Open Journal of Leadership, 2024, 13, 342-384
https://www.scirp.org/journal/ojl
ISSN Online: 2167-7751
ISSN Print: 2167-7743

Bridging the Uncanny Valley: Improving AI Chatbots for Effective Leadership Mentoring

Jörg Krauter

SYNK GROUP GmbH & Co. KG, Frankfurt, Germany
Email: joerg.krauter@synk-group.com

How to cite this paper: Krauter, J. (2024). Bridging the Uncanny Valley: Improving AI Chatbots for Effective Leadership Mentoring. *Open Journal of Leadership, 13,* 342-384.
https://doi.org/10.4236/ojl.2024.133021

Received: July 30, 2024
Accepted: September 9, 2024
Published: September 12, 2024

Abstract

Background: A significant challenge in developing AI assistants, such as chatbots as mentors or coaches for leaders, is the Uncanny Valley effect. This phenomenon describes the discomfort leaders feel when AI-chatbots appear almost human-like but not quite perfect, causing a sense of eeriness. To avoid this, a chatbot-design should focus on creating distinctive features rather than overly human-like ones and balance human-likeness with chat-elements to maintain user comfort. Emphasizing functional and straightforward designs that prioritize leaders' interaction with the chatbot and usability over appearance. **Method:** This study employed a qualitative research method, utilizing a qualitative content analysis to explore the extent of the Uncanny Valley effect in the interaction between AI chatbots and leaders. Data was collected through an online survey involving 82 leaders, of which 62 completed the entire questionnaire. Participants were shown video excerpts from an existing AI-based chatbot named "KI.m" and participated in semi-structured interviews to provide insights into their perceptions and emotional responses. **Results:** Key findings indicate that leader's expectation, voice quality, emotional expression, and social presence significantly influence user discomfort. Participants preferred more natural and human-like voices, suggesting that developers refine pronunciation and intonation. Emotional responses were often perceived as forced and insincere, highlighting the need for contextually relevant and genuine emotional expressions. Social interactions were seen as too mechanical, suggesting a balance between professional, supportive communication and subtle empathy. Personalization and accessibility through customizable settings and multilingual support were emphasized. Users preferred concise, clear answers. **Conclusion:** This study identified key themes influencing user discomfort with AI chatbots, including voice quality, emotional expression, social presence, and the need for aligning AI's cognitive abilities with user expectations. Enhancing personalization and accessibility in AI tools was also emphasized as crucial for improving user satisfaction and practical utility. The findings highlight the potential of AI mentoring tools to support leadership development, provided that key discomforts related to the Uncanny Valley effect are addressed. Future research should focus on refining AI attributes and exploring long-term impacts, cultural differences, and ethical considerations to enhance the design and implementation of AI mentoring systems.

Keywords

Uncanny Valley Effect, AI Chatbots, Leadership Mentoring, Chatflow Design Strategies

1. Introduction

The integration of AI into business is transforming leadership. AI enhances corporate performance and promotes sustainable leadership practices (Kaggwa, Eleogu, & Okonkwo, 2024). AI's role in leadership increases efficiency, allowing focus on complex tasks (Abasaheb & Subashini, 2023). While AI can automate tasks and support leadership development through tools like VR and apps, it requires understanding of AI systems and ethics (van Quaquebeke & Gerpott, 2023). Effective human-AI collaboration is vital for success, relying on leaders confidence in AI and supportive climates (Bankins, Ocampo, & Marrone, 2024). Leaders and employees often see algorithmic decisions as less fair than human ones, highlighting the need for a human element in leadership (McGuire & De Cremer, 2023). Digital cognitive assistants are well-accepted, requiring technological competence (Petrat et al., 2022). Leaders need technological knowledge, data utilization, interpersonal skills, and change leadership for successful AI integration (Ressem, 2023). A significant challenge in developing AI assistants for leadership is the Uncanny Valley effect, where almost human-like avatars cause discomfort (Strong & Terblanche, 2020; Doğan Erdinç & Uzunçarşılı Soydaş, 2024). Designers should focus on distinctive machine-like features, balancing human-likeness with robotic elements to maintain comfort (Petrat et al., 2022; Creasey & Vázquez Anido, 2020). Emphasizing functional designs that prioritize interaction and usability over appearance can enhance acceptance in leadership contexts (Rajaobelina, Prom Tep, & Arcand, 2021; McGuire, De Cremer, & Hesselbarth, 2023).

This study is organized into several key sections to comprehensively explore the Uncanny Valley effect in AI chatbots used for leadership mentoring. Following the introduction, a detailed **Literature Review** is presented, examining existing research on AI in leadership, the Uncanny Valley phenomenon, and human-AI interactions. **Methodology** section outlines the qualitative research approach, including data collection and analysis methods used to investigate leaders' perceptions of AI chatbots. Next, **Results** section presents findings from the data analysis, identifying specific indicators of the Uncanny Valley effect and suggesting strategies to mitigate user discomfort. **Discussion** section interprets these findings in the context of existing literature, exploring their implications for the design and implementation of AI chatbots in leadership mentoring. Finally, **Conclusion** offers a summary of key insights, practical recommendations, and potential avenues for future research.

1.1. Objectives

The objective is to explore the Uncanny Valley effect in AI chatbot-leader interactions and identify strategies to mitigate it, enhancing acceptance of AI chatbots as coaches and mentors.

1.2. Propositions

Proposition 1: Observable indicators of the Uncanny Valley effect exist in AI chatbots used as coaches/mentors for leaders. This study seeks to identify these indicators.

Proposition 2: The Uncanny Valley effect can be mitigated by design strategies that reduce the uncanny valley effect. This study aims to explore these strategies.

1.3. Research Questions

1) **Research Question 1:** What indicators of the Uncanny Valley effect are present in AI chatbots used as coaches/mentors for leaders?

2) **Research Question 2:** What design strategies can mitigate the Uncanny Valley effect in AI chatbots used as coaches/mentors for leaders?

2. Literature Review

This chapter delves into the multifaceted challenges and implications of integrating AI into leadership and human-computer interaction, particularly focusing on the Uncanny Valley effect and its impact on AI chatbots. The following sections synthesize findings from recent studies, offering insights into the psychological and behavioral implications of AI, human-AI collaboration, and effective strategies to mitigate the Uncanny Valley effect in AI chatbots. The discussion is structured around key themes: Leadership in the Digital Era, AI in Leadership, Human-AI Collaboration, Psychological and Behavioral Implications, the Uncanny Valley effect, and specific strategies for AI chatbots and computer speech. This comprehensive analysis aims to provide a deeper understanding of the current landscape and guide future research and development in AI-driven leadership tools.

2.1. Leadership in the Digital Era

The integration of AI into leadership roles is reshaping responsibilities and interactions between leaders, employees, and AI systems. Leaders must adapt to technological advancements to maintain effective leadership (Mihai & Crețu, 2019). Davenport and Foutty (2022) discuss the critical role of leaders in implementing AI, highlighting the need to prepare teams for AI integration and foster collaboration between people and machines. This preparation is essential for effectively using AI to enhance decision-making and operational efficiency. Leaders should adapt to changes brought by AI, manage emotional issues, and foster a decentralized leadership structure to remain agile and responsive (Banks, Dionne, Mast, & Sayama, 2022).

2.2. AI in Leadership

The rapid integration of AI into leadership roles is reshaping leadership. Abasaheb & Subashini (2023) shows that AI enhances leadership by fostering autonomy and innovation, critical for digital transformation. This shift challenges the belief that leadership is exclusively human. Van Quaquebeke and Gerpott (2023) argue that AI can meet employees' psychological needs better than human leaders, indicating a shift in leadership practices. Shick and Johnson (2023) highlight AI's role in rational decision-making by providing accurate data and decision alternatives, ending bounded rationality in decisions.

Integrating AI with emotional intelligence (EI) is crucial for effective leadership. Tîrnăcop (2023) notes that AI and EI together improve leaders' empathy, self-awareness, and social skills, enhancing overall effectiveness. This is vital as leaders navigate AI-driven environments, where technological skills must be complemented by interpersonal skills. Watson et al. (2021) discuss the future role of AI in senior leadership, emphasizing changes in organizational culture and leadership dynamics for effective integration. Aktaş (2024) also discusses AI's transformative impact on leadership paradigms and the need for new competencies in AI-driven environments. Paudel (2024) examines AI and automation's impact on leadership and workforce dynamics, noting both opportunities and challenges.

Primasatya et al. (2024) underscore the importance of adaptive leadership and self-leadership in integrating AI chatbot technology, especially in education. Arakawa and Yakura (2024) present a blended approach to leadership coaching, combining chatbots with human coaches to enhance self-reflection and growth. Pagnozzi, Renkema, and Antonelli (2024) examine AI-driven chatbots' effectiveness in supporting help-seeking behaviors, highlighting digital literacy and supportive culture as key factors.

2.3. Human-AI Collaboration and Psychological Implications

The integration of AI into leadership roles is transforming organizational dynamics, necessitating a shift from traditional hierarchies to more collaborative approaches. Waugh and Streib (2006) emphasize the need for collaboration and transformational leadership, a shift mirrored in broader contexts where AI demands new leadership paradigms. Strange (2022) proposes a multilevel framework for human-robot collaboration, highlighting the need for a holistic view of AI integration. This framework underscores the importance of cultivating a digital culture that fosters innovation and aligns organizational values with technological advancements (An, Lin, & Luo, 2024).

Trust is fundamental to effective human-AI collaboration. Zhang et al. (2023) find that performance consistency in AI teammates is crucial for building trust and cooperation. The impact of AI-oriented leadership on employee outcomes is complex; while it can enhance job performance, it may also increase unethical behavior, necessitating tailored leadership strategies that consider individual traits.

Van Quaquebeke and Gerpott (2023) argue that AI can better cater to employees' psychological needs than human leaders, prompting a reevaluation of leadership roles. However, blending human and AI leadership can blur responsibilities, impacting trust and relationship dynamics (Schafheitle, Weibel, & Rickert, 2021). Asfahani (2022) highlights AI's positive impact on job performance and efficiency but emphasizes the need for ethical design to ensure positive interactions. He et al. (2023) suggest that leaders' support for AI can stimulate job crafting and adaptability. This integration underscores the need for leaders to adapt to technological advancements while maintaining ethical practices and emotional intelligence.

2.4. Uncanny Valley Effect

In recent years, the Uncanny Valley effect has emerged as a significant topic in AI and robotics, particularly concerning human-robot interaction. It describes the discomfort caused by robots or AI closely resembling humans. This review synthesizes findings to understand factors contributing to the Uncanny Valley effect and strategies to mitigate it in AI chatbots and robotic design. Shoydin and Pazoev (2022) highlight avoiding the Uncanny Valley effect in holographic image transmission, emphasizing precise visual representation in AI. Lay (2015) explores how near-human faces evoke eerie feelings, identifying mismatched emotional expressions as a key factor, suggesting consistent emotional expressions to mitigate the effect. Tschöpe et al. (2017) reaffirm that emotions impact perceived human-likeness and uncanniness, aligning with Lay's (2015) suggestions. Ho and MacDorman (2017) provide indices for humanness and eeriness to design robots that avoid the Uncanny Valley effect.

Tu et al. (2020) reveal older adults prefer more humanlike robots, suggesting design should be tailored to age groups to avoid the effect. Skjuve et al. (2019) found transparency in conversational agents did not affect user experience negatively, identifying three crucial factors (conversation content, conversation demeanor, and conversation flow) for user assessments to mitigate the Uncanny Valley effect in chatbots. Blut et al. (2021) clarify that the impact of anthropomorphism on customer intentions varies by robot and service types, suggesting avoiding highly humanlike robots in certain contexts. Diel et al. (2022) suggest considering customer traits, sociodemographics, and robot characteristics in design. Mäkäräinen et al. (2015) show that strangeness can evoke positive emotions, suggesting design can embrace certain levels of strangeness.

In summary, the synthesis highlights the multifaceted nature of the Uncanny Valley effect, influenced by visual representation, emotional expression, demographic factors, user experience, and individual traits. The insights suggest the need for context-specific design strategies to mitigate the Uncanny Valley effect, enhancing AI chatbots and robots' acceptance and usability.

2.5. Uncanny Valley and Chatbots

The Uncanny Valley effect, where artificial agents elicit discomfort when they appear almost human-like but not quite, has significant implications for AI chatbots in leadership development. Pavlidou (2021) highlights that increased human-likeness in chatbots can decrease consumers' purchase intentions due to discomfort. This underscores the importance of balancing human-like features with abstract elements. Similarly, van Lierop (2021) found that enhancing human-likeness in customer support chatbots reduces perceived trust and satisfaction, recommending moderate human-likeness and functional efficiency.

Ekkelenkamp (2021) explored the effects of human-likeness and gender on perceived expertise and recommendation adherence of e-commerce chatbots, suggesting non-human characteristics combined with professional behavior are more effective. Nguyen and van der Lee (2021) emphasize the dual nature of human-like chatbots, which can enhance social presence but also induce discomfort, recommending simple, cartoonish designs. In therapeutic contexts, Cui, Wang, and Qi (2021) found that excessive anthropomorphism diminishes chatbot effectiveness, recommending empathetic interactions with minimal anthropomorphic features.

Lu (2021) suggests that high levels of agency and experience in chatbots can help mitigate Uncanny Valley effects by enhancing user acceptance. Wang and Shao (2022) discuss how anthropomorphic chatbots can alleviate loneliness but caution that the Uncanny Valley effect can reduce this benefit, advocating for human-like interactions without overemphasizing physical human-likeness. Betriana, Osaka, and Matsumoto (2021) suggest using abstract forms of affective communication and natural language processing to mitigate the Uncanny Valley effect. Wald, Heijselaar, and Bosse (2021) highlight the potential of chatbot customization in building user trust and avoiding the Uncanny Valley effect.

Overall, mitigating the Uncanny Valley effect in AI chatbots involves balancing human-like and machine-like characteristics, focusing on functional design, and allowing user customization. These strategies can enhance user acceptance and improve the effectiveness of AI chatbots in various applications, including leadership development.

2.6. Taxonomy of Uncanny Valley for Chatbots

The Uncanny Valley effect, where AI avatars and chatbots appear almost human-like but not quite perfect, is a significant challenge in AI development. Various studies have explored this effect, providing insights into user interaction and design strategies. Rapp, Curti, and Boldi (2021) reviewed text-based chatbots, finding that user acceptance, trust, engagement, and emotional experience are critical in human-chatbot interaction. Users often ascribe humanness to chatbots, influencing their emotional experience and trust. This review highlights the importance of designing chatbots that manage user expectations and emotional responses to avoid the Uncanny Valley effect.

Stepp Jr (2022) investigated variables reducing the Uncanny Valley effect in human-robot interaction, categorizing the effect into visual, intelligence, verbal, interaction, auditory, and nonverbal components. Visual and intelligence aspects are particularly relevant, encompassing appearance and perceived cognitive abilities. The study suggests balancing these elements to mitigate discomfort and enhance user acceptance.

Cui, Wang, and Qi (2021) explored visual and verbal anthropomorphism in therapeutic conversational agents, finding that while visual anthropomorphism can enhance social presence, it may also trigger the Uncanny Valley effect if too human-like. Verbal anthropomorphism positively influences user perception without eliciting uncanniness, suggesting a strategic combination of visual and verbal elements to enhance engagement while minimizing discomfort.

Mathur and Reichling (2020) examined mind perception as an explanation for the Uncanny Valley effect, concluding that how users perceive robot minds significantly influences their comfort with human-like appearances. This underscores designing AI with mind perception to foster trust and reduce eeriness.

Song and Shin (2022) studied the Uncanny Valley effects on chatbot trust, purchase intention, and adoption intention in e-commerce, finding that hyperrealistic avatars induce discomfort and reduce trust. They recommend using less humanized, cartoonish avatars to enhance comfort and trust.

Skjuve et al. (2019) investigated user experiences in human-chatbot interactions, finding that conversation flow and responsiveness are crucial for user experience, suggesting seamless interactions to prevent discomfort and improve satisfaction.

Mozafari, Hammerschmidt, and Weiger (2021) explored social presence in human-chatbot interactions, discovering that concealing the chatbot's identity can increase usage intentions in non-embarrassing contexts. In embarrassing contexts, higher social presence can hinder usage, indicating the need to consider interaction context to balance social presence and user comfort.

Overall, these studies provide a comprehensive understanding of the Uncanny Valley effect and suggest various strategies to mitigate it. By considering visual, verbal, and interaction elements, as well as user perception and context, designers can create AI chatbots that enhance user acceptance and engagement without triggering discomfort.

2.7. Taxonomy of Computer Speech in Chatbots

The uncanny valley effect, traditionally discussed in relation to human-like robots and avatars, is now being explored within the realm of computer speech.

Clark, Ofemile, and Cowan (2021) investigate this effect through the lens of verbal communication, focusing on vague language in synthetic and human-like voices. Their research highlights that vague language in human speech is perceived as natural, but in computer-generated voices, it can cause a perceptual mismatch and discomfort. Direct and unambiguous language aligns more closely with users' expectations of machine speech, reducing the uncanny valley effect (Clark et al., 2021). Appropriateness in computer speech is crucial, as it relates to the broader context of human-computer interaction (HCI) and users' partner models.

Cowan et al. (2015) discuss partner models, where users form mental representations of their interaction partners, including assumptions about the machine's capabilities. These models are disrupted when a machine uses socially-driven linguistic strategies such as politeness, which are reserved for human-human interaction (HHI). Social rules do not seamlessly transfer from HHI to HCI, evident in how vague language does not necessarily translate into a positive experience with a computer interface, especially if the voice is synthesized (Clark et al., 2016).

Brown and Levinson (1987) argue that politeness is tied to human interaction and maintaining face, which may be inappropriate in HCI. Research suggests that while human-like voices can moderate the uncanny valley effect, they do not eliminate perceptual mismatches. Other factors, such as the medium of interaction and the context, play significant roles in shaping what is perceived as appropriate speech. For example, a laptop providing task-based instructions may still be seen as inappropriate if it uses overly human-like language (Clark et al., 2016). The complexity of language use in HCI remains a challenge, and future research is needed to empirically test the boundaries of the verbal uncanny valley and identify what constitutes appropriate computer speech in various contexts.

3. Theoretical Framework

A theoretical framework is a foundational structure composed of established concepts and theories that guide a study and provide a basis for data analysis. It is developed from existing theories to explain phenomena and create a coherent structure for interpreting research findings. According to Varpio et al. (2020), it involves defining and connecting relevant concepts and theories, aiding in practical decisions about research design. Kivunja (2018) describes it as a synthesis of existing theories and concepts that act as a "theoretical coat hanger" for data analysis, situating research within an academic context for rigor and depth. A theoretical framework grounds research in established theories, ensuring robust data analysis and interpretation, guiding the research design, and enhancing the clarity of study conclusions.

Building on the literature review, two sources emerge for developing indicators for the Uncanny Valley effect and deriving design strategies to mitigate this effect in AI chatbots as coaches/mentors for leaders. These sources provide theoretical foundations and practical insights to guide the identification of key indicators and the formulation of design strategies to ensure AI chatbots are perceived as effective and trustworthy tools in leadership development.

Taxonomy of the Uncanny Valley Effect in Chatbots

This taxonomy synthesizes the criteria from various studies on the uncanny valley effect, focusing on computer speech and user interaction with AI chatbots. The categories created from this synthesis address different aspects of the user experience, from perceptual qualities to cognitive characteristics, and provide a comprehensive framework for understanding and mitigating the uncanny valley effect in AI chatbots. Table 1 below outlines these new combined categories, explaining their significance and detailing the specific criteria and descriptions that contribute to each category.

Table 1. Taxonomy of the uncanny valley effect in chatbots.

Category	Explanation	Criteria	Description/Source
Perceptual and Visual Characteristics	This category focuses on how users perceive the voice quality and visual appearance of the AI chatbot, including human-likeness and clarity. It combines elements of how the AI looks and sounds to the user.	• Voice Quality • Humanness • Visual Appearance • Visual Anthropomorphism	• **Voice Quality**: The degree to which the speech sounds human-like versus robotic (Clark, Ofemile, & Cowan, 2021). • **Humanness**: Degree to which users ascribe human-like qualities to chatbots (Rapp, Curti, & Boldi, 2021). • **Visual Appearance**: Human-likeness of robots' physical appearance (Stepp Jr, 2022). • **Visual Anthropomorphism**: Use of human-like visual features (Cui, Wang, & Qi, 2021).
Emotional and Social Characteristics	This category evaluates the AI chatbot's ability to express emotions appropriately and create a sense of social presence, making the interaction feel warm and human-like.	• Emotional Expression • Social Presence	• **Emotional Expression**: Chatbots' ability to express emotions appropriately (Rapp, Curti, & Boldi, 2021). • **Social Presence**: Perceived warmth and human-like interaction quality (Mozafari, Hammerschmidt, & Weiger, 2021).
Cognitive Characteristics	This category investigates the perceived cognitive abilities and mind perception of the AI chatbot, examining how users view the AI's intelligence and mental capabilities.	• Intelligence • Mind Perception	• **Intelligence**: Perceived cognitive abilities of AI (Stepp Jr, 2022). • **Mind Perception**: Users' perception of robots having a mind (Mathur & Reichling, 2020).
Linguistic and Verbal Characteristics	This category assesses the appropriateness and match of the language used by the AI chatbot against user expectations, including human-like verbal cues and the flow of conversation.	• Linguistic Appropriateness • Verbal Anthropomorphism	• **Linguistic Appropriateness**: How well the language used matches user expectations of machine versus human speech (Clark, Ofemile, & Cowan, 2021). • **Verbal Anthropomorphism**: Use of human-like verbal cues and backstories (Cui, Wang, & Qi, 2021).

Continued

Interactional Characteristics	This category evaluates the impact of social cues and the dynamics of interaction in AI communication, focusing on politeness, relational work, and specific interaction elements that stand out.	• Social Cues • Interactional Dynamics	• **Social Cues**: The presence of politeness, relational work, and other interpersonal linguistic strategies (Clark, Ofemile, & Cowan, 2021). • **Interactional Dynamics**: Responsiveness and fluidity of chatbot interactions (Skjuve et al., 2019).
Contextual and Cognitive Models	This category considers the medium and context in which the AI chatbot is used and how it aligns with users' mental representations and expectations of the AI's capabilities.	• Medium and Context of Interaction • User Partner Models	• **Medium and Context of Interaction**: The specific interface (e.g., laptop, mobile) and the context (e.g., task-based instructions) (Clark, Ofemile, & Cowan, 2021). • **User Partner Models**: Users' mental representations of the machine's capabilities and characteristics (Clark, Ofemile, & Cowan, 2021).
Consistency	This category measures the consistency between the type of voice (synthetic vs. pre-recorded human) and the language used by the AI chatbot, ensuring that the voice and language align seamlessly.	• Voice Consistency	• **Voice Consistency**: Consistency between the type of voice (synthetic vs. pre-recorded human) and the language used (Clark, Ofemile, & Cowan, 2021).

4. Methodology

4.1. Qualitative Research Methodology in Leadership Research

Qualitative research has become a crucial tool for examining leadership (Bryman, 2017; Conger, 1998). This approach provides a detailed and nuanced analysis of the complexities involved in leadership experiences and leaders' perceptions (Merriam & Grenier, 2019). Qualitative methods have been used to investigate both anticipated and unexpected phenomena in leaders' perceptions (Insch, Moore, & Murphy, 1997). Additionally, qualitative research introduces a broader array of contextual variables, including leaders' thinking about using AI-chatbots as coaches and mentors (Bryman, Bresnen, Beardsworth, & Keil, 1988), and offers a grounded perspective based on leaders' experiences, making it more accessible to researchers (Klenke, 2008). By examining leaders' perceptions (Bryman, Stephens, & a Campo, 1996), qualitative research provides valuable insights into the perception of uncanny valley effects. Given these advantages, qualitative research was selected for this study to offer a more thorough understanding of leaders' perceptions and their experiences of uncanny valley effects when considering AI-based chatbots as coaches and mentors.

4.2. Sample

In addition to the qualitative considerations outlined, the determination of the sample size for this study was also guided by established recommendations for various research designs. According to Onwuegbuzie and Collins (2007), the minimum sample size for common qualitative research designs typically ranges between 21 participants for experimental studies, 51 - 64 participants for causal-comparative research, and 64 - 82 participants for correlational studies. Given that this study involves elements of correlational research in its analysis of leaders' perceptions and the Uncanny Valley effect in AI chatbots, the target sample size was aligned with the upper range of these recommendations.

This study has examined the characteristics commonly found among leaders. A total of 82 leaders participated in the study, with 62 being selected for data gathering as they completed the entire questionnaire (see Table 2).

Table 2. Socio-demographic and role-specific characteristics of the study sample.

Socio-Demographic Characteristics	Frequency	Percentage
Gender		
Male	43	69.4%
Female	19	30.6%
Diverse	0	0.0%
Not indicated	0	0.0%
Total	62	
Age		
<30	15	24.2%
31 - 40	30	48.4%
41 - 50	11	17.7%
51 - 60	5	8.1%
>60	1	1.6%
Not indicated	0	0.0%
Total	62	
Highest Educational Qualification		
Secondary School	0	0.0%
Middle School	4	6.5%
High School	7	11.3%
Magister	3	4.8%
Diploma	1	1.6%
Bachelor	24	38.7%
Master	21	33.9%
Doctorate	2	3.2%
Not indicated	0	0.0%
Total	62	
Years of Experience as a Leader		
<1 year	3	4.8%
1 - 3 years	31	50.0%

Continued

4 - 5 years	13	21.0%
6 - 10 years	7	11.3%
>10 years	8	12.9%
Not indicated	0	0.0%
Total	62	
Hierarchical Position		
Top Management	9	14.5%
Middle Management	20	32.3%
Head of Department	15	24.2%
Team Leader	17	27.4%
Not indicated	1	1.6%
Total	62	
Area of Responsibility		
Organization	13	21.0%
Business Unit	13	21.0%
Team	35	56.5%
Not indicated	1	1.6%
Total	62	
Manager-to-Employee Ratio		
<5	21	33.9%
5 - 10	16	25.8%
11 - 20	16	25.8%
21 - 50	7	11.3%
>50	2	3.2%
Not indicated	0	0.0%
Total	62	

The participants were 69.4% male, 30.6% female, with none identifying as diverse or not indicating their gender. Regarding age, 24.2% of the participants were under 30 years old, 48.4% were between 31 and 40, 17.7% were between 41 and 50, 8.1% were between 51 and 60, and 1.6% were over 60. Educational qualifications varied, with none having only a secondary school diploma, 6.5% having a middle school diploma, 11.3% having a high school diploma, 4.8% holding a Magister degree, 1.6% having a diploma, 38.7% holding a bachelor's degree, 33.9% having a master's degree, and 3.2% holding a doctorate.

Experience as leaders also varied: 4.8% had less than one year of experience, 50.0% had between 1 and 3 years, 21.0% had between 4 and 5 years, 11.3% had between 6 and 10 years, and 12.9% had more than 10 years of experience. Regarding hierarchical positions, 14.5% were in top management, 32.3% in middle management, 24.2% were heads of departments, 27.4% were team leaders, and 1.6% did not indicate their hierarchical level.

In terms of responsibility areas, 21.0% were responsible for organizations, 21.0% for business units, and 56.5% for teams, with 1.6% not indicating their area of responsibility. The manager-to-employee ratio also varied, with 33.9% leading fewer than 5 employees, 25.8% leading between 5 and 10 employees, 25.8% leading between 11 and 20 employees, 11.3% leading between 21 and 50 employees, and 3.2% leading more than 50 employees.

The patterns visible in the sample show that the majority of participants were male (69.4%) and aged between 31 and 40 (48.4%). Most participants held either a bachelor's (38.7%) or master's degree (33.9%). In terms of experience, a significant portion had between 1 and 3 years of managerial experience (50.0%), and a majority had team responsibilities (56.5%). Most participants led fewer than 10 employees (59.7%).

4.3. Data Collection Methods

This research utilized an online survey based on a semi-structured interview questionnaire (Burgess, 2001; Kasunic, 2005), hosted on https://www.umfrageonline.com/ (Lumsden & Morgan, 2005). The usage of semi-structured interviews is aligned with the principles outlined by Barriball and While (1994), the study effectively balances structured questions with the flexibility to explore unexpected themes, enabling a deep exploration of managers' perceptions of the KI.m AI-based chatbot. The interview design includes a well-defined guide with open-ended questions, fostering consistent data collection while allowing for in-depth probing where necessary. Building rapport with participants, as facilitated by the option to take breaks or skip questions, encourages honest responses, which enhances the reliability and authenticity of the data. Ethical considerations, such as informed consent and confidentiality, are carefully maintained, supporting the validity of the study by ensuring participant trust and reducing response bias. Furthermore, the study employs rigorous post-interview data analysis, including thematic analysis, which systematically identifies patterns and ensures that findings are both reliable and valid.

Participants were informed they could take breaks or refuse to answer at any point.

The study utilized a semi-structured interview approach to gather qualitative data on the perceptions of KI.m, an AI-based chatbot designed for executive support and mentoring. The primary objective was to understand how managers perceive KI.m and identify potential areas for improvement, which could inform AI-based leadership development strategies.

4.3.1. Introduction

Participants were first introduced to the study's purpose, which centered on evaluating the effectiveness and reception of KI.m among executives. They were informed that KI.m is built on a science-based leadership model and employs advanced conversational AI technologies to provide tailored support and mentoring. This introduction set the stage for participants to comprehend the study's relevance and the role of KI.m in leadership development. KI.m is an innovative AI-based mentoring platform designed to enhance leadership development by providing personalized, data-driven insights and recommendations. Leveraging advanced machine learning algorithms and a deep understanding of leadership dynamics, KI.m offers tailored guidance to help leaders navigate complex challenges, improve decision-making, and foster growth in their teams. The platform aims to make high-quality leadership mentoring accessible and effective, utilizing conversational AI to simulate interactions with a seasoned mentor, providing real-time feedback and support. This comprehensive tool is designed to integrate seamlessly into the daily routines of leaders, offering ongoing support that adapts to individual needs and organizational contexts. By focusing on the unique aspects of each leader's situation, KI.m helps in developing the skills and strategies necessary for effective leadership in a rapidly changing business environment.

4.3.2. Demographic and Personality Data Collection

The interview commenced with demographic questions aimed at capturing essential background information about the participants. Questions included gender, age, highest educational qualification, years of professional experience as a manager, hierarchical position, field of responsibility, and the number of employees directly managed. Additionally, participants were asked to provide insights into their personality traits based on the Big Five personality model. This information was crucial for contextualizing the responses and understanding the diversity of the participant pool.

4.3.3. Thinking about KI.m Usage and Impression Assessment

Following the demographic section, participants were presented with an image highlighting the features of KI.m. Participants were asked to evaluate the potential support KI.m could provide in their leadership roles after being shown a series of images that highlighted various features of the platform. The first image (**Figure 1**) presented the key functionalities of the KI.m chatbot. Subsequent images depicted the chatflows related to order clarification and relationship management (**Figure 2(a)** and **Figure 2(b)**), as well as problem understanding and reflection (**Figure 3(a)** and **Figure 3(a)**). These visual aids aimed to demonstrate how KI.m's conversational AI could assist leaders in various aspects of their daily responsibilities.

The images in **Figure 2(a)** and **Figure 2(b)** depict the real chatflow of the KI.m chatbot. In this sequence, the chatbot begins by inquiring about the client's current emotional state to assess whether they are emotionally ready to engage in the mentoring process. Following this, the chatbot proceeds to gather information about the situation and context, ensuring that it fully understands the client's circumstances before offering guidance or support. This approach aims to tailor the mentoring experience to the client's immediate needs and emotional readiness, ensuring a more effective and personalized interaction.

Figure 1. Features of KI.m.

(a) (b)

Figure 2. Chatflow of KI.m order clarification and relationship management.

Figure 3(a) and Figure 3(b) illustrate the chatflow of the KI.m chatbot as it engages with the client to understand and reflect on their situation. In these images, the client describes their situation, problem, and specific questions they have. The KI.m chatbot then asks a clarifying question to gain a deeper understanding of the situation. Recognizing the emotional strain the client is under, the chatbot responds with empathy, acknowledging the client's feelings and providing support. This interaction is designed to ensure that the chatbot fully comprehends the client's concerns while also offering emotional reassurance.

(a) (b)

Figure 3. Chatflow of KI.m Problem Understanding and reflection.

This segment aimed to gauge initial perceptions and expectations regarding the AI mentor's capabilities.

4.3.4. Video Excerpts Evaluation

To provide a comprehensive understanding of KI.m's functionality, participants watched six short video excerpts showcasing different aspects of a mentoring session with KI.m. These excerpts included:

1) Greeting
2) Relationship building
3) Clarification of assignment (understanding the problem or question)
4) Active listening
5) Showing empathy
6) Giving recommendations and feedback on emotional situations

Participants were instructed to watch each video excerpt carefully and then respond to a series of open-ended questions designed to capture their immediate impressions and emotional reactions (Reja, Manfreda, Hlebec, & Vehovar, 2003). The questions were structured to explore their first impressions, the clarity and understandability of the content, the relevance to their expectations and needs, and the specific aspects of the videos that caught their attention. Participants were also asked to reflect on their emotional responses and the overall fit of the mentoring sequence with their professional experiences. The questions included:

1) What was your first impression of mentoring in the video?

2) How did this sequence affect you?

3) Does this question meet your expectations at this point?

4) What else would you expect at this point?

5) What emotions did you feel when watching the video?

6) How clear and understandable did you find the question in the video?

7) What aspects of the video caught your attention in particular?

8) How well does the mentoring sequence shown fit your expectations or needs?

9) You have now been able to gain a first impression of KI.m. "Hand on heart"—Would you use KI.m regularly in your everyday work? *

10) What is your suggestion for improving KI.m?

4.4. Data Protection and Confidentiality

Throughout the interview process, participants were assured of the confidentiality and anonymity of their responses. They were informed that their data would not be shared with third parties and would be treated with strict confidentiality. The voluntary nature of participation was emphasized, along with the option to withdraw from the study at any time without any repercussions. The study results were intended for use in research reports, technical articles, and conference presentations, with all data anonymized to protect participant identity.

This semi-structured interview methodology allowed for in-depth exploration of participant perceptions, providing rich qualitative data to inform the development and refinement of KI.m as an AI-based executive mentor.

5. Data Analysis Methods

Qualitative content analysis is a methodologically controlled approach to analyzing texts within their context (Mayring, 2004). This study employed deductive category application for the emotional perspective, where text passages were assigned to predefined categories based on the theoretical framework from the literature review (Fenzl & Mayring, 2017; Insch et al., 1997). Explicit definitions and examples for each deductive category were provided in a coding agenda (Roller, 2019). A derived coding schema contained category names, descriptions, and examples (Mayring, 2015, 2021).

Inductive coding was used to classify indicators of uncanny valley effects and criteria for design strategies, identifying themes or patterns in the data. This process involves collecting data, analyzing it to find patterns and meanings, and categorizing it into meaningful groups, allowing for understanding relationships and inferring general conclusions rather than verifying a hypothesis.

Combining deductive and inductive approaches helps reduce researcher bias (Insch et al., 1997), enabling coding and interpreting data with context awareness while evaluating it against existing theoretical assumptions from change models (Insch et al., 1997).

Data analysis involves breaking down data into smaller units and coding them to understand their content. This process includes four key stages: decontextualization, recontextualization, categorization, and compilation (Bengtsson, 2016; Insch et al., 1997):

- Decontextualization: Familiarizing with the data, breaking it into smaller units, and labeling them with codes using literature, frameworks, or software like QCAmap.
- Recontextualization: Reviewing the original text and unit list to ensure all content elements relate to the research.
- Categorization: Condensing units and determining themes and categories, requiring multiple reviews for accurate grouping.
- Compilation: Analyzing data in-depth to uncover underlying meanings, summarizing themes, categories, and sub-categories, and verifying with a panel of experts.

5.1. Criteria for Evaluating Trustworthiness and Validity of Qualitative Research

Key criteria for evaluating qualitative research trustworthiness include credibility, dependability, confirmability, transferability, and reflexivity (Kitto, Chesters, & Grbich, 2008; Mays & Pope, 2020; Stenfors, Kajamaa, & Bennett, 2020). To ensure validity, this study employed a narrative literature review and theoretical framework to evaluate leaders' motivational and emotional experiences in organizational change. Snowball sampling was used to collect data from 62 participants with diverse leadership backgrounds and experiences.

The research questions were clearly presented, and the theoretical framework was designed for flexibility. The data collection instrument was carefully constructed to manage the data volume, with semi-structured interviews and open-ended questions coded and verified by two leadership experts (Kitto, Chesters, & Grbich, 2008; Mays & Pope, 2020; Stenfors, Kajamaa, & Bennett, 2020).

Data was analyzed using qualitative content analysis to detect indicators of uncanny valley effects and criteria for design strategies supported by related literature and theoretical triangulation (Lincoln & Guba, 1985; Merriam, 2002).

5.2. Generalizability in Qualitative Research

This study applies the concept of "qualitative generalization" (Levitt, 2021), which emphasizes generalizing insights to the phenomena under investigation rather than to larger populations as in quantitative studies. Qualitative research offers profound insights into specific phenomena but faces skepticism regarding its potential for generalization. To address this, qualitative generalization involves inferential processes that create a variation map within the data, reflecting the complexities of the phenomena rather than the population. This approach enhances research outcomes and substantive explanations, focusing on leaders' perceptions and allowing for deeper understanding and transferability of findings (Katz, 2015; Myers, 2000).

6. Results

This chapter presents the findings from our analysis of user feedback regarding the use of AI chatbots as coaches or mentors for leaders. The analysis aims to address two primary research questions:

1) What indicators of the Uncanny Valley effect are present in the use of AI chatbots as coaches/mentors for leaders?

2) What design strategies can be employed to mitigate the Uncanny Valley effect in the use of AI chatbots as coaches/mentors for leaders?

The results derived from the data analysis are organized to provide a comprehensive understanding of the Uncanny Valley effect and its impact on user experiences. The findings for each research question are presented through detailed tables and textual analysis, highlighting key indicators and potential design improvements.

6.1. Data Presentation Research Question 1

This chapter presents the findings from our analysis of user feedback, both "not approving" (negative) and "approving" (positive), regarding the use of AI chatbots as coaches or mentors for leaders (research question 1). The aim was to identify specific indicators of the Uncanny Valley effect, which contributes to user discomfort and unease. The analysis focuses on several key indicators such as voice quality, humanness, emotional expression, social presence, and linguistic appropriateness.

The following tables provide a detailed breakdown of these indicators, highlighting the most relevant text segments (four examples) that illustrate user perceptions.

By examining both negative (not approving), see **Table 3**, and positive (approving) feedback, see **Table 4**, regarding AI-chatbot usage, we gain a nuanced understanding of the factors that contribute to the Uncanny Valley effect in AI chatbots.

Table 3. Results of "not approving" data.

Indicator	Description	Text Segments
Voice Quality	The degree to which the speech sounds human-like vs. robotic.	"The greeting was a bit wooden."/"The voice sounded artificial."/"The robotic voice was not good."/"The voice sounds mechanical."
Humanness	Degree to which users ascribe human-like qualities to chatbots.	"Nothing special, just what I expect from a chatbot."/"Negative, a small-talk AI seems inappropriate and strange."/"It doesn't sound like real empathy."/"The voice was creepy."
Emotional Expression	Chatbots' ability to express emotions appropriately.	"The computer voice sounded robotic."/"Fake empathy is still fake."/"Repeating the original question as empathy is not effective."/"Confused, slightly frustrated."

Continued

Social Presence	Perceived warmth and human-like interaction quality.	"Nothing special, just what I expect from a chatbot."/"Negative, a small-talk AI seems inappropriate and strange."/"The empathetic approach is not what I expected."/"Creepy!"
Intelligence	Perceived cognitive abilities of AI.	"The responses were longer but not comprehensive enough."/"The question was clear and concise."/"The computer voice sounded robotic."/ "Some word and grammar mistakes."
Mind Perception	Users' perception of robots having a mind.	"I tell the AI I feel insecure, and it says I have my emotions under control? The AI's answers don't seem precise."/"Fake empathy is still fake."/"The voice sounded robotic."
Linguistic Appropriateness	How well the language used matches user expectations of machine versus human speech.	"The greeting was a bit wooden."/"The voice sounded artificial."/"The question was clear and concise."/"Maybe a more conversational tone would be better."
Social Cues	The presence of politeness, relational work, and other interpersonal linguistic strategies.	"Negative, a small-talk AI seems inappropriate and strange."/"The empathetic and user-focused approach can be positive but wasn't my expectation."/"Fake empathy is still fake."
Interactional Dynamics	Responsiveness and fluidity of chatbot interactions.	"The responses were longer but not comprehensive enough."/"Creepy!"/"The computer voice sounded robotic."/"Some word and grammar mistakes."
User Partner Models	Users' mental representations of the machine's capabilities and characteristics.	"Nothing special, just what I expect from a chatbot."/"Negative, a small-talk AI seems inappropriate and strange."/"It doesn't sound like real empathy."/"Confused, slightly frustrated."
Voice Consistency	Consistency between the type of voice (synthetic vs. pre-recorded human) and the language used.	"The computer voice sounded robotic."/"The voice sounded artificial."/ "The greeting was a bit wooden."/"The voice was very bad, sounded like a robot!"

Table 4. Results of "approving" data.

Indicator	Description	Text Segments
Voice Quality	The degree to which the speech sounds human-like vs. robotic.	"The voice sounds very computerized."/"In my opinion, the simple greeting makes the AI voice sound artificial, robotic."/"I think the voice is quite robotic."/"Artificial voice and not human."
Humanness	Degree to which users ascribe human-like qualities to chatbots.	"Quick and friendly."/"A sympathetic reception."/"Uncomplicated, understandable, friendly."/"Very polite, but it seemed a bit stiff. I think it should speak more casually and change the voice to a more everyday one."
Emotional Expression	Chatbots' ability to express emotions appropriately.	"My first impression of the mentoring in the video was professional and structured."/"KI.m starts with a warm and welcoming introduction, setting a positive and professional tone for the session."/"Calming, empathetic, interested."/"Trustworthy; supportive; competent."
Social Presence	Perceived warmth and human-like interaction quality.	"It shows understanding and also that the AI is quite good."/"Friendly, interactive, supportive, appealing."/"Very human-like."/"Connected; confident; motivated; comforted."
Intelligence	Perceived cognitive abilities of AI.	"It shows understanding and also that the AI is quite good."/"The advice is particularly interesting, especially the detailed advice for the respective tips."/"The detailed solution caught my attention. Also, that it fully understood my situation."/"I think it's good because I can now choose the solution that suits me best from various suggestions."

Continued

Mind Perception	Users' perception of robots having a mind.	"It's impressive how it digs deeper and asks questions."/"The detailed solution caught my attention. Also, that it fully understood my situation."/ "I think it could be very helpful in some situations. It would be interesting to see how the AI reacts in other situations."
Linguistic Appropriateness	How well the language used matches user expectations of machine versus human speech.	"Clear and distinct."/"The question was easy and straightforward to identify."/"The question was uncomplicated and clearly recognizable."/ "The question in the text was very clear and concise, directly and specifically stated."
Social Cues	The presence of politeness, relational work, and other interpersonal linguistic strategies.	"Trustworthy; supportive; competent."/"Friendly, interactive, supportive, appealing."/"Inspired; motivated; engaged; helpful; structured; personal."/ "Intense; emotional; moving; understanding; supportive; goal-oriented."
Interactional Dynamics	Responsiveness and fluidity of chatbot interactions.	"Quick and friendly."/"Uncomplicated, understandable, friendly."/"Very polite, but it seemed a bit stiff."/"I found it clear and understandable; you knew right away what the AI was aiming for."
User Partner Models	Users' mental representations of the machine's capabilities and characteristics.	"I find the sequence for the initial classification of problems very good."/"I think it could be very helpful in some situations. It would be interesting to see how the AI reacts in other situations."/"The AI made an empathetic impression."/"I feel safe and secure, knowing I was not alone."
Voice Consistency	Consistency between the type of voice (synthetic vs. pre-recorded human) and the language used.	"The voice sounds very computerized."/"Artificial voice and not human."/ "Very polite, but it seemed a bit stiff."/"The voice was very bad, sounded like a robot!"

The key findings reveal that voice quality is a significant concern, with users frequently describing the chatbot's voice as "robotic," "artificial," and "mechanical," which creates discomfort. Users found the chatbot's attempts at humanness unconvincing and often forced, emphasizing a need for more natural and empathetic interactions. Emotional expression was seen as insincere, with many users perceiving the chatbot's attempts at empathy as "fake" and not genuinely supportive. Social presence was similarly lacking, with interactions feeling mechanical rather than warm and engaging. Additionally, linguistic appropriateness was a major issue, with the language used by the chatbot often described as stilted and inappropriate, further contributing to the overall sense of unease.

The key indicators of the Uncanny Valley effect in AI chatbots used as coaches or mentors for leaders include voice quality, humanness, emotional expression, social presence, and linguistic appropriateness. Users frequently noted that the chatbot's voice sounded robotic and artificial, highlighting the need for more natural-sounding speech. Humanness was another significant factor, with feedback indicating that interactions often felt forced and lacked genuine warmth. Emotional expression and social presence were perceived as insincere and mechanical, contributing to user discomfort. Additionally, while the chatbot's language was clear and understandable, it often felt stilted and inappropriate for natural conversation, emphasizing the need for improved linguistic appropriateness.

6.2. Research Question 2

The selected criteria for this question focus on identifying design strategies that can be implemented to reduce the Uncanny Valley effect. These strategies aim to enhance the chatbot's design and interaction features to improve user comfort and acceptance.

The data used in this analysis comes from the question "What is your suggestion for improving KI.m?" from both "not approving" (negative feedback) and "apporving" (positive feedback). By examining user suggestions and preferences, we can identify specific design improvements that can help mitigate the Uncanny Valley effect and create more effective AI chatbot interactions.

The criteria for improving AI chatbots include ensuring the voice quality sounds appropriately robotic or human-like, balancing human-like visual features (visual anthropomorphism), enhancing social presence through warmth and human-like interaction quality, increasing personalization, providing multilingual support, improving accessibility features, integrating human-like verbal cues and backstories (verbal anthropomorphism), enhancing natural language processing, ensuring fluid and responsive interactions, focusing on efficient and concise communication, adapting to ongoing situations dynamically, providing clear and direct answers, aligning AI capabilities with user expectations (user partner models), implementing mechanisms for autonomous learning and adaptation, integrating with common tools like calendars and planning software, utilizing real-time context awareness, and ensuring consistency between the type of voice and language used.

The following Table 5 provides a detailed breakdown of these strategies, highlighting the most relevant text segments (four examples) that illustrate user recommendations and preferences. By addressing these design strategies, we aim to improve the user experience and effectiveness of AI chatbots as coaches or mentors for leaders.

Table 5. Results from question: "What is your suggestion for improving KI.m?"

Indicator	Description	Text Segments
Voice Quality	Ensuring the voice sounds appropriately robotic or human-like depending on context	"I would change the voice to a less computerized one to give the employee more personality."/"Work on the pronunciation; it sounds irritating."/"Also, a male voice."/"A slightly less robotic voice."
Visual Anthropomorphism	Balancing human-like visual features to avoid excessive human-likeness	"More graphical elements, short tips, and precise questions."
Social Presence	Enhance social presence through warmth and human-like interaction quality	"The language used can change depending on the user. As I said, no informal tone ('you'), not just banal suggestions I could come up with myself, and less small talk and forced empathy, which is inappropriate for a machine."
Personalization	Increase personalization to meet user needs	"Increase personalization: Offer more customization options to meet the individual needs and situations of users."

Continued

Multilingual Support	Provide services in multiple languages	"Accessibility and Multilingual Support: Improve accessibility features to accommodate diverse user needs, including support for users with disabilities and multilingual capabilities."
Accessibility	Improve accessibility features for diverse user needs	"Accessibility and Multilingual Support: Improve accessibility features to accommodate diverse user needs, including support for users with disabilities and multilingual capabilities."
Verbal Anthropomorphism	Integrate human-like verbal cues and backstories appropriately	"Clearer answers, not by making them longer."
Natural Language Processing	Enhance the naturalness of language use	"Clearer answers, not by making them longer."
Interactional Dynamics	Ensure fluid and responsive interactions	"Work more contextually, less like a Google voice assistant."
Efficiency	Focus on efficient and concise communication	"Get to the point a bit faster. Otherwise, a very good thing."/"Shorter answers."/"More focus on efficiency, more concrete and purposeful questions."/"Communication should be more efficient, concise, and factual."
Dynamic Problem-Solving	Adapt to ongoing situations dynamically	"KI.m should monitor ongoing situations to ensure that solutions remain dynamic as they unfold."
Clear and Direct Answers	Provide clear and direct answers without unnecessary elaboration	"I would skip the initial questions about well-being and mood as they just waste time in my view."/"Clearer answers, not by making them longer."/"I would expect answers to include multiple suggestions and options, not just one."
User Partner Models	Align AI capabilities with user expectations	"I would probably try it from time to time. At the moment, it seems too generic and general to me, and above all, not constructive enough."/"I asked the AI very specific questions, but they were only answered relatively vaguely."/"In my opinion, the AI should ask for the needed information more specifically and without small talk! - to then present a very concrete answer."/"The goal-checking function would probably be more interesting to me personally than the conversation shown."
Continuous Learning and Adaptation	Implement mechanisms for autonomous learning and adaptation	"Implement mechanisms for KI.m to autonomously learn from new data sources, industry trends, and best practices in executive coaching."/"Make the AI self-learning."
Integration with Tools	Integrate with common tools like calendars and planning software	"Integration with Teams or Webex."
Real-Time Context Awareness	Save and utilize context from past interactions	"The AI should save the context if I write to it weekly, so I don't have to describe the situations every time."
Voice Consistency	Ensure consistency between the type of voice (synthetic vs. pre-recorded human) and the language used	"The voice should sound smoother; there are already so many better voices, not so robotic."

The gathered data from the question "What is your suggestion for improving KI.m?" provides valuable insights into design strategies that can mitigate the Uncanny Valley effect in AI chatbots used as coaches or mentors for leaders.

"**Voice Quality**" was frequently mentioned, with users suggesting that the voice should sound more human-like and less robotic. Improvements in pronunciation, the option for a male voice, and a more natural-sounding tone were recommended.

In terms of "**Visual Anthropomorphism**", users recommended balancing human-like visual features by incorporating more graphical elements, short tips, and precise questions. Enhancing "**Social Presence**" through warmer and more human-like interaction quality was also emphasized. Users suggested that the chatbot's language should adapt to the user, avoiding overly familiar tones and reducing small talk and artificial empathy.

"**Personalization**" was another key area, with users calling for more customization options to meet individual needs. The importance of "**Multilingual Support**" and "**Accessibility**" was highlighted, with suggestions to provide services in multiple languages and improve accessibility features for diverse user needs.

For "**Verbal Anthropomorphism**", integrating human-like verbal cues and backstories appropriately was recommended to improve interaction quality. Enhancing the naturalness of language use through "**Natural Language Processing**" was seen as critical, with users requesting clearer and more concise responses.

Ensuring fluid and responsive "**Interactional Dynamics**" was important, with users suggesting more context-aware interactions. "**Efficiency**" was also emphasized, with a desire for quicker, more direct communication. Users recommended that the chatbot should adapt to ongoing situations dynamically "**Dynamic Problem-Solving**" and provide clear and direct answers without unnecessary elaboration.

Aligning AI capabilities with user expectations "**User Partner Models**" was necessary, with feedback calling for more specific and constructive responses. Implementing mechanisms for "**Continuous Learning and Adaptation**" was suggested to enhance the chatbot's capabilities over time.

Integration with common tools like calendars and planning software "**Integration with Tools**" was recommended to enhance user experience. Saving and utilizing context from past interactions "**Real-Time Context Awareness**" was important to avoid repetitive explanations. Finally, ensuring consistency between the type of voice and the language used "**Voice Consistency**" was emphasized, with users requesting more fluid and natural-sounding voices.

In addition to these suggestions, several passages from the "suggestions for improvement" were not included in the analysis because they could not be clearly assigned to one of the criteria defined in the codebook. These passages either state that no concrete suggestions for improvement can be made, that the tool is already considered good enough, or contain general statements that do not concern specific points of improvement.

Examples include statements such as "I don't know, it looked pretty perfect", "Honestly, I don't really know what else an AI could do", and "You would have to spend more time with the chatbot for concrete suggestions for improvement".

Some users also expressed satisfaction with the current state of the tool, saying "I have nothing to add" or "It seemed very perfect, so I can't suggest anything to you".

6.3. Significant Findings

This section presents the key findings from the study, addressing both research questions. The analysis aims to identify the indicators of the Uncanny Valley effect in the use of AI chatbots as coaches or mentors for leaders and explore design strategies that can mitigate this effect. The results from the data analysis provide a comprehensive understanding of the factors contributing to user discomfort and the potential improvements needed to enhance user experience.

6.3.1. Research Question 1

This section addresses Research Question 1 by presenting the key findings derived from the analysis of user feedback. Both "not approving" (negative) and "approving" (positive) responses were examined to identify specific indicators of the Uncanny Valley effect in AI chatbots used as coaches or mentors for leaders. By understanding these indicators, we can better comprehend the factors contributing to user discomfort and the areas needing improvement to enhance the interaction experience.

Key Findings—"Not approving"

The research aimed to identify significant indications of the Uncanny Valley effect in the use of AI chatbots as coaches and mentors for leaders. Based on user feedback, the following indicators were identified as the most significant contributors to this phenomenon:

"Voice Quality": The degree to which the chatbot's speech sounds human-like versus robotic emerged as a major concern. Users frequently described the voice of the chatbot with terms such as "mechanical," "robotic," "artificial," and "unnatural." Specific feedback included comments highlighting significant discomfort due to the poor quality of the voice, such as "The voice unfortunately sounded artificial" and "The voice is very bad ... sounds like a robot!! Deal breaker."

"Emotional Expression": The chatbots' ability to express emotions appropriately was another key factor. Users perceived the chatbot's emotional expressions as forced and insincere, which detracted from the overall experience. Phrases like "Fake empathy is just fake" and "Empathy is simulated by repeating the initial question, which is not very effective" indicate a negative reaction to the chatbot's attempts to simulate human emotions.

"Social Presence": The perceived warmth and quality of human-like interaction were also significant indicators. Despite efforts to create warm and engaging interactions, users felt the chatbot's presence was more mechanical than human. Comments such as "negative, as a small-talk driven AI seems inappropriate and strange to me" reveal how the lack of genuine social presence resulted in discomfort.

"Humanness": The degree to which users ascribe human-like qualities to chatbots was another crucial aspect. Users often found the chatbot's attempts at human-like interaction unconvincing. Statements like "nothing special, what I expect from a chatbot" and "The very empathetic and user-focused approach can certainly be positive, but this is not the expectation I had of an advisor" suggest that the chatbot's efforts to mimic human qualities were perceived as unnatural and forced.

"Linguistic Appropriateness": How well the language used by the chatbot matches user expectations of machine versus human speech also played a significant role. Users noted that the chatbot's language and phrasing were sometimes stilted and inappropriate for the context. Comments like "The greeting was somewhat wooden" and "Poor pronunciation" indicate that inconsistencies in linguistic appropriateness contributed to a sense of unease.

These key findings highlight the significant contributors to the Uncanny Valley effect when using AI chatbots as coaches and mentors for leaders, emphasizing areas for improvement in voice quality, emotional expression, social presence, humanness, and linguistic appropriateness to enhance user experience.

Key Findings—"Approving": The analysis of the "approving" (positive) data segments provides crucial insights into the specific indicators that significantly contribute to the Uncanny Valley effect in AI chatbots used as coaches or mentors for leaders. These indicators are essential in understanding the design and interaction aspects that lead to user discomfort and unease. The most frequently mentioned text segments and their associated indicators highlight the primary areas that need improvement to mitigate the Uncanny Valley effect.

"**Voice Quality**" emerged as a significant factor. Users frequently highlighted the robotic, mechanical, and unnatural qualities of the chatbot's voice with comments such as "The voice sounds very computerized" and "artificial voice and not human". These descriptions indicate significant discomfort stemming from the non-human-like voice, which fails to meet users' expectations of natural interaction.

"**Humanness**" was another critical indicator. Users noted that the chatbot's attempts at human-like interaction often felt awkward and forced. Feedback such as "Very polite, but it seemed a bit stiff" and "friendly, interactive, supportive, appealing" pointed to a stiffness in its responses that detracted from the overall experience, making interactions feel unnatural and strained.

"**Emotional Expression**" also contributed to the Uncanny Valley effect. The chatbot's emotional expressions were often seen as forced and insincere. Despite attempts to show empathy and support, users found these expressions unconvincing, as indicated by comments like "The AI made an empathetic impression" and "Trustworthy; supportive; competent". This lack of genuine emotional depth heightened user discomfort.

"**Social Presence**" was generally perceived as lacking genuine warmth and human touch. Users found the conversational style inappropriate and awkward, with feedback such as "friendly, interactive, supportive, appealing" and "very polite, but it seemed a bit stiff" emphasizing the perceived mechanical nature of interactions, which exacerbated the Uncanny Valley effect.

"**Intelligence**" was acknowledged, but inconsistencies in cognitive abilities and language errors diminished trust. Comments like "The advice is particularly interesting, especially the detailed advice for the respective tips" and "The detailed solution caught my attention. Also, that it fully understood my situation" showed that while users recognized the chatbot's ability to provide relevant advice, the inconsistencies aligned with the Uncanny Valley effect where human-like reasoning was expected but not fully realized.

"**Mind Perception**" played a role in user discomfort as well. Users perceived the chatbot as having deeper cognitive abilities, which sometimes felt unsettling. Statements such as "It's impressive how it digs deeper and asks questions" and "The detailed solution caught my attention. Also, that it fully understood my situation" suggested that the chatbot's responses indicated a level of understanding that felt artificial, contributing to the Uncanny Valley effect.

"**Linguistic Appropriateness**" was another area of concern. The language used by the chatbot was clear and understandable but often felt stilted and inappropriate for natural human conversation. Feedback like "The question was uncomplicated and clearly recognizable" and "The question in the text was very clear and concise, directly and specifically stated" highlighted how this mismatch in linguistic appropriateness contributed to user unease and the Uncanny Valley effect.

"**Social Cues**" were perceived as polite and considerate but lacked genuine relational work. Users noted that these social cues felt mechanical and rehearsed, as reflected in comments like "Trustworthy; supportive; competent" and "friendly, interactive, supportive, appealing". This mechanistic nature of social cues enhanced the Uncanny Valley effect.

"**Interactional Dynamics**" were often perceived as unnatural and forced. Feedback such as "uncomplicated, understandable, friendly" and "Very polite, but it seemed a bit stiff" indicated a lack of fluidity and responsiveness in conversational dynamics, further contributing to the Uncanny Valley effect.

"**Medium and Context of Interaction**" received generally positive feedback regarding the interface's user-friendliness. Comments like "The interface seems user-friendly" and "It was easy to use" suggested a good reception of the medium and context. However, this did not significantly mitigate the Uncanny Valley effect.

"**User Partner Models**" revealed mixed expectations of the chatbot's capabilities. While some users found it helpful, others noted its limitations. Statements such as "The AI made an empathetic impression" and "I think it could be very helpful in some situations. It would be interesting to see how the AI reacts in other situations" indicated that the perception of artificial empathy and inconsistency in performance contributed to the Uncanny Valley effect.

"**Voice Consistency**" highlighted further inconsistencies that exacerbated user discomfort. Comments like "The voice sounds very computerized", "artificial voice and not human", and "Very polite, but it seemed a bit stiff" underscored the artificial nature of the interaction, significantly contributing to the Uncanny Valley effect.

These key findings underscore the areas that need improvement to enhance user experience and mitigate the Uncanny Valley effect in AI chatbots used as coaches or mentors for leaders.

6.3.2. Research Question 2

The gathered data from the question "What is your suggestion for improving KI.m?" provides valuable insights into design strategies that can mitigate the Uncanny Valley effect in AI chatbots used as coaches or mentors for leaders.

"**Voice Quality**" was frequently mentioned, with users suggesting that the voice should sound more human-like and less robotic. Improvements in pronunciation, the option for a male voice, and a more natural-sounding tone were recommended. Users highlighted the need for a "less computerized voice" and noted that "the voice should sound smoother".

In terms of "**Visual Anthropomorphism**", users recommended balancing human-like visual features by incorporating more graphical elements, short tips, and precise questions. Enhancing Social Presence through warmer and more human-like interaction quality was also emphasized. Users suggested that the chatbot's language should adapt to the user, avoiding overly familiar tones and reducing small talk and artificial empathy.

"**Personalization**" was another key area, with users calling for more customization options to meet individual needs. The importance of Multilingual Support and Accessibility was highlighted, with suggestions to provide services in multiple languages and improve accessibility features for diverse user needs.

For "**Verbal Anthropomorphism**", integrating human-like verbal cues and backstories appropriately was recommended to improve interaction quality. Enhancing the naturalness of language use through Natural Language Processing was seen as critical, with users requesting clearer and more concise responses.

Ensuring fluid and responsive "**Interactional Dynamic**" was important, with users suggesting more context-aware interactions. Efficiency was also emphasized, with a desire for quicker, more direct communication. Users recommended that the chatbot should adapt to ongoing situations dynamically "**Dynamic Problem-Solving**" and provide clear and direct answers without unnecessary elaboration.

Aligning AI capabilities with user expectations "**User Partner Models**" was necessary, with feedback calling for more specific and constructive responses. Implementing mechanisms for Continuous Learning and Adaptation was suggested to enhance the chatbot's capabilities over time.

Integration with common tools like calendars and planning software "**Inte-**

gration with Tools" was recommended to enhance user experience. Saving and utilizing context from past interactions "**Real-Time Context Awareness**" was important to avoid repetitive explanations. Finally, ensuring consistency between the type of voice and the language used "Voice Consistency" was emphasized, with users requesting more fluid and natural-sounding voices.

In addition to these suggestions, several passages from the "Suggestions for improvement" were not included in the analysis because they could not be clearly assigned to one of the criteria defined in the codebook. These passages either state that no concrete suggestions for improvement can be made, that the tool is already considered good enough, or contain general statements that do not concern specific points of improvement. Examples include statements such as "I don't know, it looked pretty perfect", "Honestly, I don't really know what else an AI could do", and "You would have to spend more time with the chatbot for concrete suggestions for improvement". Some users also expressed satisfaction with the current state of the tool, saying "I have nothing to add" or "It seemed very perfect, so I can't suggest anything to you".

7. Discussion

7.1. Interpretation of Results

7.1.1. Research Question 1

The research aimed to identify observable indicators of the Uncanny Valley effect in AI chatbots used as coaches or mentors for leaders. Analysis of both negative and positive feedback revealed several key areas contributing to user discomfort and unease, which are essential to address to enhance user experience.

"**Voice Quality**" emerged as a significant concern, with users frequently describing the chatbot's voice as "mechanical", "robotic", "artificial", and "unnatural". Feedback underscored the need for a more natural and human-like voice to mitigate discomfort. "**Emotional Expression**" also played a critical role, as users found the chatbot's attempts at showing empathy forced and insincere, indicating that the lack of genuine emotional depth significantly contributed to their discomfort.

"**Social Presence**" was another crucial factor, with users perceiving the chatbot's presence as more mechanical than human, lacking genuine warmth and human touch. This mechanical nature of interactions exacerbated the Uncanny Valley effect. Similarly, users found the chatbot's attempts at human-like interaction unconvincing and forced, emphasizing the need for more natural and less strained interactions to improve the user experience.

"**Linguistic Appropriateness**" was a concern, with users noting that the chatbot's language often felt stilted and inappropriate for natural human conversation, contributing to user unease. While users recognized the chatbot's ability to provide relevant advice, inconsistencies in cognitive abilities and language errors diminished trust, aligning with the Uncanny Valley effect. Additionally, users sometimes found the chatbot's deeper cognitive abilities unsettling, with responses that felt artificial despite showing a level of understanding.

"**Social Cues**" were seen as polite but lacking genuine relational work, which enhanced the Uncanny Valley effect due to their mechanical and rehearsed nature. Interactional dynamics were perceived as unnatural and forced, with a lack of fluidity and responsiveness in conversational dynamics further contributing to the Uncanny Valley effect.

Despite generally positive feedback on the user-friendliness of the interface, this did not significantly mitigate the Uncanny Valley effect. "**User Partner Models**" revealed mixed expectations of the chatbot's capabilities, with some users finding it helpful while others noted its limitations. The perception of artificial empathy and inconsistency in performance contributed to the Uncanny Valley effect. Lastly, "**Voice Consistency**" further exacerbated user discomfort, with inconsistencies in voice quality underscoring the artificial nature of the interaction.

7.1.2. Research Question 2

The findings suggest several design strategies to mitigate the Uncanny Valley effect in AI chatbots used as coaches or mentors for leaders. A critical design strategy is balancing human-like and machine-like voice quality. This involves refining pronunciation and intonation while selecting voice types that do not overly mimic human speech, creating a voice that balances warmth and precision to make chatbots more reliable and less unsettling.

Incorporating human-like visual features, such as graphical elements and concise tips, without making the chatbot too human-like, can maintain user comfort. This includes designing visually appealing, user-friendly interfaces that are distinctly non-human to avoid discomfort from overly anthropomorphic designs. Improving social presence by fostering warm and engaging interactions is crucial. However, this should avoid making the AI seem too human. The design should focus on professional and supportive communication styles, avoiding excessive small talk and ensuring subtle, contextually appropriate expressions of empathy.

Increasing personalization allows chatbots to adapt their interactions to individual user preferences through customizable settings. This ensures that users feel the chatbot is responsive without pretending to have human-like understanding or empathy. Ensuring the chatbot can communicate in multiple languages and is accessible to diverse user needs is fundamental. This includes integrating robust language support and accessibility options like text-to-speech and speech-to-text capabilities, making the chatbot inclusive and enhancing user satisfaction.

Improving natural language processing (NLP) capabilities for more natural and fluid conversations is important, but these conversations should remain efficient and to the point. Subtle verbal anthropomorphism, like appropriate colloquial language use or contextually relevant backstories, can make interactions engaging without being too human-like. Streamlining interactions to be fluid and responsive ensures a positive user experience. Designing the chatbot to provide concise, clear answers and adapt dynamically to ongoing conversations focuses on efficiency and fulfills user expectations while maintaining a machine-like identity. Integrating the chatbot with common tools like calendars and planning software enhances practical utility. Implementing real-time context awareness allows the chatbot to remember past interactions, reducing repetitive explanations and providing a seamless user experience. This emphasizes the chatbot as a supportive tool rather than a human-like companion.

Finally, implementing mechanisms for continuous learning ensures the chatbot evolves and improves over time. Using machine learning to refine responses based on user feedback and emerging trends enhances effectiveness while maintaining a balance between human-like and machine-like qualities.

7.2. Comparison with Existing Literature

The Uncanny Valley effect, where users feel discomfort due to near-human-like qualities of AI, is a significant challenge in AI chatbot development. This analysis compares user feedback from AI chatbots used as mentors or coaches with existing studies on the Uncanny Valley effect, identifying consistencies, differences, and gaps.

Both studies and user feedback highlight "**Voice Quality**" as a critical factor. Users described the chatbot's voice as "robotic", "artificial", and "mechanical", aligning with findings that human-like voices can moderate but not eliminate the Uncanny Valley effect (Clark, Ofemile, & Cowan, 2021). The feedback emphasizes the need for more natural-sounding voices to enhance user comfort.

While existing studies address "**Visual Anthropomorphism**" (Cui, Wang, & Qi, 2021), this was not a significant factor in user feedback, indicating a gap where further exploration might be needed regarding visual appearance and its impact on the Uncanny Valley in chatbots.

"**Emotional Expression**" played a critical role, as users found the chatbot's attempts at showing empathy forced and insincere, which significantly contributed to their discomfort. This corroborates the findings of Rapp, Curti, and Boldi (2021) that inappropriate or mismatched emotional expressions can trigger discomfort.

Findings on "**Social Presence**" being a double-edged sword are echoed in user feedback, which described interactions as "mechanical" rather than warm (Mozafari, Hammerschmidt, & Weiger, 2021). This suggests a need for enhancing the human-like interaction quality to make social presence feel more genuine and less discomforting.

Research on the perceived intelligence of AI creating unease if not perfect matches user feedback highlighting insufficient and imprecise responses (Stepp Jr, 2022). Users expect more concrete and helpful suggestions, indicating a need for improving AI's perceived cognitive abilities to align with user expectations and reduce discomfort.

The study on mind perception causing discomfort if robots appear too sentient aligns with user feedback (Mathur & Reichling, 2020). Users found the chatbot's responses often lacked depth, reinforcing the importance of balancing perceived intelligence to avoid the Uncanny Valley.

Research on "**Linguistic Appropriateness**" matches user feedback, which often found the chatbot's language stilted and inappropriate (Clark, Ofemile, & Cowan, 2021). Ensuring that language use aligns with user expectations is essential to avoid creating a perceptual mismatch that leads to discomfort.

Emphasis on "**Verbal Anthropomorphism**" enhancing user engagement but potentially contributing to the Uncanny Valley if overdone aligns with the feedback (Cui, Wang, & Qi, 2021). Users appreciated clear and direct answers, indicating a preference for minimalistic and appropriate verbal cues.

Findings on "**Social Cues**" and "**Interaction Dynamics**" causing discomfort if perceived as inappropriate are reflected in the feedback (Clark, Ofemile, & Cowan, 2021). Users noted that the chatbot's attempts at empathy and social interaction often felt forced, emphasizing the need for natural and fluid conversation flows.

Feedback on user interface and context, such as suggestions for integration with common tools and context awareness, highlights the importance of aligning the chatbot's capabilities with user expectations (Clark, Ofemile, & Cowan, 2021). Ensuring seamless integration and context awareness can help mitigate the Uncanny Valley effect.

Aligning AI capabilities with user expectations is crucial. Feedback indicates that users found the chatbot too generic and not tailored enough, reflecting the need for personalized interactions to meet individual user needs effectively.

The importance of consistency between voice type and language used is highlighted (Clark, Ofemile, & Cowan, 2021), which is echoed in user feedback. Users emphasized the need for a more natural and consistent voice to reduce discomfort.

User suggestions for improving chatbots include enhancing voice quality, balancing visual anthropomorphism, increasing social presence, and improving personalization. These strategies align with the criteria outlined in existing studies, indicating that addressing these areas can effectively reduce the Uncanny Valley effect.

The degree of **humanness** ascribed to the chatbots was another critical issue. Users often found the chatbot's human-like attempts unconvincing and forced. This aligns with Rapp, Curti, and Boldi's (2021) assertion that higher humanness can increase perceived eeriness if not perfectly human-like.

7.3. Implications

7.3.1. Theoretical Implications

The research aimed to identify observable indicators of the Uncanny Valley effect in AI chatbots used as coaches or mentors for leaders. Findings from both negative and positive feedback highlight several critical areas contributing to user discomfort and unease. These findings have several theoretical implications:

1) Refinement of the Uncanny Valley Theory: The study reinforces and refines the existing theoretical framework by identifying specific chatbot attributes—such as voice quality, emotional expression, and social presence—that significantly influence user discomfort. These insights help delineate the boundary conditions of the Uncanny Valley effect more clearly, suggesting the theory needs more nuanced guidelines on threshold levels of these attributes.

2) Emotional Expression and AI: The findings corroborate the theoretical stance that mismatched or inappropriate emotional expressions by AI can trigger discomfort (Rapp, Curti, & Boldi, 2021). This suggests the theory should include guidelines for designing AI with emotional expression capabilities that are contextually appropriate and perceived as genuine by users, accounting for the complexity of human emotions.

3) Social Presence and Engagement: The research supports the idea that social presence is a double-edged sword in AI interactions (Mozafari, Hammerschmidt, & Weiger, 2021). The theoretical framework should consider the balance between warmth and professionalism in AI interactions to avoid making the AI seem overly human, which can lead to unease. This requires exploring the social cues and relational dynamics that AI should exhibit to foster trust and engagement without crossing into the uncanny territory.

4) Cognitive Abilities and User Perception: The findings highlight the importance of aligning AI's cognitive abilities with user expectations. The theory needs to address how perceived intelligence and mind perception contribute to the Uncanny Valley effect, emphasizing the need for AI to exhibit appropriate levels of understanding and cognitive capabilities without appearing too sentient (Mathur & Reichling, 2020).

5) Linguistic and Interactional Dynamics: The results suggest that linguistic appropriateness and fluid conversational dynamics are crucial for reducing user discomfort. The theoretical framework should integrate these aspects, emphasizing the need for natural and contextually appropriate language use in AI chatbots (Clark, Ofemile, & Cowan, 2021). It should explore how different linguistic styles and interactional strategies impact user perceptions and comfort levels.

7.3.2. Practical Implications

The practical implications of these findings are significant for the design and development of AI chatbots used as coaches or mentors for leaders. Key strategies to mitigate the Uncanny Valley effect include:

1) Enhancing Voice Quality: Developers should create more natural and human-like voices for chatbots by refining pronunciation and intonation. Implementing advanced text-to-speech technologies can help bridge the gap between machine and human interaction, balancing warmth and precision in the chatbot's voice.

2) Improving Emotional Expression: AI chatbots should express emotions appropriately, avoiding forced or insincere empathy. Programming AI to respond with contextually relevant and genuine emotional expressions can enhance the perceived naturalness of interactions. Developing algorithms that better understand and replicate human emotions will make emotional expressions more convincing.

3) Balancing Social Presence: Enhancing social presence through warm and engaging interactions is crucial, but AI should not appear overly human. Fostering professional and supportive communication styles, avoiding excessive small talk, and ensuring subtle empathy can make AI seem more genuine and less mechanical.

4) Personalization and Accessibility: Increasing personalization through customizable settings allows chatbots to adapt to individual user preferences. Providing multilingual support and accessibility options like text-to-speech and speech-to-text capabilities can make the chatbot more inclusive and enhance user satisfaction. AI systems should learn from user interactions to tailor responses more effectively.

5) Natural Language Processing and Verbal Anthropomorphism: Improving NLP capabilities for more natural and fluid conversations is important. Subtle verbal anthropomorphism, such as using appropriate colloquial language and contextually relevant backstories, can make interactions engaging without being too human-like. Investing in NLP technologies that understand and generate human-like text while maintaining appropriate anthropomorphism can improve user engagement.

6) Streamlined Interactions and Practical Integration: Designing chatbots to provide concise, clear answers and adapt dynamically to ongoing conversations ensures a positive user experience. Integrating the chatbot with common tools like calendars and planning software enhances its practical utility, while real-time context awareness provides a seamless user experience. This integration ensures the chatbot remains a useful tool rather than an unsettling companion.

7) Continuous Learning and Adaptation: Implementing mechanisms for continuous learning ensures that the chatbot evolves and improves over time. Using machine learning to refine responses based on user feedback and emerging trends can enhance the effectiveness of the chatbot while maintaining a consistent machine-like persona. Continuous improvement processes should regularly update the chatbot's capabilities based on user interactions and feedback.

These practical implications extend beyond immediate design improvements to a broader strategy for AI development. Enhancing voice quality and emotional expression are foundational to reducing user discomfort, as they address the most visceral aspects of human-computer interaction. The refinement of emotional expression in AI is challenging but necessary, requiring a nuanced approach where chatbots express appropriate levels of empathy and understanding based on context.

Balancing social presence and engagement is crucial, aiming for interactions that are supportive and professional while avoiding excessive human-like behaviors. Personalization and accessibility ensure AI chatbots meet diverse user needs, with customizable settings and robust language support enhancing user satisfaction.

Improving NLP capabilities is essential for creating fluid and natural interactions. Chatbots must understand and generate language that feels natural to users, using subtle anthropomorphic cues without overstepping into the uncanny. Streamlining interactions to be clear, concise, and dynamically adaptive enhances the practical utility of chatbots.

Continuous learning and adaptation are fundamental for long-term success. Implementing machine learning algorithms to refine chatbot responses based on user feedback ensures that AI evolves to meet user expectations effectively, maintaining a consistent and reliable machine-like persona crucial for reducing the Uncanny Valley effect.

7.4. Limitations

This study has certain limitations that need to be acknowledged to provide a comprehensive understanding of the findings and their applicability. One key limitation is the sample size and composition. With 62 participants, the sample may not fully represent the diversity of leadership experiences and perspectives. The gender distribution, predominantly male (69.4%), and age concentration between 31 and 40 years (48.4%), may also limit the generalizability of the results to broader leadership populations.

Another limitation is the reliance on self-reported data, which can introduce biases such as social desirability bias or inaccurate self-assessment. Participants' responses might be influenced by their perceptions of the study's purpose or their expectations of desirable traits and behaviors in leaders.

The study's design, utilizing a semi-structured online survey and qualitative content analysis, while providing rich, in-depth insights, may present challenges in terms of replicability and objectivity. The subjective nature of qualitative analysis and potential variability in participants' interpretations of the questions and video excerpts could affect the consistency of the findings.

Additionally, the focus on an AI-based chatbot for leader's support and mentoring, while innovative, may have inherent limitations due to the technology's current developmental stage. Participants' familiarity with and attitudes toward AI technology could vary widely, influencing their feedback and potentially introducing bias towards either optimism or skepticism about AI capabilities.

The study's scope, centered on perceptions of a specific AI tool, may not capture the full spectrum of potential applications and challenges associated with AI in leadership contexts. The results are specific to the functionalities and design of KI.m and may not be applicable to other AI systems with different features and interaction models.

Furthermore, the context of use and situational factors during the survey (e.g., the environment in which participants engaged with the online survey and videos) could impact the responses. Variations in the setting, distractions, and participants' mental state at the time of survey completion might affect their perceptions and feedback.

Lastly, while the study compares its findings with existing literature on the Uncanny Valley effect and AI chatbot interactions, the rapidly evolving nature of AI technology means these comparisons may quickly become outdated. Continuous advancements in AI could alter user perceptions and the relevance of current theoretical frameworks.

In conclusion, while this study provides valuable insights into leaders' perceptions of an AI-based mentoring tool, the limitations outlined here suggest that further research with more diverse samples, varied methodologies, and consideration of technological advancements is needed to validate and extend these findings.

7.5. Future Research Directions

Future research should build on this study's findings to further explore AI-based mentoring tools and their impact on leadership development. One key area is examining more diverse and larger sample sizes to ensure findings are generalizable across different demographics and leadership contexts. Including a more balanced gender distribution and a wider age range will help understand how various groups perceive and interact with AI mentoring tools.

Investigating the longitudinal effects of AI mentoring is crucial. Understanding how prolonged use of AI mentors influences leadership skills, decision-making, and overall performance over time is essential. Longitudinal studies can identify long-term benefits, potential drawbacks, and changes in user perceptions as they become more accustomed to AI interactions.

The impact of cultural differences on the acceptance and effectiveness of AI mentoring tools is another valuable research avenue. Exploring these differences can provide deeper insights into customizing AI tools for global use.

Further research should delve into specific AI attributes that affect user comfort and effectiveness. Investigating voice quality, emotional expression, and social presence can lead to more user-friendly AI designs. Experimental studies can offer concrete guidelines for AI development.

Future studies could also examine integrating AI mentoring tools with other technologies used in professional settings. Understanding how AI can blend into existing workflows and enhance productivity is critical for its widespread adoption.

Exploring the ethical implications of AI in leadership mentoring is another important research direction. Issues such as data privacy, potential bias in AI algorithms, and transparency of AI decision-making should be thoroughly examined to ensure ethical standards and build user trust.

Finally, future research should investigate the customization and personalization of AI mentoring tools. Developing adaptive AI systems that tailor interactions based on user preferences, learning styles, and leadership challenges can enhance their effectiveness and user satisfaction. Research can focus on the algorithms and techniques necessary to achieve high levels of personalization.

8. Conclusion

This study provides a comprehensive analysis of the Uncanny Valley effect in AI chatbots used for leadership mentoring, focusing on how this phenomenon influences user perceptions and exploring potential strategies to mitigate it. Through a detailed examination of both positive and negative user feedback, several key findings were identified that highlight the critical factors contributing to user discomfort and unease when interacting with AI chatbots designed for executive support.

One of the primary sources of discomfort was the robotic and artificial nature of the chatbot's voice. Users consistently expressed a preference for more natural and human-like voices, indicating that improvements in voice synthesis are crucial for reducing the Uncanny Valley effect. Additionally, the chatbot's emotional expressions were often perceived as forced and insincere, further contributing to user unease. Participants emphasized the need for more contextually appropriate and genuine emotional responses from the chatbot to improve the overall interaction experience.

The mechanical and unnatural social interactions provided by the chatbot also diminished user engagement. A balance between professional support and subtle empathy is necessary to enhance the chatbot's social presence without triggering the Uncanny Valley effect. Moreover, while the chatbot's language was clear, it often felt stilted and inappropriate for natural human conversation. Enhancing the naturalness and appropriateness of language use is essential for improving user experience.

Furthermore, users highlighted the importance of personalization and accessibility features, including multilingual support and customizable settings. These enhancements can significantly improve the practical utility and user satisfaction of AI chatbots in leadership mentoring contexts.

The findings of this study underscore the importance of refining AI chatbots' design to align more closely with user expectations, particularly in areas related to voice quality, emotional expression, and social interaction. As AI technology continues to evolve, future research should focus on exploring long-term impacts, cultural differences, and ethical considerations in the use of AI for leadership mentoring. By addressing these key discomforts, AI mentoring tools can be more effectively integrated into leadership development strategies, ultimately supporting the growth and success of leaders in a rapidly changing business environment.

Conflicts of Interest

The author declares no conflicts of interest regarding the publication of this paper.

References

Abasaheb, S. A., & Subashini, R. (2023). Maneuvering of Digital Transformation: Role of Artificial Intelligence in Empowering Leadership—An Empirical Overview. *International Journal of Professional Business Review, 8,* e01838. https://typeset.io/papers/maneuvering-of-digital-transformation-role-of-artificial-2vl1j0hu?utm_source=chatgpt

Aktaş, B. N. (2024). Artificial Intelligence: A New Paradigm for Leadersh. In F. Mızrak (Ed.), *Trends, Challenges, and Practices in Contemporary Strategic Management* (PP. 33-50). IGI Global. https://doi.org/10.4018/979-8-3693-1155-4.ch002

An, M., Lin, J., & Luo, X. (2024). The Impact of Human AI Skills on Organizational Innovation: The Moderating Role of Digital Organizational Culture. *Journal of Business Research, 182,* Article ID: 114786. https://doi.org/10.1016/j.jbusres.2024.114786

Arakawa, R., & Yakura, H. (2024). Coaching Copilot: Blended Form of an LLM-Powered Chatbot and a Human Coach to Effectively Support Self-Reflection for Leadership Growth. In *ACM Conversational User Interfaces 2024* (pp. 1-14). ACM. https://doi.org/10.1145/3640794.3665549

Asfahani, A. M. (2022). The Impact of Artificial Intelligence on Industrial-Organizational Psychology: A Systematic Review. *The Journal of Behavioral Science, 17,* 125-139. https://www.researchgate.net/publication/364421222_The_Impact_of_Artificial_Intelligence_on_Industrial-Organizational_Psychology_A_Systematic_Review

Bankins, S., Ocampo, A. C., Marrone, M., Restubog, S. L. D., & Woo, S. E. (2024). A Multilevel Review of Artificial Intelligence in Organizations: Implications for Organizational Behavior Research and Practice. *Journal of Organizational Behavior, 45,* 159-182. https://doi.org/10.1002/job.2735

Banks, G. C., Dionne, S. D., Mast, M. S., & Sayama, H. (2022). Leadership in the Digital Era: A Review of Who, What, When, Where, and Why. *The Leadership Quarterly, 33,* Article ID: 101634. https://doi.org/10.1016/j.leaqua.2022.101634

Barriball, K. L., & While, A. (1994). Collecting Data Using a Semi-Structured Interview: A Discussion Paper. *Journal of Advanced Nursing, 19,* 328-335. https://doi.org/10.1111/j.1365-2648.1994.tb01088.x

Bengtsson, M. (2016). How to Plan and Perform a Qualitative Study Using Content Analysis. *NursingPlus Open, 2,* 8-14. https://doi.org/10.1016/j.npls.2016.01.001

Betriana, F., Osaka, K., Matsumoto, K., Tanioka, T., & Locsin, R. C. (2021). Relating Mori's Uncanny Valley in Generating Conversations with Artificial Affective Communication and Natural Language Processing. *Nursing Philosophy, 22,* e12322. https://doi.org/10.1111/nup.12322

Blut, M., Wang, C., & Schoefer, K. (2021). This Meta-Analysis Clarifies That the Impact of Anthropomorphism on Customer Intentions to Use Service Robots Depends on Robot and Service Types. *Journal of Service Research, 24,* 3-20.

Brown, P., Levinson, S. C., & Gumperz, J. J. (1987). *Politeness: Some Universals in Language Usage.* Cambridge University Press. https://doi.org/10.1017/cbo9780511813085

Bryman, A. (2017). *Social Research Methods* (5th ed.). Oxford University Press.

Bryman, A., Bresnen, M., Beardsworth, A., & Keil, T. (1988). Qualitative Research and the Study of Leadership. *Human Relations, 41,* 13-29. https://doi.org/10.1177/001872678804100102

Bryman, A., Stephens, M., & Campo, C. (1996). The Importance of Context: Qualitative Research and the Study of Leadership. *Leadership & Organization Development Journal, 17,* 36-45.

Burgess, T. F. (2001). *Guide to the Design of Questionnaires.* University of Leeds.

Clark, L., Ofemile, A., & Cowan, B. R. (2021). Exploring Verbal Uncanny Valley Effects with Vague Language in Computer Speech. In B. Weiss, J. Trouvain, M. Barkat-Defradas, & J. J. Ohala (Eds.), *Voice Attractiveness* (pp. 317-330). Springer. https://doi.org/10.1007/978-981-15-6627-1_17

Clark, L., Ofemile, A., Adolphs, S., & Rodden, T. (2016). A Multimodal Approach to Assessing User Experiences with Agent Helpers. *ACM Transactions on Interactive Intelligent Systems, 6,* 1-31. https://doi.org/10.1145/2983926

Conger, J. A. (1998). Qualitative Research as the Cornerstone Methodology for Understanding Leadership. *The Leadership Quarterly, 9,* 107-121. https://doi.org/10.1016/s1048-9843(98)90044-3

Cowan, B. R., Branigan, H. P., Obregón, M., Bugis, E., & Beale, R. (2015). Voice Anthropomorphism, Interlocutor Modelling and Alignment Effects on Syntactic Choices in Human−computer Dialogue. *International Journal of Human-Computer Studies, 83,* 27-42. https://doi.org/10.1016/j.ijhcs.2015.05.008

Creasey, M. C., & Vázquez Anido, A. (2020). *Virtual Influencing: Uncharted Frontier in the Uncanny Valley.* Lund University. https://lup.lub.lu.se/luur/download?func=downloadFile&recordOId=9015731&fileOId=9015749

Cui, T., Wang, X., & Qi, J. (2021). *Designing Anthropomorphic Therapeutic Conversational Agents: An Uncanny Valley Perspective.* ECIS.

Davenport, T. H., & Foutty, J. (2022). *AI-Driven Leadership.* MIT Sloan Management Review. https://sloanreview.mit.edu/article/ai-driven-leadership/

Diel, V., MacDorman, K. F., & Bhatt, S. (2022). This Meta-Analysis Provides a Comprehensive Model of Factors Influencing the Uncanny Valley Effect. *Journal of Robotics Research, 41,* 723-742.

Doğan Erdinç, E., & Uzunçarşılı Soydaş, A. (2024). Tekinsiz Vadi Teorisi Bağlamında Yapay Zeka Etkileyicileri. *Selçuk İletişim, 17*, 1-38. https://doi.org/10.18094/josc.1390778

Ekkelenkamp, J. (2021). *The Uncanny Valley: The Effects of Human-Likeness and Gender of E-Commerce Chatbots on Perceived Expertise and Recommendation Adherence.* Master's Thesis, Tilburg University.

Fenzl, T., & Mayring, P. (2017). Qualitative Content Analysis. In U. Flick (Ed.), *The SAGE Handbook of Qualitative Data Analysis* (pp. 214-228). SAGE Publications.

He, G., Liu, P., Zheng, X., Zheng, L., Hewlin, P. F., & Yuan, L. (2023). Being Proactive in the Age of AI: Exploring the Effectiveness of Leaders' AI Symbolization in Stimulating Employee Job Crafting. *Management Decision, 61,* 2896-2919. https://typeset.io/papers/being-proactive-in-the-age-of-ai-exploring-the-effectiveness-24t04zf8?utm_source=chatgpt

Ho, C., & MacDorman, K. F. (2017). Measuring the Uncanny Valley Effect: Refinements to Indices for Perceived Humanness, Attractiveness, and Eeriness. *International Journal of Social Robotics, 9,* 129-139. https://doi.org/10.1007/s12369-016-0380-9

Insch, G. S., Moore, J. E., & Murphy, L. D. (1997). Content Analysis in Leadership Research: Examples, Procedures, and Suggestions for Future Use. *The Leadership Quarterly, 8,* 1-25. https://doi.org/10.1016/s1048-9843(97)90028-x

Kaggwa, S., Eleogu, T. F., Okonkwo, F., Farayola, O. A., Uwaoma, P. U., & Akinoso, A. (2024). AI in Decision Making: Transforming Business Strategies. *International Journal of Research and Scientific Innovation, X,* 423-444.

https://doi.org/10.51244/ijrsi.2023.1012032

Kasunic, M. (2005). *Designing an Effective Survey.* Carnegie Mellon University.

Katz, J. (2015). A Theory of Qualitative Methodology: The Social System of Analytic Fieldwork. *Méthod(e)s: African Review of Social Sciences Methodology, 1,* 131-146. https://doi.org/10.1080/23754745.2015.1017282

Kitto, S. C., Chesters, J., & Grbich, C. (2008). Quality in Qualitative Research. *Medical Journal of Australia, 188,* 243-246. https://doi.org/10.5694/j.1326-5377.2008.tb01595.x

Kivunja, C. (2018). Distinguishing between Theory, Theoretical Framework, and Conceptual Framework: A Systematic Review of Lessons from the Field. *International Journal of Higher Education, 7,* 44-53. https://doi.org/10.5430/ijhe.v7n6p44

Klenke, K. (2008). *Qualitative Research in the Study of Leadership.* Emerald Group Publishing.

Lay, S. (2015). The Uncanny Valley Effect. PhD Thesis, Open University (United Kingdom).

Levitt, H. M. (2021). Qualitative Generalization, Not to the Population but to the Phenomenon: Reconceptualizing Variation in Qualitative Research. *Qualitative Psychology, 8,* 95-110. https://doi.org/10.1037/qup0000184

Lincoln, Y. S., Guba, E. G., & Pilotta, J. J. (1985). Naturalistic Inquiry. *International Journal of Intercultural Relations, 9,* 438-439. https://doi.org/10.1016/0147-1767(85)90062-8

Lu, E. M. (2021). Behind the Uncanny Valley of Mind: Investigating the Effects of Agency and Experience in Chatbot Interactions (Identity No. 1282271). Master's Thesis. Eindhoven University of Technology. https://research.tue.nl/en/studentTheses/behind-the-uncanny-valley-of-mind

Lumsden, J., & Morgan, W. (2005). Online-Questionnaire Design: Establishing Guidelines and Evaluating Existing Support. In *16th Annual Information Resources Management Association International Conference* (pp. 407-410). IGI Global.

Mäkäräinen, M., Saari, T., & Takala, T. (2015). This Study Contradicts the Traditional View of the Uncanny Valley by Showing That Strangeness Can Be Associated with Positive Emotional Reactions. *Journal of Experimental Psychology, 11,* 341-357.

Mathur, M. B., & Reichling, D. B. (2020). Does Mind Perception Explain the Uncanny Valley? A Meta-Regression. *Journal of Experimental Social Psychology, 85,* Article ID: 103855.

Mayring, P. (2004). 5.12 Qualitative Content Analysis. In U. Flick, E. von Kardoff, & I. Steinke (Eds.), *A Companion to Qualitative Research* (pp. 266-269). SAGE Publications Ltd.

Mayring, P. (2015). Qualitative Content Analysis: Theoretical Background and Procedures. In A. Bikner-Ahsbahs, C. Knipping, & N. Presmeg (Eds.), *Approaches to Qualitative Research in Mathematics Education* (pp. 365-380). Springer. https://doi.org/10.1007/978-94-017-9181-6_13

Mayring, P. (2021). *Qualitative Content Analysis: A Step-by-Step Guide.* SAGE Publications Ltd.

Mays, N., & Pope, C. (2020). Quality in Qualitative Research. In C. Pope, & N. Mays (Eds.), *Qualitative Research in Health Care* (pp. 211-233). Wiley-Blackwell.

McGuire, J., & De Cremer, D. (2023). Algorithms, Leadership, and Morality: Why a Mere Human Effect Drives the Preference for Human over Algorithmic Leadership. *AI and Ethics, 3,* 601-618. https://doi.org/10.1007/s43681-022-00192-2

McGuire, J., De Cremer, D., Hesselbarth, Y., De Schutter, L., Mai, K. M., & Van Hiel, A. (2023). The Reputational and Ethical Consequences of Deceptive Chatbot Use. *Scientific Reports, 13,* Article No. 16246. https://doi.org/10.1038/s41598-023-41692-3

Merriam, S. B. (2002). Introduction to Qualitative Research. *Qualitative Research in Practice: Examples for Discussion and Analysis, 1,* 1-17.

Merriam, S. B., & Grenier, R. S. (2019). *Qualitative Research in Practice: Examples for Discussion and Analysis*. John Wiley & Sons.

Mihai, R., & Crețu, A. (2019). Leadership in the Digital Era. *Valahian Journal of Economic Studies, 10,* 65-72. https://doi.org/10.2478/vjes-2019-0006

Mozafari, N., Hammerschmidt, M., & Weiger, W. (2021). That's So Embarrassing! When Not to Design for Social Presence in Human-Chatbot Interactions. In *Proceedings of the International Conference on Information Systems*. https://ssrn.com/abstract=4305288

Myers, M. (2000). Qualitative Research and the Generalizability Question: Standing Firm with Proteus. *The Qualitative Report, 4*, 1-9. https://doi.org/10.46743/2160-3715/2000.2925

Nguyen, T., & van der Lee, P. (2021). *The Uncanny Valley Effects: Influence of Enhancing the Human-Likeness of E-Commerce Chatbot on Perceived Social Presence and Patronage Intention.*

Onwuegbuzie, A. J., & Collins, K. M. (2007). A Typology of Mixed Methods Sampling Designs in Social Science Research. *Qualitative Report, 12*, 281-316.

Pagnozzi, F., Renkema, M., & Antonelli, G. (2024). Help-Seeking Practices and Behaviours in Organizations and Ai-Driven Chatbot: A Scoping Review. *puntOorg International Journal, 1*. https://doi.org/10.19245/25.05.of.13.06.24.1

Paudel, R. (2024). The Impact of Automation and Artificial Intelligence (AI) on Leadership and the Workforce. *Indonesian Journal of Banking and Financial Technology, 2,* 109-124. https://doi.org/10.55927/fintech.v2i2.8904

Pavlidou, P. (2021). *The Uncanny Valley: The Human-Likeness of Chatbots and Its Paradoxical Impact on Consumers' Purchase Intention in E-Commerce.* Master's Thesis, Tilburg University.

Petrat, D., Yenice, I., Bier, L., & Subtil, I. (2022). Acceptance of Artificial Intelligence as Organizational Leadership: A Survey. *TATuP, 31,* 64-69.

Primasatya, R. D., Labbaik, M., Mujtaba, M. I. E., Wahono, R. D., & Wahono, R. D. (2024). Self-Leadership Dalam Menyikapi Perkembangan Teknologi Chatbots AI di Dunia Pendidikan Akuntansi: Tinjauan Perspektif Adaptive Leadership. *Owner, 8,* 1944-1955. https://doi.org/10.33395/owner.v8i2.2313

Rajaobelina, L., Prom Tep, S., Arcand, M., & Ricard, L. (2021). Creepiness: Its Antecedents and Impact on Loyalty When Interacting with a Chatbot. *Psychology & Marketing, 38,* 2339-2356. https://doi.org/10.1002/mar.21548

Rapp, A., Curti, L., & Boldi, A. (2021). The Human Side of Human-Chatbot Interaction: A Systematic Literature Review of Ten Years of Research on Text-Based Chatbots. *International Journal of Human-Computer Studies, 151,* Article ID: 102630. https://doi.org/10.1016/j.ijhcs.2021.102630

Reja, U., Manfreda, K. L., Hlebec, V., & Vehovar, V. (2003). Open-Ended vs. Close-Ended Questions in Web Questionnaires. *Developments in Applied Statistics, 19*, 159-177.

Ressem, V. (2023). *Implementing AI in Organizations: The Role of Leaders and Its Impact on Leadership.* Master's Thesis, Norwegian University of Science and Technology.

Roller, M. R. (2019). A Quality Approach to Qualitative Content Analysis: Similarities and Differences Compared to Other Qualitative Methods. *Forum Qualitative Sozialforschung Forum: Qualitative Social Research, 20*.

Schafheitle, S., Weibel, A., & Rickert, A. (2021). The Bermuda Triangle of Leadership in the AI Era? Emerging Trust Implications from “Two-Leader-Situations” in the Eyes of Employees. In *Proceedings of the Annual Hawaii International Conference on System Sciences* (p. 5473). Hawaii International Conference on System Sciences. https://doi.org/10.24251/hicss.2021.665

Shick, M., Johnson, N., & Fan, Y. (2023). Artificial Intelligence and the End of Bounded Rationality: A New Era in Organizational Decision Making. *Development and Learning in Organizations: An International Journal, 38,* 1-3. https://doi.org/10.1108/dlo-02-2023-0048

Shoydin, M., & Pazoev, Y. (2022). The Study Highlights the Importance of Avoiding the Uncanny Valley Effect When Compressing and Transmitting Complex Holographic Images. *Journal of Holographic Imaging, 34,* 123-138.

Skjuve, M., Haugstveit, I. M., Følstad, A., & Brandtzaeg, P. B. (2019). Help! Is My Chatbot Falling into the Uncanny Valley? An Empirical Study of User Experience in Human-Chatbot Interaction. *Human Technology, 15,* 30-54. https://doi.org/10.17011/ht/urn.201902201607

Song, S. W., & Shin, M. (2022). Uncanny Valley Effects on Chatbot Trust, Purchase Intention, and Adoption Intention in the Context of E-Commerce: The Moderating Role of Avatar Familiarity. *International Journal of Human-Computer Interaction, 40,* 441-456. https://doi.org/10.1080/10447318.2022.2121038

Stenfors, T., Kajamaa, A., & Bennett, D. (2020). How to… Assess the Quality of Qualitative Research. *The Clinical Teacher, 17,* 596-599. https://doi.org/10.1111/tct.13242

Stepp Jr., D. C. (2022). *Investigating Variables to Reduce the Uncanny Valley Effect in Human-Robot Interaction: A Systematic Literature Review*. Ph.D. Thesis, University of Arizona Global Campus.

Strange, J. (2022). Human-Robot Collaboration: A Multilevel and Integrated Leadership Framework. *Leadership Quarterly, 33,* Article ID: 101594.

Strong, N., & Terblanche, N. (2020). Chatbots as an Instance of an Artificial Intelligence Coach. In R. Wegener (Ed.), *Coaching im digitalen Wandel* (pp. 51-62). Vandenhoeck & Ruprecht. https://doi.org/10.13109/9783666407420.51

Tîrnăcop, A. B. (2023). Leadership in the Digital Era: Exploring the AI-EI Nexus. *Cactus, 5,* 45-52. https://doi.org/10.24818/cts/5/2023/1.05

Tschöpe, N., Reiser, J. E., & Oehl, M. (2017). Exploring the Uncanny Valley Effect in Social Robotics. In *Proceedings of the Companion of the 2017 ACM/IEEE International Conference on Human-Robot Interaction* (pp. 307-308). ACM. https://doi.org/10.1145/3029798.3038319

Tu, T., Cheng, H., & Chu, Y. (2020). This Research Shows That Older Adults Do Not Experience the Uncanny Valley Effect as Younger and Middle-Aged Adults Do. *Journal of Geriatric Robotics, 5,* 65-78.

Van Lierop, T. (2021). *The Uncanny Valley: The Paradoxical Effects of Enhancing the Human-Likeness of Customer Support Chatbots on Perceived Trust and Satisfaction.*

Van Quaquebeke, N., & Gerpott, F. H. (2023). The Now, New, and Next of Digital Leadership: How Artificial Intelligence (AI) Will Take over and Change Leadership as We Know It. *Journal of Leadership & Organizational Studies, 30,* 265-275. https://doi.org/10.1177/15480518231181731

Varpio, L., Paradis, E., Uijtdehaage, S., & Young, M. (2020). The Distinctions between

Theory, Theoretical Framework, and Conceptual Framework. *Academic Medicine, 95,* 989-994. https://doi.org/10.1097/acm.0000000000003075

Wald, R., Heijselaar, E., & Bosse, T. (2021). Make Your Own: The Potential of Chatbot Customization for the Development of User Trust. In *Adjunct Proceedings of the 29th ACM Conference on User Modeling, Adaptation and Personalization* (pp. 382-387). ACM. https://doi.org/10.1145/3450614.3463600

Wang, P., & Shao, J. (2022). Escaping Loneliness through Tourist-Chatbot Interactions. In J. L. Stienmetz, B. Ferrer-Rosell, & D. Massimo, (Eds.), *Information and Communication Technologies in Tourism 2022* pp. (473-485). Springer. https://doi.org/10.1007/978-3-030-94751-4_44

Watson, G. J., Desouza, K. C., Ribiere, V. M., & Lindič, J. (2021). Will AI Ever Sit at the C-Suite Table? The Future of Senior Leadership. *Business Horizons, 64,* 465-474. https://doi.org/10.1016/j.bushor.2021.02.011

Waugh, W. L., & Streib, G. (2006). Collaboration and Leadership for Effective Emergency Management. *Public Administration Review, 66,* 131-140. https://doi.org/10.1111/j.1540-6210.2006.00673.x

Zhang, G., Chong, L., Kotovsky, K., & Cagan, J. (2023). Trust in an AI versus a Human Teammate: The Effects of Teammate Identity and Performance on Human-AI Cooperation. *Computers in Human Behavior, 139,* Article ID: 107536. https://doi.org/10.1016/j.chb.2022.107536