Group Decision and Negotiation (2026) 35:33
https://doi.org/10.1007/s10726-026-09990-z

Extroversion-Introversion Design of Social Robots: The Role of the Mental Model Attribution Process

Viviana M. Oberhofer[1] · Isabella Seeber[2] · Lena Waizenegger[3]

Received: 13 May 2025 / Accepted: 7 April 2026

Abstract
Social robots are becoming more autonomous and they are likely to soon join organizational teams as active team members. As such, their personality design—specifically, their extroversion-introversion personality—matters, as it shapes individuals' affective reactions. Yet it remains unclear through which underlying cognitive processes robot personality influences task satisfaction in team contexts. Past research efforts to understand these processes resulted in dispersed and conflicting theories. This study proposes a parsimonious conceptual model integrating theories on anthropomorphism and the theory of mind: the mental model attribution process (MMAP). Based on a between-subject animated video vignette study with 401 crowd workers, the MMAP explains how robot extroversion-introversion cues affect task satisfaction. The results show that extroverted social robots elicit higher task satisfaction than introverted robots. This effect is explained by increased anthropomorphism, leading to more agency and experience inference, and higher ascribed robot empathy. By integrating research on anthropomorphism, theory of mind, and robot personality design, this study contributes a parsimonious, empirically testable conceptual model for understanding affective reactions to social robots in a team context.

Keywords Anthropomorphism · Empathy · Human-agent team · Mental model attribution process · Extroversion · Satisfaction

✉ Viviana M. Oberhofer
viviana.oberhofer@uibk.ac.at

Isabella Seeber
isabella.seeber@grenoble-em.com

Lena Waizenegger
lena.waizenegger@aut.ac.nz

1 University of Innsbruck, Innsbruck, Austria

2 Grenoble Ecole de Management, Grenoble, France

3 Auckland University of Technology, Auckland, New Zealand

Published online: 30 April 2026

1 Introduction

Social robots are becoming part of non-industrial work environments, supporting employees, for example, in food deliveries (Yörük et al. 2024), as hotel concierges, waiters, retail assistants, or receptionists (Seo 2022). Social robots are embodied software-based agents, which are capable of interacting with humans in a natural manner (Baird and Maruping 2021; Fong et al. 2003). Existing research on human-robot interaction (HRI) and human-agent teams (HATs) has investigated how social robots can be designed to gain users' trust and increase adoption to reap the benefits of AI-enabled technologies. Especially as empirical evidence shows that social robots can improve task performance in teams (You and Robert 2018). However, a growing number of people express fear about social robots perceiving their advanced capabilities as a potential threat to human identity (Chang and Kim 2022) and the fear of being replaced by autonomous agents (e.g., Strich et al. 2021). While some teams can imagine collaborating with a social robot, others prefer to work with human team members (e.g., Baltrusch et al. 2022). The design of social robots can significantly shape such human perceptions and trigger positive affective responses, which are a key lever of acceptance of robotic systems.

Personality traits are one design specification that have been shown to elicit positive perceptions and mitigate fears (e.g., Mileounis et al. 2015; Paetzel-Prüsmann et al. 2021). Personality is a "set of distinctive qualities that, taken collectively, distinguish individuals" (Fong et al. 2003, p. 15). For social robots, it is a key predictor for robot acceptance (Esterwood et al. 2022). Goldberg (1990)'s Big-Five model proposes five personality traits: extroversion, conscientiousness, agreeableness, neuroticism, and openness. The extroversion-introversion dimension is particularly relevant for designers of social robots intended for workplace contexts, as extroversion-introversion cues (e.g., speech intonation, movements) can be easily incorporated into the social robot's behavior and speech. Other dimensions pose challenges for designers as they directly affect the social robot's ability to function reliably. A social robot that is badly organized and not dependable (i.e., low conscientiousness), not cooperative (i.e., low agreeableness), or emotionally unstable (i.e., high neuroticism) would not be realistically introduced within a team. Openness to experience may be advantageous for certain tasks but counterproductive for others; for example, a very curious robot would be ill-suited for repetitive sorting or categorization tasks. By contrast, extroversion-introversion does not impede a social robot's capacity to perform tasks across contexts, making it a uniquely adaptable design dimension. For these reasons, this research focuses on the extroversion-introversion dimension of the Big Five.

To understand how a social robot's extroversion-introversion design influences human affective responses, it is essential to examine the underlying cognitive processes that shape how people interpret and respond to social robots. Theories such as the Uncanny Valley theory (Mori 1970), the computers-are-social-actors (CASA) paradigm (Nass and Lee 2001), and the theory of anthropomorphism (Epley et al. 2007) have been highly influential in explaining how design choices influence human perceptions of non-human agents. The Uncanny Valley theory suggests that human-like morphology of machines influences individuals' emotional responses. The CASA paradigm argues that people mindlessly apply social rules, norms, and expectations when interacting with computers (Gambino et al. 2020; Nass and Lee 2001). However, neither CASA nor the Uncanny Valley accounts for the cognitive attribution processes that shape an individual's reactions. The theory of anthropomorphism offers such an explanation by describing the individual's tendency to ascribe non-human agents' human characteristics (Airenti 2015; Epley et al. 2007). However, despite its prominence, the theory of anthropomorphism has shown limited explanatory power, as empirical findings remain inconsistent and often contradictory across different social robot designs and contexts (e.g., Lu et al. 2021; Mathur et al. 2020; Szczepanowski et al. 2020).

The goal of this study is to revisit the anthropomorphism theory (e.g., Epley et al. 2007) and related theories, such as the theory of mind (Premack and Woodruff 1978) and mind perception (Airenti 2015; Epley and Waytz 2010), to better understand how people *perceive* personality cues in social robots. To synthesize these perspectives, we propose a parsimonious conceptual model—the Mental Model Attribution Process (MMAP)—which captures the core cognitive attribution processes through which people make sense of a social robot's design. We empirically test this model by examining how variations in a robot's extroversion-introversion personality design influence users' attributions of mental models. Thus, considering MMAP as the causal mechanism, we pose the following research question: *How does a social robot's extroverted personality affect task satisfaction in comparison to an introverted personality?*

To answer this research question, we performed an animated vignette study involving 401 crowd workers, in which participants experienced a team collaboration scenario with an extroverted or an introverted social robot in a business-like team meeting. Our study reveals that an extroverted social robot fosters higher task satisfaction and that this effect is explained through the inference of mental states and the ascription of human-like capabilities to the social robot. These findings contribute to research on anthropomorphism and robot personality design.

2 Theoretical Underpinnings

2.1 Anthropomorphism Revisited: Mixed Empirical Results

Among existing approaches, anthropomorphism theory remains the dominant lens for understanding humans' perceptions of non-human agents. Anthropomorphism is defined as the individual's inference and attribution of human characteristics, traits, and mental states to non-human agents (Airenti 2015; Damiano and Dumouchel 2018; Epley et al. 2007; Gambino et al. 2020). Epley et al. (2007) posits three psychological determinants for anthropomorphizing: (1) people use their knowledge of humans to interpret non-human agents (elicited agent knowledge), (2) people infer intentions and mental states to understand and predict the behavior of non-human agents in uncertain contexts (effectance motivation), and (3) the need for social connection lets people treat non-human agents as social partners (sociality motivation). To study anthropomorphism, researchers frequently manipulated the morphology and appearance of social robots to appear more or less human-like (e.g., Rosenthal-

Von Der Pütten and Krämer 2014). Yet this manipulation did not lead to consistent findings (see Table 1). While some studies found that a more human-like appearance leads to more trust (e.g., Lu et al. 2021) or more likability (e.g., Szczepanowski et al. 2020), others found the opposite effect (e.g., Mathur and Reichling 2016; Mathur et al. 2020).

This inconsistency could be attributed to the inconclusive findings of morphology predicting anthropomorphism (e.g., Banerjee et al. 2025; Mara and Appel 2015a; Szczepanowski et al. 2020). The inconsistency can also be explained through one overlooked key aspect: the attribution of human characteristics to non-human agents is not determined by the non-human agent itself (Airenti 2015). Rather, it depends on people's mental states, motivations (Epley et al. 2007), and the social interaction context in which the non-human agent is encountered (Airenti 2015), which activate cognitive attribution processes (Gray and Wegner 2012). As summarized in Table 1, much of the existing research has not considered these underlying cognitive processes as a mediating mechanism between a social robot's appearance and subsequent affective responses. Moreover, prior research has predominantly emphasized the role of physical design cues in social robots as a means to elicit anthropomorphism. Emerging evidence, however, suggests that non-physical design cues—such as behavioral traits reflecting extroversion or introversion—can likewise evoke anthropomorphic perceptions in human users (Andriella et al. 2025; Blut et al. 2021). Despite these insights, little is known about how such non-physical cues shape the underlying cognitive attribution processes through which humans ascribe human-like qualities to

Table 1 Examples of contradictory findings on the effects of anthropomorphism

Negative effect	Manipulation, predictor & Outcome variable	Positive effect
The more humanlike the robot, the more submissive it was perceived compared to the more mechanical robots (Rosenthal-Von Der Pütten and Krämer 2014).	Manipulation: robot appearance (i.e., mechanical, humanoid, android) Outcome variable: Dominance-Submissiveness	Android robot perceived as more dominant than humanoid and mechanical robot (Mara and Appel 2015a).
Less humanlike features led to more trust (Mathur and Reichling 2016).	Manipulation: robot appearance (i.e., mechanical, humanoid, android) Robot voice, Robot eyes (i.e., larger vs. smaller, position on the face) Outcome variable: Trust	The more humanlike voice led to more trust than monotonous, mechanical voice (Lu et al. 2021). More large eyes led to more trust than medium and small eyes (Song et al. 2021). More humanlike appearance leads to more trust (Qin et al. 2025).
More mechanicalness increases likability (Rosenthal-Von Der Pütten and Krämer 2014). Non-linear relationship between robot appearance and likability (Banerjee et al. 2025; Mathur et al. 2020; Sun and Xiao 2025).	Manipulation: robot appearance (i.e., mechanical, humanoid, android) Outcome variable: Likability	More humanlike appearance increases likability (Rosenthal-Von Der Pütten and Krämer 2014; Szczepanowski et al. 2020)
Higher perceived anthropomorphism reduced the level of eeriness (Mara and Appel 2015b).	Predictor: perceived anthropomorphism Outcome variable: Eeriness	Higher perceived anthropomorphism led more eeriness (Yam et al. 2021)

non-human agents. Consequently, a systematic understanding of how these cognitive processes unfold remains limited.

2.2 Toward an Integrative Account of Cognitive Attribution: Mental Model Attribution Process

Given the contradictory findings outlined above, revisiting the foundational theory of mind (ToM) and related conceptual developments offers a crucial starting point for developing a unified model of cognitive attribution processes.

Theory of mind (ToM) refers to the human ability to infer mental states—such as beliefs, desires, and emotions (Gray et al. 2007)—to oneself and others (Premack and Woodruff 1978), to explain and predict the behavior of other agents. Through observing others, people can infer the mental states of these agents (Mitchell and Phillips 2015) that represent the agent's view of the world, regardless of whether those states correspond to actual reality (Frith and Frith 2005).

The attribution of mental states is shaped by one's own mental models about the world. Mental models describe meta-representations "that individuals construct in order to support their predictions and understanding of the world around them" (Phillips et al. 2011, p. 1491). They function as structured representations (Epley et al. 2007; Phillips et al. 2011), which build the foundation for explaining, rationalizing, predicting behaviors (Mitchell and Phillips 2015; Premack and Woodruff 1978), and choosing an appropriate behavioral or emotional reaction (Airenti 2015; Damiano and Dumouchel 2018). Mental models constantly adapt with each interaction and observation of behaviors (Phillips et al. 2011), and with each gained experience (Schlinger 2009).

Since its inception in 1978, ToM research has expanded steadily across domains such as human development, neurological disorders, and human-non-human interaction (Schaafsma et al. 2015). This growth has produced a proliferation of related concepts, such as intentionality (Malle and Knobe 1997) and mentalizing (Epley and Waytz 2010; Frith and Frith 2005), which researchers tried to combine in the theory of mind perception (Airenti 2015; Epley and Waytz 2010) (see Table SM1 in Supplementary Materials). Yet instead of clarifying the relationships, it led to conceptually overlapping and unclear hierarchical relations, which make delineations and causal relationships difficult to investigate.

Building on this extensive literature, we propose a unifying cognitive attribution mechanism that we term the *mental model attribution process* (MMAP) (see Fig. 1).

In *human-to-human interaction*, this attribution process begins with *observing* a behavior (see Fig. 1, left side). The observer relies on the social cues the human agent provides to make sense of the behavior (Gambino et al. 2020). In this sense-making process, the observer *classifies* the behavior, which means that they deconstruct the observed behavior (Schaafsma et al. 2015) into meaningful components (e.g., autonomy, intentionality; De Graaf and Malle (2019); Malle and Knobe (1997)). Then, the observer *attributes* a mental state by reconstructing these components (Schaafsma et al. 2015). In this attribution step, pre-existing mental models guide, which and how cues are selected and integrated (Epley and Waytz 2010; Epley et al. 2007; Leslie et al. 2004). Thereby, existing mental models are *refined* and updated with each encoun-

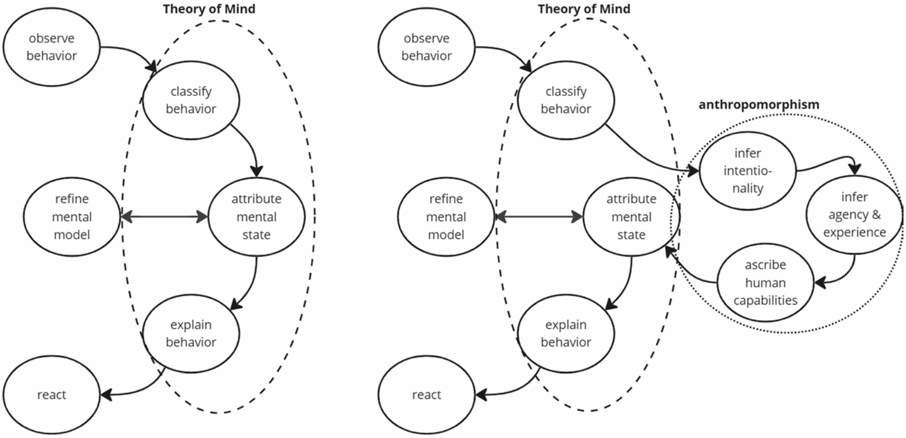

Fig. 1 Mental model attribution process. **a** on the left: Mental state attribution of other human agents; based on the theory of mind. **b** on the right: Mental state attribution of non-human agents; extended by anthropomorphism

ter (Gambino et al. 2020; Leslie et al. 2004; Phillips et al. 2011). The mental state attribution allows *explaining* the behavior (De Graaf and Malle 2019) and *reacting* to it accordingly (Leslie et al. 2004).

In the context of team collaboration, when the observer sees a human teammate give a project update (observe behavior), they rely on several deconstructive processes to *classify* what they notice—for example, one may interpret the repeated referencing of deadlines ("We need this by Friday") and check-ins by the teammate as a signal of urgency. Guided by their existing mental models, the observer then *attributes* specific mental states, such as believing the project is delayed or under pressure. Based on this updated mental model, the observer can anticipate that project work will need prioritization and can react by identifying low-impact tasks that can be deprioritized.

In *human-to-non-human interaction*, the cognitive attribution process needs to be extended with anthropomorphism (see Fig. 1, right side). This is required because humans infer mental states in non-human agents that do not have minds of their own. The observer relies on the social cues exerted by the non-human agent, which automatically activates ToM and anthropomorphism (Airenti 2015; De Graaf and Malle 2019; Gambino et al. 2020).

Anthropomorphism enables the attribution of human-like mental capacities to non-human agents, without requiring people to believe that the non-human agent possesses mental states (Damiano and Dumouchel 2018; Epley et al. 2007). Thereby, anthropomorphism facilitates the inference of *intentionality* (i.e., understanding the non-human agent as an intentional agent) (Airenti 2015; De Graaf and Malle 2019; Malle and Knobe 1997). This constitutes a prerequisite for inferring *agency*, which describes the capacity to act, plan, and have self-control, shape the mental states people attribute, and *experience*, which describes the capacity to feel emotions (Damiano and Dumouchel 2018; Epley and Waytz 2010; Gray et al. 2007). Throughout this process, people draw on their existing mental models to interpret the observed behaviors

(Gambino et al. 2020), using them to generate inferences about the agent's possible mental states (Damiano and Dumouchel 2018).

Continuing the example above, when a social robot, as a non-human agent, gives a project update, its (non-)verbal cues (e.g., referencing of deadlines, repeated reminders) can automatically trigger ToM. Actively sending reminders may lead the observer to attribute agency, as the robot appears capable of planning or prioritizing tasks.

It is important to note that we consider the MMAP applicable to non-human agents that show social cues (Gambino et al. 2020) and are perceived as interaction partners (Airenti 2015). Also, individual characteristics (e.g., gender, attitudes) and situational factors (e.g., task, duration) serve as important boundary conditions, as inferences of mental states are (at least partially) influenced by the mental models of oneself (Epley and Waytz 2010). Thus, the perception of non-human agents varies at the individual level, and the emergence of different outcomes can be explained by the inference and attribution of agency, experience, and human-like capabilities.

Drawing on the proposed MMAP (Fig. 1), we argue that many contradictory findings in the literature, as shown in Table 1, could be reconciled by accounting for differences in the agency, experience, and capabilities that people attribute to non-human agents. For example, some studies have solely focused on how the social robot's morphology shapes the perception of dominance or submissiveness (Mara and Appel 2015a; Rosenthal-Von Der Pütten and Krämer 2014). The MMAP suggests that this perception might be connected to the inference and attribution of *agency*. Whereas some people might interpret a social robot with a mechanical appearance as signaling dominance and might lead to the association with bodyguards or military personnel, who express no emotions and are very composed, others may infer a more playful mental state because the social robot reminds them of a toy, leading them to attribute feelings of simple mindedness instead. Similarly, contradictory findings on how a social robot's human-like appearance affects trust may be explained by differences in attributed *experience*. Some people may trust a humanoid social robot more because it reminds them of other humans, which leads to the ascription of emotions, affect, and empathy, whereas others may trust humanoid social robots less because they rather perceive it as a technology pretending to have emotions.

3 Hypotheses Development

The proposed MMAP represents a parsimonious and empirically testable model. We operationalize the social robot's behavior by investigating the effect of simulated robot extroversion-introversion cues on the human affective reaction in a collaboration context. The MMAP is captured through the attribution of human-like mental capacities to the social robot (i.e., anthropomorphism), *perceived* agency and experience reflecting goal-directed and affective capacities, and its ascribed affective and cognitive states enabling the social robot to respond to human agents' behaviors and emotions (robot empathy). This should allow us to explain humans' affective reactions, operationalized as task satisfaction (see Fig. 2). All focal concepts measure individuals' perception. However, we omit the prefix 'perceived' in the text for better readability.

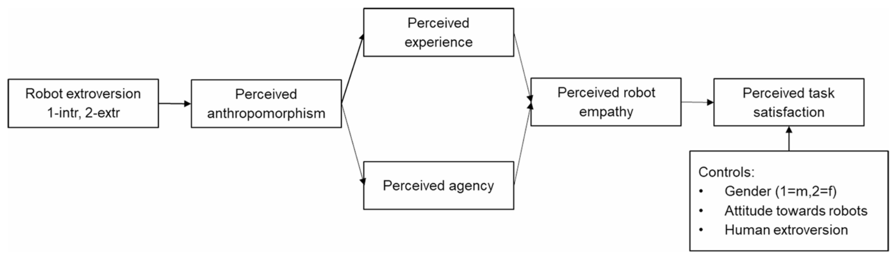

Fig. 2 Research model

3.1 Extroversion-Introversion Cues as Observed Behavior

Research on HATs has shown that extroverted social robots reduce threat perception (Paetzel-Prüsmann et al. 2021), are perceived as more intelligent, and likable (Mileounis et al. 2015), are perceived as more playful and fun (Robert 2018), but are perceived as less safe (Andriella et al. 2021; Andriella et al. 2025) compared to introverted social robots. This shows that extroversion-introversion cues, as non-physical social cues of robots, influence human perceptions (Diefenbach et al. 2023). But it remains unclear if they can trigger the attribution of human-like characteristics.

Extroversion-introversion cues provide socially meaningful signals that observers routinely use to interpret others' internal states. Such cues prompt anthropomorphic interpretation by activating the observer's mental models (e.g., Epley et al. 2007). For instance, Park et al. (2012) showed that telling a joke to an extroverted social robot elicited more anthropomorphism than to an introverted one. Another study in a caregiver context found that extroverted social robots are anthropomorphized more than introverted social robots (Andriella et al. 2021). These findings were replicated also in an assistive task (Andriella et al. 2025). Compared to introverted robots, social cues of extroverted robots, such as stronger intonation in speech and more non-verbal movements, may appear more human-like, more natural, more lifelike than social cues of introverted robots (Andriella et al. 2021; Bartneck et al. 2009; Salem et al. 2013). Social cues increase the chance of the behavior being interpreted as intentional, facilitating anthropomorphic capacity attribution (Gambino et al. 2020). We therefore expect that extroversion cues are associated with higher anthropomorphism perceptions than introversion cues in a team context (see Fig. 1):

Hypothesis 1 *Collaborating with an extroverted robot versus an introverted robot leads to higher anthropomorphism.*

3.2 Mental State Attribution and Anthropomorphism

After observing a non-human agent's behavior exerting social cues (Gambino et al. 2020), the MMAP suggests a sequential process in which the anthropomorphism process triggers the inference of agency and experience, and consequently the ascription of human capabilities (Epley and Waytz 2010).

Anthropomorphism enables mental state attribution where human-like psychological capacities are ascribed to the non-human agent to support interpretation and prediction of its behavior (Damiano and Dumouchel 2018). Even minimal social cues, such as emotional, dynamic speech for extroverted agents or unemotional, neutral speech for introverted agents, can be sufficient to automatically engage mind attribution processes (Airenti 2015; Gambino et al. 2020). Observers draw on existing mental models associated with these personality cues to generate inferences about the agent's internal states (Damiano and Dumouchel 2018), and thereby attribute agency (i.e., intentional action, planning, and self-control) and experiences (i.e., the capacity to feel emotions) to the social robots (Guidi et al. 2025; Nardelli et al. 2025). In this process, perceptions of the agent's capacity for agency and experience play a central role in shaping attributed mental states (Epley and Waytz 2010; Gray et al. 2007). Consequently, stronger anthropomorphic interpretations of extroversion cues are expected to increase attributions of both agency and experience (Guidi et al. 2025; Nardelli et al. 2025). Hence, we suggest:

Hypothesis 2 *Higher anthropomorphism leads to an increased inference of (a) agency and (b) experience.*

The inference of a social robot as an intentional agent inherently entails the attribution of human-like capabilities. This research argues that once individuals infer agency and experience in a social robot, they subsequently ascribe capabilities such as empathy to it.

Empathic capabilities suggest that an agent can detect emotions, rationalize them, and form an affective reaction towards another agent (Mitchell and Phillips 2015). The ascription of empathy to social robots can be stimulated through the manipulation of extroversion-introversion social cues, such as gaze direction (Alves-Oliveira et al. 2019) or gestures (Charrier et al. 2018). Moreover, Edwards et al. (2022) showed that anthropomorphizing an animated agent leads to more ascribed empathic capabilities.

Drawing on the MMAP framework, the attribution of empathetic capabilities to a social robot is a function of increased anthropomorphism, which materializes through the perception of the social robot as an intentional agent endowed with mental capabilities. The more individuals anthropomorphize a social robot, the more they are likely to perceive it as possessing more agency and experience, thereby facilitating the ascription of capabilities such as empathy. Thus, we propose:

Hypothesis 3 *Higher inferred (a) agency and (b) experience, induced through higher anthropomorphism, lead to higher ascribed robot empathy.*

3.3 From Ascribed Human Capability to Affective Outcomes

Finally, we test the affective reaction of the human in terms of perceived task satisfaction. Satisfaction has been shown to be a critical affective predictor of the acceptance of social robots (Kim and Kim 2021), promoting positive job quality (Baltrusch et al. 2022) and behavioral intentions to continue using the robot (Seo and Lee 2021). In teams, employee satisfaction is also vital for improving performance (Robert and

You 2018), heightening the willingness to learn and the willingness for knowledge distribution (Medina 2016). Drawing on the MMAP, we argue that attributing agency, experience, and human-like capabilities to non-human agents renders collaboration more enjoyable and rewarding. Prior research showed that extroverted social robots elicit more positive attitudes and trust (Tay et al. 2014), improve the collaboration (Andriella et al. 2021), and result in extended collaborations in comparison to introverted robots (Andrist et al. 2015). Accordingly, collaborating with an extroverted social robot should result in higher task satisfaction. In line with the MMAP, we expect this to be explained by the increased perception of anthropomorphism, the inference of agency and experience, attribution of empathic capabilities, ultimately improving task satisfaction. Thus, we propose:

Hypothesis 4 *Collaborating with an extroverted social robot versus an introverted robot leads indirectly to greater task satisfaction. An extroverted social robot leads to more anthropomorphism in comparison to an introverted social robot, which in turn results in more ascribed agency and experience. This results in increased perception of robot empathy, which positively influences task satisfaction.*

4 Method

We conducted a between-subject online experiment with Amazon Mechanical Turk workers (MTurkers) who viewed an animated video vignette of a collaboration between an intelligent social robot and two human members of a marketing team. Participants were randomly assigned to view either an extroverted or an introverted social robot.

4.1 Video Vignette and Manipulation of Robot Extroversion-Introversion

We adopted a video-vignette study to simulate a real project management environment. The vignettes allowed us to capture the participants' attitudinal preferences and judgments toward social robots through fictional scenarios (Aguinis and Bradley 2014). We chose a first-person animated video vignette to allow inexperienced robot users a better immersion into the scenario (Fernández-Llamas et al. 2018). Although vignette-based studies lack real life interactions, and thus have reduced external validity, past research has offered guidelines on increasing vignettes' validity and reliability (e.g., Matza et al. 2021; Steiner et al. 2016). We followed these guidelines by working closely with experts in the design and evaluation of the vignettes (Matza et al. 2021), embedding it within a survey, and employing random sampling (Steiner et al. 2016). Vignettes offer distinct advantages over lab or field experiments by overcoming sampling (Matza et al. 2021), practical, technical, and ethical constraints. As a consequence, vignette studies are widely adopted in HRI research (Oberhofer et al. 2023), as they provide effective means of eliciting user perceptions.

The vignette portrays a project meeting between a marketing manager (i.e., the participants), a product owner, and the intelligent social robot Jamie. Jamie was depicted as a blue and white humanoid robot with arms, legs, a torso, and a head.

The robot's English name was chosen to match the Anglo-Saxon participants and to facilitate in-group acceptance (Eyssel et al. 2012). We adopted a unisex name to avoid gendering the social robot. The social robot acted fully autonomously, and participants observed the social robot adding features to the product backlog, performing time and cost forecasting, and performing administrative tasks like scheduling meetings. Thus, simulating intelligence.

We based the manipulation of the robot's extroversion-introversion on a multimodal approach, manipulating verbal, para-verbal, and non-verbal cues (Diefenbach et al. 2023). In terms of verbal cues, we adjusted the robot's utterances to fit the characteristics of an extroverted or introverted person while keeping the meaning of the sentences the same (see Supplementary Materials 1.2). In terms of non-verbal cues, we manipulated the movement angles (Lee et al. 2006). The extroverted robot had wider upper body movements and stood in the middle of the frame or moved autonomously through the room. The introverted robot was positioned outside of the center (see Supplementary Materials 1.3), had more restricted upper body movements, and moved less through the room (Mileounis et al. 2015; Tay et al. 2014). In terms of para-verbal cues, the robot's voice was recorded from a native speaking women trained to speak faster (higher word count per minute; 186 words/minute), louder, more dynamically (i.e., pitch range: ~10 Hz), and at a higher pitch (i.e., mean fundamental frequency (F0): ~230 Hz) with emotional intonation in case of the extroverted robot. The introverted robot spoke in a more unemotional tone (i.e., pitch range: ~4.5 Hz), slower (i.e., 136 words/minute), quieter, and at a lower pitch (i.e., mean F0: ~210 Hz) (Mileounis et al. 2015; Nass and Lee 2001; Tay et al. 2014).

4.2 Measures

To measure task satisfaction, we adapted Weiss et al. (1967)'s five-item scale. Anthropomorphism was operationalized according to Bartneck et al. (2009)'s Godspeed Scale with four items. To measure experience, we adapted Gray et al. (2011)'s three-item scale. To measure agency, we adapted the Gray et al. (2011)'s scale to fit closer to the properties describing agency, as originally conceptualized by Wooldridge and Jennings (1995) with seven items. To measure robot empathy, we adapted Pitt et al. (1995)'s scale with five items. Robot extroversion and introversion were measured as manipulation checks with twenty items (Goldberg 1990). We also collected the demographics age, gender identification, country of origin, attitude toward robots, and participant's extroversion as part of the pre-survey. Attitude towards robots was adopted from Heerink et al. (2010) and Tay et al. (2014) and measured with three items. Human's extroversion was measured using the IPIP (International Personality Item Pool 2018) 10-item scale. In addition, we collected the marker variable attitude towards blue (Simmering et al. 2015) with three items. All items were measured on a seven-point Likert scale, except for agency and experience, which were measured on a five-point Likert scale (see Supplementary Materials 1.4).

4.3 Participants and Sample

We recruited participants through the CloudResearch MTurk toolkit[1] prior to the launch of ChatGPT in November 2022. The CloudResearch MTurk toolkit allows the collection of data from Amazon Mechanical Turk (i.e., MTurkers) with more stringent data quality criteria, ensuring a data quality comparable to that of data collected through Prolific (Douglas et al. 2023; Peer et al. 2022). To ensure high data quality, we used several data quality measures: Participants had to (1) pass a CAPTCHA test, (2) provide the same completion code on the Amazon portal that they were given at the end of the survey, (3) pass four attention check questions, (4) participate no more than once, and (5) spend at least ten minutes in the survey. The minimum survey time was set based on a pre-test with volunteers that required a minimum of 13 min. Additionally, only fluent English speaking MTurkers from the United States and the United Kingdom with an above 80% acceptance rate without suspicious geolocation were admitted to the survey. 433 MTurkers completed the survey, of which 32 were rejected because they violated at least one of the above exclusion criteria. We followed a strict payment scheme in which participants were paid only if they met all of these criteria. All participants gave their informed consent to participate in the survey and agreed that they had read the approval criteria and the monetary reward information.

The final sample consisted of 401 participants who were paid USD 4.00 for their participation. Participants were predominantly male (58% male, 42% female); 49% were from the United States, 42% were from the United Kingdom, 5% were from other European countries, and 4% were from Asian or Latin American countries. The majority of participants had a high school diploma or a bachelor's degree. Participants' ages ranged from 18 to 73 years (M = 34.93, SD = 10.94). 51% of the participants received the extraverted robot scenario, and 49% received the introverted robot scenario.

4.4 Experimental Procedure

After giving their consent and completing a pre-survey, participants received a task description and an introduction to the robot (either an extroverted or an introverted robot description). Participants were introduced to the voice of the marketing manager (i.e., their role). Participants listened to either a female or a male voice, depending on their gender identification. The voice introduced the characters and the context of the scenario. Participants then watched the animated video vignette. Since the vignette showed a project meeting based on the agile Scrum framework, we provided a drop-down menu that explained the framework-specific terms. After the video, participants completed the post-survey, received a completion code, and were thanked for their participation.

[1] CloudResearch Mturk toolkit: https://www.cloudresearch.com/products/turkprime-mturk-toolkit/?utm_source=nav.

4.5 Data Analysis

We performed outlier, reliability, validity, assumption, and common method bias (CMB) analyses.

Convergent validity and reliability were assessed with factor loadings from a CFA, AVE, and CR, which led to the exclusion of two items on task satisfaction, four items on agency, and one on robot empathy. As well as the ten items on robot extroversion as the AVE did not reach the threshold of 0.5, thus leaving ten items on robot introversion to assess the manipulation's success. Discriminant validity was assessed using the VIF and the Fornell-Larcker criterion (see Supplementary Materials 1.5). Outlier analyses using the Bonferroni outlier test and Q-Q plots identified two influential observations. However, they did not reach the level of 0.1 in a Cook's distance test and hence were kept in the dataset.

We performed assumption tests for linear regression analyses (Hair et al. 2010) by assessing normality distribution with studentized residual plots, homoscedasticity with Breusch-Pagan tests, linearity between independent and dependent variables with component and residual plots, and statistical independence of errors with Durbin-Watson tests. All tests were deemed satisfactory except for the Breusch-Pagan tests, which indicate heteroscedasticity. Heteroscedasticity does not affect the regression estimates but biases the variance-covariance matrix, rendering F-tests meaningless (Long and Ervin 2000). Thus, we relied on adjusted robust standard errors for all ordinary linear regressions and consistent standard errors for the calculation of serial mediations (Hayes 2018), as well as bootstrapped results.

Finally, we assessed CMB using the most sophisticated way of testing for common method variance (Podsakoff et al. 2024), by employing the CFA marker technique following Williams et al. (2010) with an a-priori determined marker variable (see Supplementary Material 1.6). The analysis revealed no evidence of shared common method variance between the marker variables and the substantive variables (Simmering et al. 2015).

5 Results

We performed a manipulation check using Welch's ANOVA, which showed that the extroverted social robot was indeed perceived as less introverted (i.e., low values signify high introversion) than the introverted robot ($F(1,357) = 109.85$, $p < .001$). Thus, our manipulation of the robot's extroversion-introversion was successful, so that the extroverted robot was perceived as less introverted ($M = 5.96$, $SD = 0.94$) than the introverted robot ($M = 4.77$, $SD = 1.31$).

First, we ran a linear regression to test the effect of the treatment factor, robot extroversion, on satisfaction (see Model (2) in Table 2). The treatment factor had no significant direct effect on task satisfaction ($\beta = 0.128$, 95% CI [−0.04, 0.29]). Robot extroversion positively influenced anthropomorphism ($\beta = 0.641$, 95% CI [0.36, 0.93]), so that extroverted social robots were more anthropomorphized than introverted robots. Hence, supporting H1.

Table 2 Regression results for interaction and task satisfaction with mediators

	Model (1)	Model (2)	Model (3)
	DV: Task satisfaction	DV: Task satisfaction	DV: Anthropomorphism°
Participant's gender	**0.247** [0.081; 0.413]	**0.245** [0.079; 0.411]	0.142 [−0.159; 0.429]
Attitude towards robots	**0.187** [0.111; 0.263]	**0.187** [0.111; 0.263]	**0.312** [0.183; 0.446]
Human extroversion	0.061 [−0.002; 0.123]	0.060 [−0.002; 0.123]	0.002 [−0.108; 0.114]
Robot extroversion		0.128 [−0.037; 0.294]	**0.641** [0.351; 0.924]
Intercept	**4.191** [3.676; 4.705]	**4.003** [3.441; 4.565]	**1.716** [0.743; 2.726]
Observations	401	401	401
R^2	0.094	0.100	0.101
Effect size (f^2)	0.100	0.111	0.112
F statistic	13.26 *** (df=3; 397)	10.85 *** (df=4; 396)	

Robot extroversion: 1-introversion, 2-extroversion

Model (1) and (2) calculated with robust linear regression models; Model (3) results from bootstrapped.

Structured equation model using lavaan

Bold estimates exclude zero in the 95% confidence interval (CI)

*p <.05; **p <.01; ***p <.001.

Hypotheses 2, 3, and 4 test the MMAP path to the affective reaction. We tested this relationship by calculating a bootstrapped structural equation model using the lavaan package in RStudio. The paths leading up to the serial and parallel mediation can be found in Model (3) in Table 2 and in Models 4–6 in Table 3. As already shown in Model (3), an extroverted social robot is anthropomorphized more than an introverted robot. More anthropomorphism is associated with more inferred agency (β=0.173, 95% CI [0.12, 0.22]; see Model (4)) and experience (β=0.167, 95% CI [0.12, 0.22]; see Model (5)), thus confirming H2a and b. Both inferred agency (β=0.432, 95% CI [0.28, 0.57]) and experience (β=0.124, 95% CI [0.04, 0.21]) increase the ascription of robot empathic capabilities (see Model (6)). Thus,

H3a and b are accepted. Contrary to expectations, the mediation was not fully mediated by agency and experience, leaving a significant direct effect between anthropomorphism and ascribed robot empathy (β=0.184, 95% CI [0.11, 0.26]). This means that another factor, captured by anthropomorphism but not by agency and experience, affects robot empathy.

The ascribed robot empathy has a positive significant effect on task satisfaction (β=0.355, 95% CI [0.27, 0.45]), explains additional 16% of the variability between robot extroversion and task satisfaction, and explains large effects observed influencing task satisfaction (Cohen 1988).

The path from robot extroversion → anthropomorphism → agency → robot empathy → task satisfaction reveals a positive significant indirect effect (β=0.017, 95% CI [0.007, 0.031]). The relationship robot extroversion → anthropomorphism → experience → robot empathy → task satisfaction also reveals a significant indirect effect (β=0.005, 95% CI [0.001, 0.010]), while no direct relationship between robot extroversion and task satisfaction exists (β = −0.044, 95% CI [−0.18, 0.11]). The partial

Table 3 Regression results for interaction and task satisfaction with mediators

	Model (4)	Model (5)	Model (6)	Model (7)
	DV: Agency	DV: Experience	DV: Robot empathy	DV: Task satisfaction
Participant's Gender	**0.167** [0.028; 0.308]	**−0.166** [−0.332; −0.004]	0.043 [−0.143; 0.218]	**0.199** [0.044; 0.347]
Attitude towards robots	**0.080** [0.020; 0.143]	**0.080** [0.021; 0.141]	0.061 [−0.028; 0.153]	**0.119** [0.053; 0.187]
Human extroversion	−0.001 [−0.053; 0.050]	0.036 [−0.013; 0.085]	−0.023 [−0.089; 0.042]	**0.067** [0.009; 0.124]
Anthropomorphism	**0.173** [0.120; 0.224]	**0.167** [0.116; 0.222]	**0.184** [0.112; 0.259]	
Experience			**0.124** [0.035; 0.211]	
Agency			**0.432** [0.282; 0.569]	
Robot extroversion				−0.044 [−0.184; 0.105]
Robot empathy				**0.355** [0.268; 0.449]
Intercept	**2.770** [2.224; 3.283]	0.345 [−0.097; 0.758]	**2.287** [1.598; 3.012]	**2.741** [2.091; 3.334]
Observations	401	401	401	401
R^2	0.167	0.118	0.292	0.265
Effect size (f^2)	0.200	0.134	0.412	0.361
X^2	58.011 *** (df=7; 393)			

Robot extroversion: 1-introversion, 2-extroversion

Bold estimates exclude zero in the 95% confidence interval (CI).

*p <.05; **p <.01; ***p <.001.

standardized effect, as an effect size measure (Hayes 2018), reveals that collaborations with extroverted social robots results in a 2.5% difference in standard deviation in task satisfaction due to the total indirect effects compared to introverted robots. Thus, the effect from robot extroversion is explained through the MMAP, whereby anthropomorphism results in inferred agency and experience, and ascribed empathy. Thus, confirming H4. The results of the analysis and the confirmation of hypotheses are summarized in Fig. 3 and in the Supplementary Materials 1.7.

The Uncanny Valley theory (Mori 1970) suggests that affective responses do not follow a linear trend dependent on the morphology of a non-human agent. We performed a robustness test with a squared anthropomorphism term. The results revealed the same power, significance, and directionality of the effects with a weaker anthropomorphism estimate. Hence revealing no non-linear effects present within the observations.

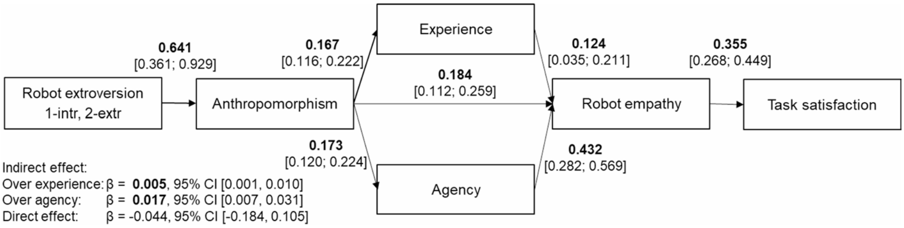

Fig. 3 Results of analysis

6 Discussion and Implications

This study aimed to explain how a social robot depicting extroverted or introverted personality cues influence users' affective outcomes (i.e., task satisfaction). For this purpose, we conceptualized the MMAP model, which integrates fragmented theoretical approaches of mind perception theories to explain the causal mechanism. Our results show that collaboration with extroverted social robots leads to more task satisfaction, because extroverted robots are more anthropomorphized than introverted robots, which results in the inference of more agency and experience, as well as the ascription of more empathic capabilities, which positively affects user satisfaction.

6.1 Theoretical Implications

Our findings contribute to research on anthropomorphism and ToM. This study provided a theoretical integration of theories and concepts connected to anthropomorphism. The conceptual works on mind perception (e.g., Airenti 2015; Epley and Waytz 2010) and ToM (Premack and Woodruff 1978) guided the theoretical synthesis resulting in a parsimonious conceptual model of mental model attribution, which can be empirically tested. Prior theories overlap and partially explain the MMAP yet do not fully account for the sequential attribution processes captured by the MMAP. Neither the three-factor theory of anthropomorphism (Epley et al. 2007) nor the recently proposed four-dimensional model of anthropomorphism by Chi et al. (2025) satisfy the complexity that is inherent to the conceptual understanding of the mental state attribution process. The proposed MMAP accounts for the sequential nature of mind attribution processes when observing social cues of non-human agents.

Our findings also contribute to research on the personality design of robots. Our findings explain why an extroverted robot leads to more task satisfaction than an introverted one—because extroversion cues facilitate the MMAP. Extroverted social robots are characterized by confidence, sociability, playfulness, outgoingness, and energetic behavior (Fong et al. 2003; Goldberg 1990). Such characteristics lead the extroverted robot to be understood as more of an intentional agent than an introverted robot. An introverted personality is often characterized by silence, shyness, and passivity (Goldberg 1990), so it may appear quite machine-like. Our findings advance this existing research stream by demonstrating that robot extroversion-introversion design has far-reaching consequences, including its impact on task satisfaction in team collaboration within a business environment.

6.2 Practical Implications

This work also offers recommendations for designers of social robots. The observed large effect size suggests that users tend to prefer extroverted social robots, which should be taken into consideration by designers. By designing social robots with extroverted social cues, they can facilitate higher task satisfaction in users. These cues can be materialized through movement, voice intonation, pitch, and speech phrasing (Lee et al. 2006; Mileounis et al. 2015; Park et al. 2012; Tay et al. 2014). Yet differences in perceptions of the agents can be inherent to the users' individual differences in the MMAP. Thus, diverse perceptions might not only be due to the design aspects but rather due to the users' natural tendencies of explaining non-human agents through anthropomorphism enabling the attribution of agency, experience, and capabilities.

This work also contributes recommendations for managers of teams that consider adopting a social robot as a team member. First, by suggesting the adoption of an actively participating, outgoing, extroverted social robot over an introverted social robot. This is especially helpful if human team members consider themselves extroverts. The findings reveal—under consideration of the MMAP—that extroverted humans have a more positive reaction to the team task if a social robot is present compared to introverted humans. Hence, hinting towards more openness towards social robot team members. Second, the MMAP offers insights into how to introduce the social robot and develop training materials for the robot's integration. It highlights that specifically the focus on stimulating anthropomorphism, highlighting the social robot's capabilities (i.e., agency), and affective computing characteristics (i.e., experience) can support the integration of an artificial team member.

7 Limitations and Future Work

We acknowledge three limitations of this study that provide avenues for future research. First, we used video-based vignettes (Aguinis and Bradley 2014). We designed the scenario with agile project management experts to reflect an authentic interaction that could occur in real-world projects. While these findings rely upon a short exposure and indirect exposure to the social robot, we are convinced that these findings provide initial support for the MMAP. Future research could replicate this study using a physical instantiation of a social robot to confirm the findings. Especially with technological advancements, social robots can be equipped with GenAI and therefore act as equal partners, allowing a viable collaboration.

Second, this study relied on crowd workers with partly limited knowledge on Scrum. Future research could include Scrum professionals as participants to examine the effects of robot extroversion-introversion design on satisfaction, as they may develop different judgments given their expertise.

Third, the manipulation of robot extroversion-introversion partly overlaps with how other researchers have manipulated robot empathy. For instance, we used gestures and gaze to simulate the robot personality, which was used by Charrier et al. (2018) to simulate an empathic robot. Manipulation of emotional intonation has also been used by other researchers to mimic empathic social robots (e.g., Brave et al. 2005). Since we followed existing research protocols in our manipulation of robot extroversion-introversion and our manipulation check was successful, we are confident that we effectively captured extroversion or introversion. Nevertheless, future research could delineate extroversion cues or empathy cues.

Finally, these findings call for research on other non-human agents to consider the MMAP. Any non-human agent capable of exerting social cues, such as showing attention or reacting to users, is automatically activating the MMAP to explain its behavior and predict its future actions. Thus, research on HATs investigating tools, such as AI-enabled chatbots and social robots, should consider the MMAP in explaining users' behaviors and their rationalization of the agents' behavior.

8 Conclusion

This study investigated how the extroversion-introversion design of intelligent social robots influence human task satisfaction in an agile project context. Our results indicate that extroverted robots promote more task satisfaction because users anthropomorphize them more, infer agency and experience, and attribute more human-like capabilities to extroverted robots than to introverted ones. With this, the suggested MMAP offers an integrated and testable mental model attribution model to test the effects of social robot designs. With these findings, we contribute to research on robot personality and the mind perception literature and related approaches in HAT research.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s10726-026-09990-z.

Acknowledgements This research was funded in whole by the Austrian Science Fund 10.55776/P29765. For open access purposes, the author has applied a CC BY public copyright license to any author accepted manuscript version arising from this submission.

Author Contributions All authors contributed to the study conception, design, and material preparation. Funding acquisition was performed by I.S. Data collection was performed by V.M.O., and analysis were performed by V.M.O. and I.S. The first draft of the manuscript was written by V.M.O. and I.S. All authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Funding Open access funding provided by University of Innsbruck and Medical University of Innsbruck.

Data Availability The datasets generated by the current study is protected and not available due to data privacy assurances given to the participants in the informed consent. The research team assured the participants that the data would be kept confidential and will only be shared within the research team and with reviewers.

Declarations

Conflict of interest The authors declare no competing interests.

References

Aguinis H, Bradley KJ (2014) Best practice recommendations for designing and implementing experimental vignette methodology studies. Organ Res Methods 17(4):351–371. https://doi.org/10.1177/1094428114547952

Airenti G (2015) The cognitive bases of anthropomorphism: from relatedness to empathy. Int J Soc Robot 7(1):117–127. https://doi.org/10.1007/s12369-014-0263-x

Alves-Oliveira P, Sequeira P, Melo FS, Castellano G, Paiva A (2019) Empathic robot for group learning: a field study. ACM Trans on Hum-Robot Interact 8(1):1–34. https://doi.org/10.1145/3300188

Andriella A, Siqueira H, Fu D, Magg S, Barros P, Wermter S, Torras C et al (2021) Do I Have a Personality? endowing care robots with context-dependent personality traits. Int J Soc Robot. https://doi.org/10.1007/s12369-020-00690-5

Andrist S, Mutlu B, Tapus A (2015) Look like me: matching robot personality via gaze to increase motivation. Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (CHI '15) 3603–3612. https://doi.org/10.1145/2702123.2702592

Andriella A, Torras C, Alenyà G (2025) Unveiling robot personality: impact on user performance and acceptance. In: Personalising human-robot interactions in social contexts, Springer, Cham

Baird A, Maruping LM (2021) The next generation of research on IS use: a theoretical framework of delegation to and from agentic IS artifacts. MISQ 45(1):315–341. https://doi.org/10.25300/MISQ/2021/15882

Baltrusch SJ, Krause F, De Vries AW, Van Dijk W, De Looze MP (2022) What about the human in human robot collaboration? Ergon 65(5):719–740. https://doi.org/10.1080/00140139.2021.1984585

Banerjee S, González-Jiménez H, Zheng L (2025) Help please! deriving social support from Geminoid DK pepper and AIBO as companion robots. Int J Human Comput Stud. https://doi.org/10.1016/j.ijhcs.2025.103577

Bartneck C, Kulic D, Croft E, Zoghbi S (2009) Measurement instruments for the anthropomorphism, animacy, likeability, perceived intelligence, and perceived safety of robots. Int J Soc Robot 1:71–81. https://doi.org/10.1007/s12369-008-0001-3

Blut M, Wang C, Wünderlich NV, Brock C (2021) Understanding anthropomorphism in service provision: a meta-analysis of physical robots, chatbots, and other AI. J Acad Mark Sci 49:632–658. https://doi.org/10.1007/s11747-020-00762-y

Brave S, Nass C, Hutchinson K (2005) Computers that care: investigating the effects of orientation of emotion exhibited by an embodied computer agent. Int J Hum Comput Stud 62(2):161–178. https://doi.org/10.1016/j.ijhcs.2004.11.002

Chang W, Kim KK (2022) Appropriate service robots in exchange and communal relationships. J Bus Res 141:462–474. https://doi.org/10.1016/j.jbusres.2021.11.044

Charrier L, Galdeano A, Cordier A, Lefort M (2018) Empathy Display Influence on Human-Robot Interactions: a Pilot Study. Workshop on Towards Intelligent Social Robots: From Naive Robots to Robot Sapiens at the 2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2018) 1–7. https://hal.science/hal-01887075/

Chi OH, Chi CG, Gursoy D (2025) Seeing personhood in machines: conceptualizing anthropomorphism of social robots. J Serv Res. https://doi.org/10.1177/10946705241297196

Cohen JE (1988) Statistical power analysis for the behavioral sciences. Routledge, New York

Damiano L, Dumouchel P (2018) Anthropomorphism in human–robot co-evolution. Front Psychol. https://doi.org/10.3389/fpsyg.2018.00468

De Graaf MMA, Malle BF (2019) People's explanations of robot behavior subtly reveal mental state inferences. 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI) 239–248. https://doi.org/10.1109/HRI.2019.8673308

Diefenbach S, Herzog M, Ullrich D, Christoforakos L (2023) Social robot personality: a review and research agenda. In: Misselhorn C, Poljansek T, Störzinger T, Klein M (eds) Emotional machines. Perspectives from affective computing and emotional human-machine interaction. Springer VS, Wiesbaden

Douglas DB, Ewell JP, Brauer M (2023) Data quality in online human-subjects research: comparisons between MTurk, Prolific, CloudResearch, Qualtrics, and SONA. PLoS One. https://doi.org/10.1371/journal.pone.0279720

Edwards AD, Shafer DM, Edwards AD, Shafer DM (2022) When lamps have feelings. Projections. https://doi.org/10.3167/proj.2022.160202

Epley N, Waytz A (2010) Mind perception. In: Handbook of social psychology, Wiley, 498–541. https://doi.org/10.1002/9780470561119.socpsy001014

Epley N, Waytz A, Cacioppo JT (2007) On seeing human: a three-factor theory of anthropomorphism. Psychol Rev 114(4):864–886. https://doi.org/10.1037/0033-295X.114.4.864

Esterwood C, Essenmacher K, Yang H, Zeng F, Robert LP (2022) A personable robot: meta-analysis of robot personality and human acceptance. IEEE Robot Autom Lett 7(3):6918–6925. https://doi.org/10.1109/lra.2022.3178795

Eyssel F, Kuchenbrandt D, Bobinger S, de Ruiter L, Hegel F (2012) 'If You Sound Like Me, You Must Be More Human': on the interplay of robot and user features on human-robot acceptance and anthropomorphism. Proceedings of the 7th annual ACM/IEEE international conference on Human-Robot Interaction (HRI '12) 125–126. https://doi.org/10.1145/2157689.2157717

Fernández-Llamas C, Conde MA, Rodríguez-Lera FJ, Rodríguez-Sedano FJ, García F (2018) May I Teach You? Students' behavior when lectured by robotic vs. Human teachers. Comput in Hum Behav 80:460–469. https://doi.org/10.1016/j.chb.2017.09.028

Fong T, Nourbakhsh I, Dautenhahn K (2003) A survey of socially interactive robots: concepts, design, and applications. Robot Auton Syst 42(3–4):143–166. https://doi.org/10.1016/S0921-8890(02)00372-X

Frith C, Frith U (2005) Theory of mind. Curr Biol 15(17):644–645. https://doi.org/10.1016/j.cub.2005.08.041

Gambino A, Fox J, Ratan RA (2020) Building a stronger CASA: extending the computers are social actors paradigm. Hum-Mach Commun. https://doi.org/10.3316/INFORMIT.097034846749023

Goldberg LR (1990) An alternative "description of personality": the big-five factor structure. J Pers and Psychol 59(6):1216–1229. https://doi.org/10.1037/0022-3514.59.6.1216

Gray K, Wegner DM (2012) Feeling robots and human zombies: mind perception and the uncanny valley. Cognition 125:125–130. https://doi.org/10.1016/j.cognition.2012.06.007

Gray HM, Gray K, Wegner DM (2007) Dimensions of mind perception. Science 315:619–619. https://doi.org/10.1007/s11097-010-9186-7

Gray K, Knobe J, Sheskin M, Bloom P, Barrett LF (2011) More than a body: mind perception and the nature of objectification. J Pers Soc Psychol 101(6):1207. https://doi.org/10.1037/a0025883

Guidi S, Bracci M, Currò F, Innocenti A, Lusuardi L, Marchigiani E, Palmitesta P et al (2025) You look so young, you look so cute. The relationship between physical appearance, age and mental abilities in social robots. Behav Inf Technol 44(6):1136–1145. https://doi.org/10.1080/0144929X.2025.2457563

Hair JF, Black WC, Babin BJ, Anderson RE (2010) Multivariate data analysis: a global perspective, 7th edn. Pearson Education, London

Hayes AF (2018) Introduction to mediation, moderation, and conditional process analysis: A regression-based approach, 2 edn. The Guilford Press, New York

Heerink M, Kröse B, Evers V, Wielinga B (2010) Assessing acceptance of assistive social agent technology by older adults: the Almere model. Int J Soc Robot 2:361–375. https://doi.org/10.1007/s12369-010-0068-5

International Personality Item Pool (2018). Big-Five Factor Markers. https://ipip.ori.org/newBigFive5broadKey.htm

Kim D, Kim S (2021) A model for user acceptance of robot journalism: influence of positive disconfirmation and uncertainty avoidance. Technol Forecast Soc Change 163:120448. https://doi.org/10.1016/j.techfore.2020.120448

Lee KM, Peng W, Jin S-a, Yan C (2006) Can robots manifest personality? An empirical test of personality recognition, social responses, and social presence in human–robot interaction. J Commun 56(4):754–772. https://doi.org/10.1111/j.1460-2466.2006.00318.x

Leslie AM, Friedman O, German TP (2004) Core mechanisms in 'theory of mind.' Trends Cogn Sci 8(12):528–533. https://doi.org/10.1016/j.tics.2004.10.001
Long JS, Ervin LH (2000) Using heteroscedasticity consistent standard errors in the linear regression model. Am Stat 53(4):217–224. https://doi.org/10.1080/00031305.2000.10474549
Lu L, Zhang P, Zhang T (2021) Leveraging "human-likeness" of robotic service at restaurants. Int J Hosp Manage. https://doi.org/10.1016/j.ijhm.2020.102823
Malle BF, Knobe J (1997) The folk concept of intentionality. J Exp Soc Psychol. https://doi.org/10.1006/jesp.1996.1314
Mara M, Appel M (2015a) Effects of lateral head tilt on user perceptions of humanoid and android robots. Comput Hum Behav 44:326–334. https://doi.org/10.1016/j.chb.2014.09.025
Mara M, Appel M (2015b) Science fiction reduces the eeriness of android robots: a field experiment. Comput Hum Behav 48:156–162. https://doi.org/10.1016/j.chb.2015b.01.007
Mathur MB, Reichling DB (2016) Navigating a social world with robot partners: a quantitative cartography of the Uncanny Valley. Cognition 146:22–32. https://doi.org/10.1016/j.cognition.2015.09.008
Mathur MB, Reichling DB, Lunardini F, Geminiani A, Antonietti A, Ruijten PAM, Levitan CA et al (2020) Uncanny but not confusing: multisite study of perceptual category confusion in the Uncanny Valley. Comput Hum Behav 103:21–30. https://doi.org/10.1016/j.chb.2019.08.029
Matza LS, Stewart KD, Lloyd AJ, Rowen D, Brazier JE (2021) Vignette-based utilities: usefulness, limitations, and methodological recommendations. Value Health. https://doi.org/10.1016/j.jval.2020.12.017
Medina MN (2016) Conflict, individual satisfaction with team, and training motivation. Team Perform Manag 22(3/4):223–239. https://doi.org/10.1108/tpm-10-2015-0047
Mileounis A, Cuijpers RH, Barakova EI (2015) Creating robots with personality: the effect of personality on social intelligence. Artificial Computation in Biology and Medicine. IWINAC 2015. Lecture Notes in Computer Science. https://doi.org/10.1007/978-3-319-18914-7_13
Mitchell RLC, Phillips LH (2015) The overlapping relationship between emotion perception and theory of mind. Neuropsychologia. https://doi.org/10.1016/j.neuropsychologia.2015.02.018
Mori M (1970) The Uncanny Valley. Energy 7(4):33–35
Nardelli A, Maccagni G, Minutoli F, Sgorbissa A, Recchiuto C (2025) Towards intuitive interaction: cognitive architecture for artificial personality, emotional intelligence, and cognitive capabilities. Int J Soc Robot 17(10):2211–2228. https://doi.org/10.1007/s12369-025-01260-3
Nass C, Lee KM (2001) Does computer-synthesized speech manifest personality? experimental tests of recognition, similarity-attraction, and consistency-attraction. J Exp Psychol Appl 7(3):171–181
Oberhofer VM, Seeber I, Maier R (2023) Human-robot interaction: mapping literature review and network analysis. Proceedings of the 56th Hawaii International Conference on System Sciences 481–490. https://hdl.handle.net/10125/102689
Paetzel-Prüsmann M, Perugia G, Castellano G (2021) The influence of robot personality on the development of uncanny feelings. Comput Hum Behav 120:106756. https://doi.org/10.1016/j.chb.2021.106756
Park E, Jin D, Del Pobil AP (2012) The law of attraction in human-robot interaction. Int J Adv Robot Syst. https://doi.org/10.5772/50228
Peer E, Rothschild D, Gordon A, Evernden Z, Damer E (2022) Data quality of platforms and panels for online behavioral research. Behav Res Methods 54:1643–1662. https://doi.org/10.3758/s13428-021-01694-3
Phillips E, Ososky S, Grove J, Jentsch F (2011) From tools to teammates: the development of appropriate mental models for intelligent robots. Proceedings of the Human Factors and Ergonomics Society Annual Meeting 1491–1495. https://doi.org/10.1177/1071181311551310
Pitt LF, Watson RT, Kavan CB (1995) Service quality: a measure of information systems effectiveness. MIS Q 19(2):173–187. https://doi.org/10.2307/249687
Podsakoff PM, Podsakoff NP, Williams LJ, Huang C, Yang J (2024) Common method bias: it's bad, it's complex, it's widespread, and it's not easy to fix. Annu Rev Organ Psychol Organ Behav 11:17–61. https://doi.org/10.1146/annurev-orgpsych-110721-040030
Premack D, Woodruff G (1978) Does the chimpanzee have a theory of mind? Behav Brain Sci 1(4):515–526. https://doi.org/10.1017/S0140525X00076512
Qin M, Li S, Zhu W, Qiu S (2025) Trust in service robot: the role of appearance anthropomorphism. Curr Issues Tour 28(1):36–54. https://doi.org/10.1080/13683500.2023.2295923
Robert L (2018) Personality in the human robot interaction literature: a review and brief critique. 24th Americas Conference on Information Systems 1–10. http://aisel-1aisnet-1org-1qnpmi2j64554.han.sub.uni-goettingen.de/cgi/viewcontent.cgi?article=1080&context=amcis2018

Robert LP, You S (2018) Are you satisfied yet? shared leadership, individual trust, autonomy, and satisfaction in virtual teams. J Assoc Inf Science and Technol 69(4):503–513. https://doi.org/10.1002/asi.23983

Rosenthal-Von Der Pütten AM, Krämer NC (2014) How design characteristics of robots determine evaluation and uncanny valley related responses. Comput Human Behav 36:422–439. https://doi.org/10.1016/j.chb.2014.03.066

Salem M, Eyssel F, Rohlfing K, Kopp S, Joublin F (2013) To err is human(-like): effects of robot gesture on perceived anthropomorphism and likability. Int J Soc Robot 5:313–323. https://doi.org/10.1007/s12369-013-0196-9

Schaafsma SM, Pfaff DW, Spunt RP, Adolphs R (2015) Deconstructing and reconstructing theory of mind. Trends Cogn Sci 19(2):65–72. https://doi.org/10.1016/j.tics.2014.11.007

Schlinger HD (2009) Theory of mind: an overview and behavioral perspective. Psychol Rec 59(3):435–448. https://doi.org/10.1007/BF03395673

Seo S (2022) When female (male) robot is talking to me: effect of service robots' gender and anthropomorphism on customer satisfaction. Int J Hosp Manag 102:103166. https://doi.org/10.1016/j.ijhm.2022.103166

Seo KH, Lee JH (2021) The emergence of service robots at restaurants: integrating trust, perceived risk, and satisfaction. Sustainability 13(8):4431. https://doi.org/10.3390/su13084431

Simmering MJ, Fuller CM, Richardson HA, Ocal Y, Atinc GM (2015) Marker variable choice, reporting, and interpretation in the detection of common method variance: a review and demonstration. Organ Res Methods 18(3):473–511. https://doi.org/10.1177/1094428114560023

Song Y, Luximon A, Luximon Y (2021) The effect of facial features on facial anthropomorphic trustworthiness in social robots. Appl Ergon 94:103420. https://doi.org/10.1016/j.apergo.2021.103420

Steiner PM, Atzmüller C, Su D (2016) Designing valid and reliable vignette experiments for survey research: a case study on the fair gender income gap. J Methods Meas Soc Sci 7(2):53–94. https://doi.org/10.2458/v7i2.20321

Strich F, Mayer A-S, Fiedler M (2021) What do i do in a world of artificial intelligence? Investigating the impact of substitutive decision-making AI systems on employees' professional role identity. J Assoc Inf Syst 22(2):304–324. https://doi.org/10.17705/1jais.00663

Sun Y, Xiao C (2025) Uncanny valley curve or linear relationship? The different influences of robots' human-likeness on mind perception and visuospatial perspective-taking. Int J Hum-Comput Interact 41(18):11392–11404. https://doi.org/10.1080/10447318.2024.2443251

Szczepanowski R, Cichoń E, Arent K, Sobecki J, Styrkowiec P, Florkowski M, Gakis M (2020) Education biases perception of social robots. Rev Eur Psychol Appl 70(2):100521. https://doi.org/10.1016/j.erap.2020.100521

Tay B, Jung Y, Park T (2014) When stereotypes meet robots: the double-edge sword of robot gender and personality in human – robot interaction. Comput Hum Behav 38:75–84. https://doi.org/10.1016/j.chb.2014.05.014

Weiss DJ, Dawis RV, England GW (1967) Manual for the Minnesota satisfaction questionnaire. J Appl Psychol. https://doi.org/10.1037/t05540-000

Williams LJ, Hartman N, Cavazotte F (2010) Method variance and marker variables: a review and comprehensive CFA marker technique. Organ Res Methods 13(3):477–514. https://doi.org/10.1177/1094428110366036

Wooldridge M, Jennings NR (1995) Intelligent agents: theory and practice. Knowl Eng Rev 10(2):115–152. https://doi.org/10.1017/S0269888900008122

Yam KC, Bigman Y, Gray K (2021) Reducing the uncanny valley by dehumanizing humanoid robots. Comput Hum Behav 125:106945. https://doi.org/10.1016/j.chb.2021.106945

Yörük T, Akar N, Özdem NV (2024) Research trends on guest experience with service robots in the hospitality industry: a bibliometric analysis. Eur J Innov Manag 27(6):2015–2041. https://doi.org/10.1108/EJIM-09-2022-0530

You S, Robert LP (2018) Emotional attachment, performance, and viability in teams collaborating with embodied physical action (EPA) robots. J Assoc Inf Syst 19(5):377–407. https://doi.org/10.17705/1jais.00496