A. I., Smyth, N., Hall, S. J., Torres, S. J., Hussein, M., Jayasinghe, S. U., Ball, K., & Clow, A. J. (2020). Psychological stress reactivity and future health and disease outcomes: A systematic review of prospective evidence. *Psychoneuroendocrinology*, *114*, Article 104599. https://doi.org/10.1016/j.psyneuen.2020.104599

Ullman, D., & Malle, B. F. (2018). What does it mean to trust a robot? Steps toward a multidimensional measure of trust. Companion of the 2018 acm/ieee international conference on human-robot interaction,

Urhal, P., Weightman, A., Diver, C., & Bartolo, P. (2019). Robot assisted additive manufacturing: A review. *Robotics and Computer-Integrated Manufacturing*, *59*, 335-345.

van den Berghe, R., Verhagen, J., Oudgenoeg-Paz, O., Van der Ven, S., & Leseman, P. (2019). Social robots for language learning: A review. *Review of Educational Research*, *89*(2), 259-295.

Van Kleef, G. A., & Cote, S. (2022). The Social Effects of Emotions. *Annual Review of Psychology*, *73*, 629-658. https://doi.org/10.1146/annurev-psych-020821-010855

Vardiman, P. D., Houghton, J. D., & Jinkerson, D. L. (2006). Environmental leadership development: Toward a contextual model of leader selection and effectiveness. *Leadership & Organization Development Journal*.

Vial, G. (2019). Understanding digital transformation: A review and a research agenda. *Journal of Strategic Information Systems*, *28*(2), 118-144. https://doi.org/10.1016/j.jsis.2019.01.003

Vigil-Colet, A., Ruiz-Pamies, M., Anguiano-Carrasco, C., & Lorenzo-Seva, U. (2012). The impact of social desirability on psychometric measures of aggression. *Psicothema*, *24*(2), 310-315.

Vroom, V. H., & Jago, A. G. (2007). The role of the situation in leadership. *American Psychologist*, *62*(1), 17.

Walters, M. L. (2008). *The design space for robot appearance and behaviour for social robot companions*

Walters, M. L., Koay, K. L., Syrdal, D. S., Dautenhahn, K., & Te Boekhorst, R. (2009). Preferences and perceptions of robot appearance and embodiment in human-robot interaction trials. *Procs of New Frontiers in Human-Robot Interaction*.

Walters, M. L., Syrdal, D. S., Dautenhahn, K., Te Boekhorst, R., & Koay, K. L. (2008). Avoiding the uncanny valley: robot appearance, personality and consistency of behavior in an attention-seeking home scenario for a robot companion. *Autonomous Robots*, *24*(2), 159-178.

Wang, G., Van Iddekinge, C. H., Zhang, L., & Bishoff, J. (2019). Meta-analytic and primary investigations of the role of followers in ratings of leadership behavior in organizations. *J Appl Psychol*, *104*(1), 70-106. https://doi.org/10.1037/apl0000345

Wang, Z., Zaman, S., Rasool, S. F., uz Zaman, Q., & Amin, A. (2020). Exploring the relationships between a toxic workplace environment, workplace stress, and project success with the moderating effect of organizational support: Empirical evidence from Pakistan. *Risk Management and Healthcare Policy*, *13*, 1055.

Warner, K. S. R., & Wager, M. (2019). Building dynamic capabilities for digital transformation: An ongoing process of strategic renewal. *Long Range Planning*, *52*(3), 326-349. https://doi.org/10.1016/j.lrp.2018.12.001

Weischer, A. E., Weibler, J., & Petersen, M. (2013). “To thine own self be true”: The effects of enactment and life storytelling on perceived leader authenticity. *The Leadership Quarterly*, *24*(4), 477-495.

Westlund, J. M. K., Martinez, M., Archie, M., Das, M., & Breazeal, C. (2016). Effects of framing a robot as a social agent or as a machine on children's social behavior. 2016 25th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN),

Whetzel, D. L., & McDaniel, M. A. (2009). Situational judgment tests: An overview of current research. *Human Resource Management Review*, *19*(3), 188-202.

Winston, B. E., & Patterson, K. (2006). An integrative definition of leadership. *International journal of Leadership Studies*, *1*(2), 6-66.

Wirtz, J., Patterson, P. G., Kunz, W. H., Gruber, T., Lu, V. N., Paluch, S., & Martins, A. (2018). Brave new world: service robots in the frontline. *Journal of Service Management*, *29*(5), 907-931. https://doi.org/10.1108/josm-04-2018-0119

Yasin Ghadi, M., Fernando, M., & Caputi, P. (2013). Transformational leadership and work engagement: The mediating effect of meaning in work. *Leadership & Organization Development Journal*, *34*(6), 532-550.

Yi, Y. D., Wu, Z., & Tung, L. L. (2005). How individual differences influence technology usage behavior? Toward an integrated framework. *Journal of Computer Information Systems*, *46*(2), 52-63. <Go to ISI>://WOS:000235019800006

Young, J., & Cormier, D. (2014). Can robots be managers, too. *Harvard Business Review*.

Yukl, G. (1989). Managerial leadership: A review of theory and research. *Journal of Management*, *15*(2), 251-289.

Yukl, G. (2006). *Leadership in organizations (6th ed.)*. Upper Saddle River, NJ: Pearson-Prentice Hall.

Zagermann, J., Pfeil, U., & Reiterer, H. (2016). Measuring cognitive load using eye tracking technology in visual computing. Proceedings of the sixth workshop on beyond time and errors on novel evaluation methods for visualization,

Zeng, Z. J., Chen, P. J., & Lew, A. A. (2020). From high-touch to high-tech: COVID-19 drives robotics adoption. *Tourism Geographies*, *22*(3), 724-734. https://doi.org/10.1080/14616688.2020.1762118

Zhang, L., Wade, J., Bian, D., Fan, J., Swanson, A., Weitlauf, A., Warren, Z., & Sarkar, N. (2017). Cognitive load measurement in a virtual reality-based driving system for autism intervention. *IEEE Transactions on Affective Computing*, *8*(2), 176-189.

Zhu, Q. (2020). Ethics, society, and technology: A Confucian role ethics perspective. *Technology in Society*, *63*, Article 101424. https://doi.org/10.1016/j.techsoc.2020.101424