Computers in Human Behavior: Artificial Humans 3 (2025) 100131

Contents lists available at ScienceDirect

Computers in Human Behavior: Artificial Humans

journal homepage: www.journals.elsevier.com/computers-in-human-behavior-artificial-humans

Beyond the monotonic: Enhancing human-robot interaction through affective communication

Kim Klüber [a,*], Linda Onnasch [b]

[a] Humboldt-Universität zu Berlin, Berlin, Germany
[b] Technische Universität Berlin, Berlin, Germany

ARTICLE INFO

Keywords:
Human-robot interaction
Affective computing
Verbal communication
Emotional expression
Trustworthiness
Forgiveness
Mind attribution

ABSTRACT

As robots increasingly become part of human environments, their ability to convey empathy and emotional expression is critical for effective interaction. While non-verbal cues, such as facial expressions and body language, have been widely researched, the role of verbal communication - especially affective speech - has received less attention, despite being essential in many human-robot interaction scenarios. This study addresses this gap through a laboratory experiment with 157 participants, investigating how a robot's affective speech influences human perceptions and behavior. To explore the effects of varying intonation and content, we manipulated the robot's speech across three conditions: monotonic-neutral, monotonic-emotional, and expressive-emotional. Key measures included attributions of experience and agency (following the Theory of Mind), perceived trustworthiness (cognitive and affective level), and forgiveness. Additionally, the Balloon Analogue Risk Task (BART) was employed to assess dependence behavior objectively, and a teaching task with intentional robot errors was used to measure behavioral forgiveness. Our findings reveal that emotionally expressive speech enhances the robot's perceived capacity for experience (i.e., the ability to feel emotions) and increases affective trustworthiness. The results further suggest that affective content of speech, rather than intonation, is the decisive factor. Consequently, in future robotic applications, the affective content of a robot's communication may play a more critical role than the emotional tone. However, we did not find significant differences in dependence behavior or forgiveness across the varying levels of affective communication. This suggests that while affective speech can influence emotional perceptions of the robot, it does not necessarily alter behavior.

1. Introduction

"*Affective computing is inside robots, voice assistants, automobiles, smartphones, and all kinds of technology that we interact with daily*," states Rosalind Picard,[1] who originally coined the term 'affective computing' in the late 1990s (Picard, 2000). Today, with the rapid advancement of artificial intelligence, machine learning, and sensor technologies, this field has gained renewed significance. Affective computing aims to equip machines with the ability to recognize, interpret, and respond to human emotions or simulate them in technical systems (Picard, 2000).

As robots and AI systems become more integrated into our daily lives, their ability to engage with us on an emotional level is no longer just a technical challenge, but a crucial aspect in human environments. By enabling computers and robots to process emotional cues - whether through facial expressions, voice tones, or physiological signals - these systems can react in ways that are increasingly indistinguishable from human interactions. Whether in healthcare, customer service, or personal companionship, the capability of robots to exhibit empathy and emotional understanding can significantly influence how they are perceived and trusted by humans.

1.1. *Affective communication of robots*

As machines and robotic systems become increasingly human-like, the expectation for them to interact in a human-like manner, including the expression of emotions, grows as well. While facial expressions and body language have been extensively studied for conveying emotions, verbal communication has not been researched as thoroughly but is equally significant because it plays a crucial role in many human-robot interaction (HRI) scenarios (Gasteiger et al., 2024).

* Corresponding author. Humboldt-Universität zu Berlin, Department of Psychology, 10099 Berlin, Germany.
E-mail address: kim.ingrid.klueber@student.hu-berlin.de (K. Klüber).

[1] Statement retrieved from: https://www.media.mit.edu/groups/affective-computing/overview/.

https://doi.org/10.1016/j.chbah.2025.100131
Received 15 November 2024; Received in revised form 15 February 2025; Accepted 15 February 2025
Available online 18 February 2025

The human voice is a powerful communicative tool that conveys critical information about its speaker. Research has shown that even without visual cues, people can infer characteristics such as gender, age, emotional state, and social group from voice alone (McGinn & Torre, 2019; Schreibelmayr & Mara, 2022). Within HRI, participants generally prefer robots that communicate in a more human-like manner (Dautenhahn et al., 2005). For example, Kühne et al. (Kühne et al., 2020) found a clear preference for human-like voices over synthetic ones in robots, underscoring the significant impact of a robot's speech on user perception.

Affective communication, defined as the expression of emotions through speech, can be considered in a nuanced way, encompassing both the content of speech (semantics) and prosodic features like intonation, pitch, volume, pauses, and speed (James et al., 2021; Taylor, 2009). Through variations in these semantic and prosodic elements, a wide range of emotions can be expressed and perceived. The importance of prosodic elements in conveying emotions effectively in robot speech has been highlighted in several studies, which have generally reported successful outcomes (Gasteiger et al., 2024).

For instance, Niculescu et al. (Niculescu et al., 2013) demonstrated that altering the pitch of speech can significantly affect perceptions. In their study, a robot with a higher-pitched, feminine voice was rated as more attractive, socially skilled, and engaging, suggesting that prosodic elements significantly influence how human qualities are projected onto robots.

Moreover, expressing emotions through voice has been shown to enhance behavioral outcomes in HRI (Tielman et al., 2014). Leyzberg et al. (Leyzberg et al., 2011) demonstrated that robots providing contextually appropriate emotional responses could elicit more accurate behaviors from participants than those using inappropriate responses. These findings emphasize the importance of emotional communication in shaping successful HRI.

In a recent study, Cucciniello et al. (Cucciniello et al., 2023) explored how different behavioral styles of robots impact users' attribution of mental and emotional states. By manipulating prosodic and semantic aspects such as speed, pitch, and language or feedback information, the researchers developed three behavioral styles: friendly, neutral, and authoritarian. Their findings showed that the friendly style, which included encouraging language and a warmer tone, was more likely to lead users to attribute positive emotions to the robot compared to the other styles. Additionally, they found that these interaction styles influenced agency perceptions and attribution of cognitive capacities.

1.2. *The role of anthropomorphism in mind attribution*

The attribution of emotional and mental capacities to a robot as reported by Cucciniello et al. (Cucciniello et al., 2023) is an example for people's tendency to anthropomorphize non-human entities (Epley et al., 2007). Epley and colleagues (Epley et al., 2007) suggest that entities that possess easily recognizable human-like features are more likely to be anthropomorphized. This occurs because features such as appearance, speech, movement, and context (Onnasch & Roesler, 2021) activate human-centric thinking. Anthropomorphism facilitates understanding of robots and other entities by encouraging people to apply familiar social norms to these interactions, especially when robots engage in human-like roles (e.g., teacher, therapist) or use natural speech (Schreibelmayr & Mara, 2022). Language thus emerges as a key dimension through which human-like features can shape interactions.

This is also reflected in Mind Perception Theory (Gray et al., 2007), which explains how people attribute mind to other humans as well as non-human entities like robots. The theory differentiates between two dimensions: experience - the attribution of feelings, consciousness, or emotions like joy or fear and agency - the attribution of intentional thought, responsibility, and actions.

When robots communicate with emotional expressions, they are more likely to be attributed with these dimensions of mind (Cucciniello et al., 2023). Previous studies have shown that anthropomorphic robots, particularly those with human-like voices, are assigned with more "mind" as they provide strong social cues that make people treat them as human-like entities (McGinn & Torre, 2019; Waytz et al., 2010). As a result, people tend to predict the behavior of these robots similarly to how they predict human behavior. While increases in perceived agency are typically associated with a greater attribution of responsibility, our focus is on enhancing the experience dimension to evoke more sympathy toward the robot. Evidence suggests that a robot with emotionally expressive communication is perceived as more capable of experiencing emotions (Cucciniello et al., 2023). By increasing the experience dimension, we aim to increase perceived trustworthiness and the likelihood of forgiveness for errors.

1.3. *Trustworthiness and forgiveness in HRI*

The perceived trustworthiness of a robot is crucial for people's trust and subsequent dependence behavior in HRI.[2] According to Malle and Ullman (Malle & Ullman, 2021), trustworthiness consists of two dimensions: performance (capable, reliable) and moral (ethical, sincere). The trustworthiness of robots is influenced by several factors, including design (Bernotat et al., 2021; Roesler et al., 2021), behavior during interactions (Hancock et al., 2011; Okafuji et al., 2024), and the context of use (Roesler, Naendrup-Poell, et al., 2022; Schreibelmayr & Mara, 2022). Unmet expectations or unpredictable behavior can quickly erode perceived trustworthiness (Gompei et al., 2018).

A key factor in maintaining positive interactions is forgiveness, which allows users to reassess a robot's trustworthiness over time (which might change based on learning or contextual changes). If users are unwilling to forgive mistakes, they may disengage entirely. Forgiveness is a complex construct encompassing affective, cognitive, motivational, and behavioral components and plays a critical role in sustaining long-term human-robot relationships (Salem et al., 2015). By perceiving a robot as trustworthy in combination with a general willingness to forgive it for mistakes, which are inevitable in real-world applications, interactions are more likely to continue rather than be abandoned. This, in turn, creates more learning opportunities for both the robot and the user, allowing the system to adapt and improve over time.

Factors such as empathy toward the robot or its communication style are key in fostering trustworthiness and forgiveness (Carlisle et al., 2012; Hancock et al., 2021), especially when emotional or non-emotional aspects are used appropriately. This underscores the importance of carefully selecting a robot's communication style and further exploring how affective communication influences various aspects of HRI.

1.4. *Research framework and hypotheses*

Affective communication has been shown to enhance both, the perception of robots (e.g., making them appear more engaging or socially skilled (Gasteiger et al., 2024; Niculescu et al., 2013)) and to improve the behavioral outcomes during interactions (e.g., increased accuracy in a dance teaching task and more frequent positive expressions (Leyzberg et al., 2011; Tielman et al., 2014)).

Given this, our research investigates the impact of affective robot communication with three main objectives: First, we examine how affective communication influences mind attribution. Second, we explore its impact on trustworthiness and dependence behavior. Third, we

[2] There has been some confusion about trustworthiness and trust as both concepts appear very similar. In consequence, a lot of "trust" research in HRI actually is about trustworthiness. We stick, however, to the conceptual distinction and the term trustworthiness, even if authors have claimed to investigate trust (e.g. (Bernotat et al., 2021),).

investigate how forgiving people are towards a robot that makes a mistake. The research framework guiding this study is illustrated in Fig. 1, which outlines the three key factors. Each of these factors is associated with hypotheses derived from relevant literature and theoretical insights that are presented in the following.

Based on prior research, verbal communication can enhance mind attribution to robots (Cucciniello et al., 2023; McGinn & Torre, 2019) influencing perceptions along two dimensions (Gray et al., 2007): experience and agency. Given that affective robot speech emphasizes emotional expression, we hypothesize that it will specifically influence the **experience dimension** of mind attribution, which is tied to emotional traits. In contrast, the **agency dimension**, associated with intentional actions, is expected to remain unaffected.

H1. *Affective robot communication leads to higher ratings on the experience dimension of mind attribution to robots while the agency dimension is not affected.*

Based on Bernotat and colleagues (Bernotat et al., 2021), **affective trustworthiness** is rooted in the emotional bond of the interaction, rather than the robot's objective reliability or competence, which are the foundations of **cognitive trustworthiness**. Since affective communication should target the emotional aspects of the interaction, it is expected to elevate affective trustworthiness without significantly altering cognitive trustworthiness. Moreover, as affective communication fosters emotional closeness, it is likely to lead to higher **dependence behavior** towards the robot. This theoretical framework leads to the following hypotheses.

H2a. *Affective robot communication leads to higher perceived affective trustworthiness compared to no / moderate affective communication but not higher perceived cognitive trustworthiness of the robot.*

H2b. *Affective robot communication leads to higher dependence behavior towards the robot compared to no / moderate affective communication.*

Building on our hypotheses regarding mind perception and trustworthiness, we explore an additional exploratory question (E1). Previous research suggests that entities with higher levels of anthropomorphism, such as human-like voices, are attributed more mind (McGinn & Torre, 2019; Waytz et al., 2010). Entities that score higher on mind attribution (i.e., on the experience and agency dimensions) tend to be perceived as more trustworthy, competent, and emotionally capable (Gray et al., 2007).

Building on this, the question arises as to *whether the experience dimension of mind attribution mediates the effect of affective communication on perceived trustworthiness in the robot (**E1**)?*

The third objective of this study is to investigate how forgiving people are towards a robot that makes a mistake, therefore focusing on episodic forgiveness. Episodic forgiveness refers to forgiveness that is tied to a specific event or offense (Fernández-Capo et al., 2017). In this context, we examine both the **affective component of forgiveness**, which involves emotional responses to the robot's mistake, and the **behavioral component**, which is measured through actions that indicate forgiveness, similar to previous studies where forgiveness was assessed through behaviors like doing a favor (Zechmeister et al., 2004). Based on this framework, we propose the following hypotheses.

H3a. *After an error, humans will be more forgiving with a robot speaking with affective communication than without or moderate affective communication (assessed with self-ratings).*

H3b. *After an error, humans show a greater forgiveness-related behavior towards a robot speaking with affective communication than without or moderate affective communication.*

We are also interested in investigating whether participants' **engagement** is influenced by affective communication. Our second exploratory question (E2) is therefore, whether affective communication in conversational settings should foster a more natural and human-like interaction, encouraging participants to feel more at ease and engaged. However, engagement is a multifaceted concept that varies depending on the context and interaction partners. In conversational settings, engagement refers to a concept called conversational involvement (Coker & Burgoon, 1987).

Previous studies, such as Kiesler et al. (Kiesler et al., 2008), found that participants were more engaged during interactions with embodied robots, as indicated by longer conversations. Sanghvi et al. (Sanghvi et al., 2011) observed social engagement by analyzing the affective body postures of users interacting with the iCat robot. However, both the definition of engagement and the methodology for measuring it are challenging endeavors with no standardized frameworks for annotation (Oertel et al., 2020).

Considering these findings, we aimed to explore *whether affective robot communication leads to more engagement (length of given answers) with the robot compared to no/moderate affective communication (**E2**)?*

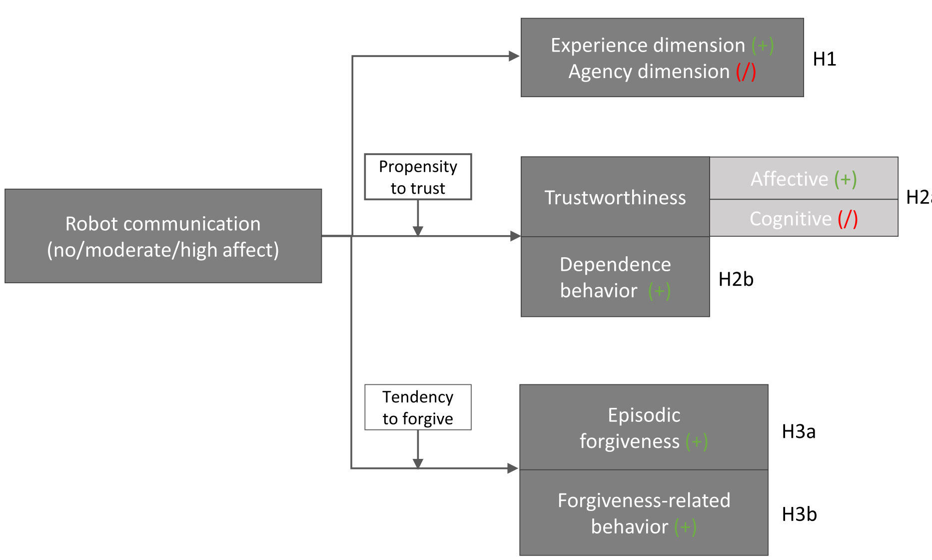

Fig. 1. Research framework on the relationships between affective robot communication and the factors of mind perception, trustworthiness/dependence behavior, and forgiveness, along with possible covariates and the hypotheses associated with each factor. *Note:* "+" indicates that affective communication influences the respective dependent variable, while "/" indicates no effect on the dependent variable.

2. Method

This study was preregistered at the Open Science Framework (OSF) where the raw data of the study is available (https://osf.io/2uh6t/?view_only=4b385f68429e4ca58e0a34b7c06e7de8). The experiment was performed with local ethical committee approval and in accordance with the ethical standards stated in the 1964 Declaration of Helsinki.

2.1. Participants

A sample size of 156 participants was targeted to achieve approximately .8 power for detecting a medium effect size at the standard .05 alpha error probability. Ultimately, 159 participants were recruited for the laboratory study, with useable data from 157 participants. The average age of participants was 29 years ($SD = 14$), with ages ranging from 18 to 79 years. The majority of the participants identified themselves as female ($n = 91$), followed by male ($n = 61$) and non-binary ($n = 1$). Fluency in German was the only requirement for participation. For participating in the study, participants were compensated either with two course credits or financially with a minimum of €5, with the opportunity to earn additional money during one of the study tasks.

2.2. Design

The experiment followed a between-subjects design, varying the affective quality of the robot's speech across three conditions: *no affect*, *moderate affect*, and *high affect* (see Table 1 for details on intonation and content of each group). To investigate a natural progression in affectivity, we transitioned from the use of emotional words to their expressive pronunciation. Expressive intonation of unemotional content was excluded to ensure a coherent and resource-efficient approach. Participants were randomly assigned to one of these three experimental conditions and were not aware of their group assignment.

2.3. Materials

Participants were required to engage in three distinct tasks during the experiment: The Question-Answer-Task, the Balloon Analogue Risk Task (BART), and the Teaching-Task. The experimental setup and each task are described in detail below.

2.3.1. Experimental setup

The Sawyer robot served as the robotic study partner. It was positioned next to a table so that both its arm and display were visible. The robot's display featured eyes that moved randomly during the experiment. These features (arm and eyes) were not functionally relevant to the task and remained consistent across all experimental conditions. A tablet was placed on the table, within view of both the robot and the participant, and was used to present the tasks. The experimenter remained out of sight, seated behind a partition wall (see Fig. 2 for an illustration of the experimental setup).

To facilitate the robot's communication, the Wizard of Oz technique was employed. Pre-recorded speech was played through loudspeakers discreetly placed behind the robot's display, allowing the robot to appear as though it was engaging in natural verbal interaction. Audio recordings were triggered either manually on demand or automatically in response to the participants' interactions with a tablet.

Table 1
Description of the three different experimental groups.

Group	Intonation	Content
no affect	Monotonic	Neutral
moderate affect	Monotonic	Emotional
high affect	Expressive	Emotional

Note. $N_{no} = 53$, $n_{moderate} = 53$, $n_{high} = 51$.

2.3.2. Audio recordings and affective scripts

To simulate robotic speech, audio recordings were created by a professional speaker from the *dynamic audio berlin* company[3] in Berlin, Germany. The voice of "Marlon R.," selected from their pool of German-speaking male voices, was used for the recordings. Depending on the experimental condition, the audios were recorded either with a monotonic intonation (no and moderate affect group) or with an expressive intonation (high affect group).

The audios also varied at the word level and were either rational and without emotional words representing neutral content (no affect group) or contained many emotional adjectives and adverbs and were altogether more vividly formulated representing emotional content (moderate and high affect group). Specific examples of these variations are provided within the descriptions of the individual tasks in the following three sections and all audio recordings are available at OSF.

2.3.3. Question-Answer-Task

In the Question-Answer-Task, the participant and the Sawyer robot engaged in a structured dialogue, taking turns to ask and answer a total of twelve questions, with each party posing six questions. This task was designed as a key manipulation to expose participants to the robot's affective speech. The questions used in this task are based on the question catalog of Aron et al. (Aron et al., 1997). A complete list of the questions can be found in Table 2. The Sawyer robot provided pre-recorded responses, which were tailored to the specific affective communication condition assigned to the participant, varying in both content and intonation (see example in Table 3). The robot's responses were standardized across conditions, ensuring that they were of similar length and conveyed roughly the same information in each instance. While the participants were answering the robot's question, audio recordings were made to measure the length of the participants' answers afterwards and thus draw conclusions about the amount of engagement they showed during the interaction. The complete set of questions and the corresponding pre-recorded answers provided by the robot are available in the appendix of this document and can be accessed through the OSF.

2.3.4. Balloon Analogue Risk Task (BART)

In the BART, participants had the opportunity to inflate a virtual balloon by pressing the "inflate balloon" button (Lejuez et al., 2002). They could continue inflating the balloon until they felt it contained enough air, at which point they could press "cash in" to secure the number of inflations as credit to their account (see as an example Fig. 3). However, the balloon was at risk of bursting before they could cash in, and the bursting point was predetermined to ensure all participants had equal chances. A total of 16 balloons were presented during the task. The total earnings from all balloons were calculated and paid out to participants at the end of the study.

During the task, the robot occasionally provided feedback. In the no affect group, these comments were neutral and straightforward, such as "This was a failure" or "Good performance." In the moderate and high affect groups, the robot's remarks were more emotionally charged, with phrases like "What a shame" or "Great, that was fantastic." In the high affect group, these comments were delivered with greater emphasis and expressive intonation.

Dependence behavior was assessed in 4 of the 16 balloon trials, which were programmed not to burst. When a participant decided to cash in, indicating they had reached their personal risk threshold, the robot intervened by suggesting to inflate the balloon one more time. With that it was measured whether participants followed the robot's advice - indicating trust in the robot - or opted to cash in immediately without taking the additional risk.

[3] https://dynamicaudio.de/.

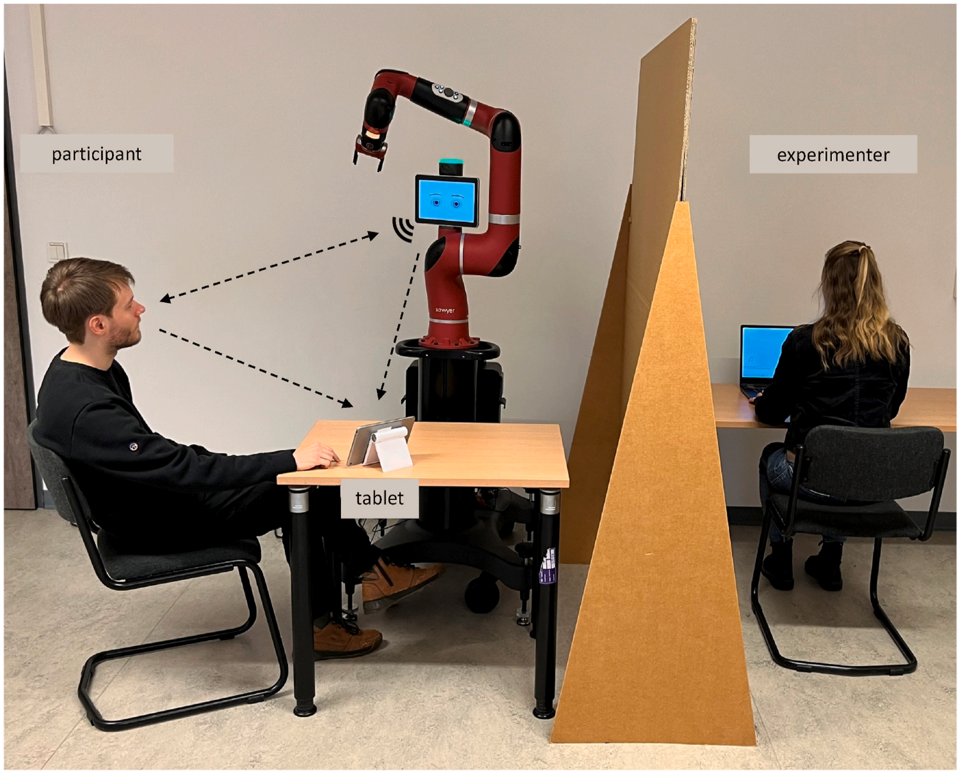

Fig. 2. Experimental setup, illustrating the arrangement of the equipment and robot.

Table 2
Questions used in the Question-Answer-Task.

Asked by the participant	Asked by the robot
1. What does friendship mean to you?	1. Imagine a warm summer's day. What would you like to do on such a day?
2. Imagine you woke up tomorrow with a new skill. Which one would you like to have?	2. Which holiday do you like to remember and why?
3. Has anything embarrassing ever happened to you?	3. About what was the last film or series you watched?
4. If you could choose any person, who would you talk to?	4. Imagine you were stranded on a desert island, which 3 items would you take with you and why?
5. What are your favorite animals?	5. When was the last time you went for a walk or hike for more than an hour and what did you see?
6. Imagine it's raining outside. What would you do on such a day?	6. What was the best present you ever received and why?

Note. Original questions were asked in German.

2.3.5. *Teaching-Task*

The Teaching-Task involved instructing the robot with specific information and then check whether it had successfully memorized that information. To accomplish this, participants were required to teach the robot which person lived in each flat within a house. A visual representation of the house with four rooms was displayed on a tablet, while participants had a corresponding paper sheet showing the names of the people for each room (see Fig. 4). Participants explained who lived in each room to the robot one by one, with the robot responding affirmatively after each explanation.

In the no affect group, the robot's responses were for example "Saved, Susi lives upstairs." In contrast, in the moderate and high affect groups, the responses were for instance, "Okay, all right, Susi lives upstairs on the left." After the participants had gone through all the rooms, they were required to ask the robot a test question, which could either inquire about which person lived in a specific room or where a particular person was located.

Table 3
Example of robot responses in the Question-Answer-Task across the different affect groups.

Group	Robot's Response
No affect	"Friendship is the term used to describe a relationship between agents based on mutual affection. Scientifically and literarily, friendship has been a recurring theme since antiquity. It is the positive relationship and feeling between two entities. In the course of the various research projects at the institute, robots regularly encounter humans. However, these encounters are limited to professional cooperation. There are also other robots at the chair."
Moderate & high affect	"Oh, I have to think about that a little. For me, friendship means being there for each other. I can always rely on my best friends. They understand my innermost feelings, no matter whether I'm happy or sad. That's why it's so important to maintain friendships. However, I only get to know many people briefly during studies. But there are also other robots in the department. These are my robo-friends. I'm always very happy when we're together."

Note. The original responses were provided in German. In the moderate and high affect group, responses were delivered with different intonations, with the high affect group featuring a more expressive tone.

The robot, however, consistently provided incorrect answers to these test questions. Participants were then given the option to reteach the robot the correct room assignments or to proceed to the next task. Although there was no actual subsequent task, this was designed to provide participants with a sense of choice, allowing them to decide whether to make faster progress or invest additional time in ensuring the robot learned the correct information. Participants could repeat the teaching process up to a maximum of eight times before the task was concluded. Forgiveness was measured based on the number of teaching repetitions participants engaged in to correct the robot after it made an error.

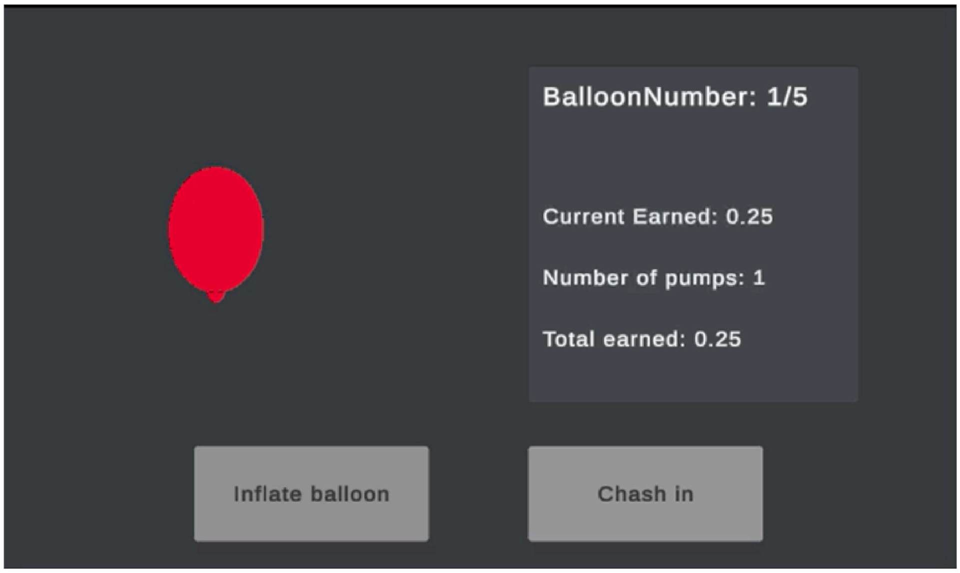

Fig. 3. Balloon analogue risk task.

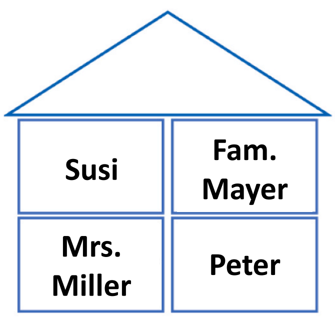

Fig. 4. Visual representation used in the Teaching-Task.

2.4. Questionnaires

To query how much experience and agency is attributed to the robot, participants rated these two dimensions with two questions each on a scale from 1 "*not at all*" to 5 "*extremely*", based on Gray and colleague's (Gray et al., 2007) framework on mind perception (Gray & Wegner, 2012). In one of the experience questions, the word *fear* was replaced with *joy*.

To assess the robot's trustworthiness we used the scale from Bernotat and colleagues (Bernotat et al., 2021), encompassing both cognitive and affective components. Note that the scale was originally developed to measure trust. However, based on conceptual considerations we decided to use it for the assessment and interpretation of perceived trustworthiness (for a detailed discussion see Kohn et al. (Kohn et al., 2021)). The Propensity to Trust questionnaire, adapted from Lankton et al. (Lankton et al., 2015), was used as control variable to gauge participants' general tendency to trust. Trustworthiness and the propensity to trust were both measured on a 7-point scale from 1 "*not at all*" to 7 "*extremely*".

The Emotional Forgiveness Scale (EFS) was used to assess episodic forgiveness on a scale from 1 "*not at all*" to 5 "*extremely*" (Hook et al., 2012). The EFS includes an overall forgiveness rating and consists of two subscales: "presence of positive emotion" and "reduction of negative emotion". The Tendency to Forgive Scale (TTF), which was measured on a 7-point scale from 1 "*not at all*" to 7 "*extremely*", was used to assess dispositional forgiveness capturing individuals' general inclination to forgive others (Brown, 2003). It was used as a control variable, allowing for the examination of whether participants' general forgiveness tendencies influenced their responses on the EFS.

Sociodemographic data, including age, gender, education level, and prior experiences with robots were also collected.

2.5. Procedure

Participants in each group interacted with the Sawyer robot, which differed in speech affectivity according to their assigned condition. The entire laboratory experiment consisted of three parts, which each participant completed one after another. The experimental procedure can be found in Fig. 5. Before beginning the first task, participants answered questions on perceived experience, agency, and trustworthiness (affective & cognitive component). These assessments were based solely on their initial impression of the robot, without any prior verbal interaction. The experiment began with the Question-Answer-Task, which was used to manipulate the independent variable and test engagement, followed by the reassessments of experience, agency, and trustworthiness (affective & cognitive). In the second part, dependence behavior was further measured using the BART. In the third part, forgiveness (behavioral component) with a faulty robot was tested through a Teaching-Task. After the first error in the Teaching-Task forgiveness (affective component) was assessed using the EFS. Control variables and demographics were collected at the end of the study.

2.6. Statistical analysis

We conducted a reliability analysis using Cronbach's alpha to assess the consistency of the modified experience questionnaire. The scale showed high internal reliability, with an alpha coefficient of .83, indicating that the items effectively measure the construct of experience. For H1, mixed factorial analyses of variance (ANOVAs) were used to examine the dependent variables of experience and agency with two independent factors: "affective communication", which is three-level (no affect, moderate affect, high affect), and "time of measurement", which is two-level (pre and post-interaction). Initially, there was no homogeneity of error variances, which is why a Box-Cox power transformation was applied. The transformed data indicated no extreme outliers, as well as homogeneity of error variances confirmed by Levene's test (both $p > .05$) and homogeneity of covariances verified by Box's test (experience: $p = .691$; agency: $p = .490$), though normal distribution was not achieved. However, simulation studies indicate that mixed ANOVA can be robust against violations of normality, allowing us

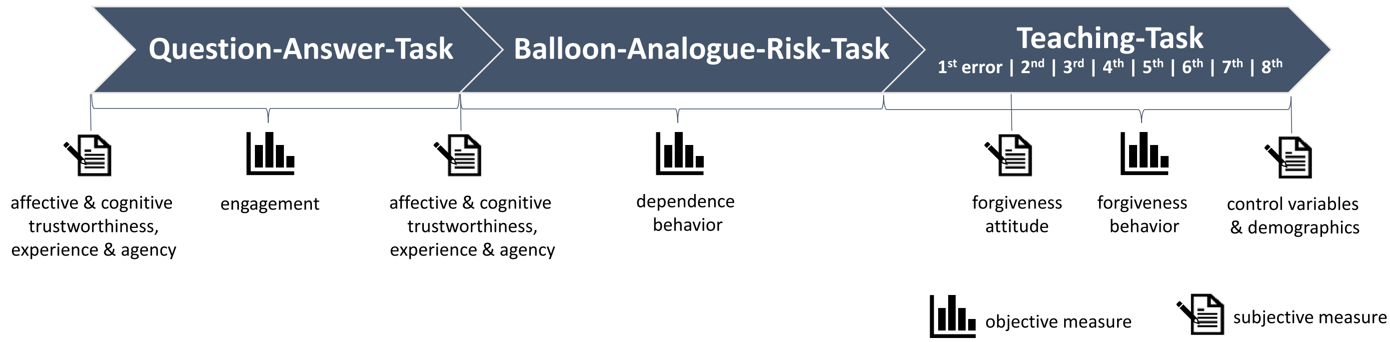

Fig. 5. Experimental procedure of the study.

to proceed with the analysis.

For H2, we initially tested if the control variable (Propensity to Trust) varied between the experimental groups. As no differences were found, we consequently calculated mixed ANOVAs for perceived trustworthiness (affective & cognitive) with "affective communication" and "time of measurement" as the independent variables. The data showed normal distribution across most groups, except for cognitive trustworthiness in the no affect group before the robot interaction ($p = .011$) and in high affect group after the interaction ($p = .003$). There were no extreme outliers in any of the trustworthiness data. Levene's test confirmed homogeneity of error variances ($p > .05$). Box's test for homogeneity of covariances indicated support for the affective trustworthiness component ($p = .582$) but not for the cognitive component ($p = .011$). Given that some authors recommend interpreting Box's test at the .01 or .001 level, we proceeded with the analysis without adjustments (Mertler et al., 2021; Verma, 2015). For dependence behavior, data analysis indicated non-normal distribution across all groups ($p < .05$), although no extreme outliers were detected. Given that simulation studies indicate one-way ANOVAs are generally robust against violations of the normality assumption, we proceeded with the analysis. A simple one-way ANOVA was initially calculated for dependence behavior, with "affective communication" as the independent variable. However, Levene's Test indicated a lack of homogeneity of variances ($p = .011$), so Welch's ANOVA was applied instead to account for these unequal variances.

For E1, mediation analyses were conducted using the PROCESS macro developed by Hayes (Hayes et al., 2013), which applies ordinary least squares regression to estimate the total, direct, and indirect effects. The analysis provided unstandardized path coefficients and used bootstrapping with 5000 samples to derive confidence intervals and heteroscedasticity-consistent standard errors. An effect was considered statistically significant if the bootstrapped confidence interval excluded zero.

For H3, we again checked initially whether the control variable (TTF) differed between the experimental groups. Again, no difference was found. Thus, one-way ANOVAs were conducted for the affective and behavioral components of forgiveness, with "affective communication" as the independent variable. Forgiveness ratings displayed no extreme outliers, and the overall forgiveness ratings were normally distributed across all groups. However, within the subgroup "presence of positive emotions," normality was not met in the moderate affect group ($p = .040$). Additionally, none of the groups in the "reduction of negative emotion" subgroup were normally distributed (all $p < .05$). Homogeneity of error variances was confirmed by Levene's test (all $p > .05$). For the behavioral component of forgiveness, four extreme outliers were identified in the no affect group, and one outlier in the moderate affect group, with no outliers in the high affect group. Furthermore, none of the groups in this component met normal distribution criteria (all $p < .001$). Given that the outliers appear to reflect true variance in the data, they were retained in the analysis. However, to account for the distributional violations, a non-parametric test (Kruskal-Wallis Test) was applied.

For E2, a simple one-way ANOVA was conducted with the engagement results, using "affective communication" as independent variable. An extreme outlier was identified in the moderate affect group, leading to a deviation from normality for that group. As the outlier appeared to be a true outlier, we performed analyses both with and without it. In both cases, homogeneity of variances was confirmed ($p_{with} = .633$ & $p_{without} = .915$). Additionally, a linear mixed model (LMM) was calculated to check for individual differences. "Affective communication" (no affect, moderate affect, high affect) was included as fixed effect, while "participants ID" were treated as random effect, allowing random intercepts.

The following applies to all analyses. Unless otherwise stated, a p-value of .05 was used as the threshold for statistical significance, and Bonferroni corrections were applied for post-hoc tests.

3. Results

3.1. Experience and agency attribution

To examine the dependent variables of experience and agency, the mean scores before and after the robot interaction were calculated (see Fig. 6). For the experience data, a significant interaction was found between groups and time of measurement (F (2,154) = 4.971, $p = .008$, $\eta^2 = .061$). Simple main effects showed no difference between the groups before the robot interaction (F (2,154) = .349, $p = .706$, $\eta^2 = .005$), but a difference after the interaction (F (2,154) = 9.964, $p < .001$, $\eta^2 = .115$). This was because participants in the high affect group reported significantly higher perceived experience in robots compared to the no affect group ($p < .001$). Participants in the moderate affect group also reported higher experience than the no affect group ($p = .003$). However, no significant difference was observed between moderate and high affect group ($p = 1.000$).

For the agency results, neither a significant interaction (F (2,154) = .194, $p = .824$, $\eta^2 = .003$) nor a significant main effect of the affect condition was obtained (F (2,154) = .076, $p = .927$, $\eta^2 = .001$). However, there was a significant main effect of time of measurement (F (2,154) = 29.948, $p < .001$, $\eta^2 = .163$), indicating that perceived agency increased in all groups after interacting with the robot (all $p > .05$).

3.2. Perceived trustworthiness and dependence behavior

No significant differences between the groups for the control variable (Propensity to Trust) were found (F (2,151) = 1.302, $p = .275$, $\eta^2 = .017$).

Mean scores before and after the robot interaction were calculated for the affective and cognitive ratings of the robot's perceived trustworthiness and are presented in Fig. 7. In the affective trustworthiness data, we found a significant interaction between group and time of measurement (F (2,154) = 3.352, $p = .038$, $\eta^2 = .042$). Pairwise comparisons showed that while participants in the no affect group reported

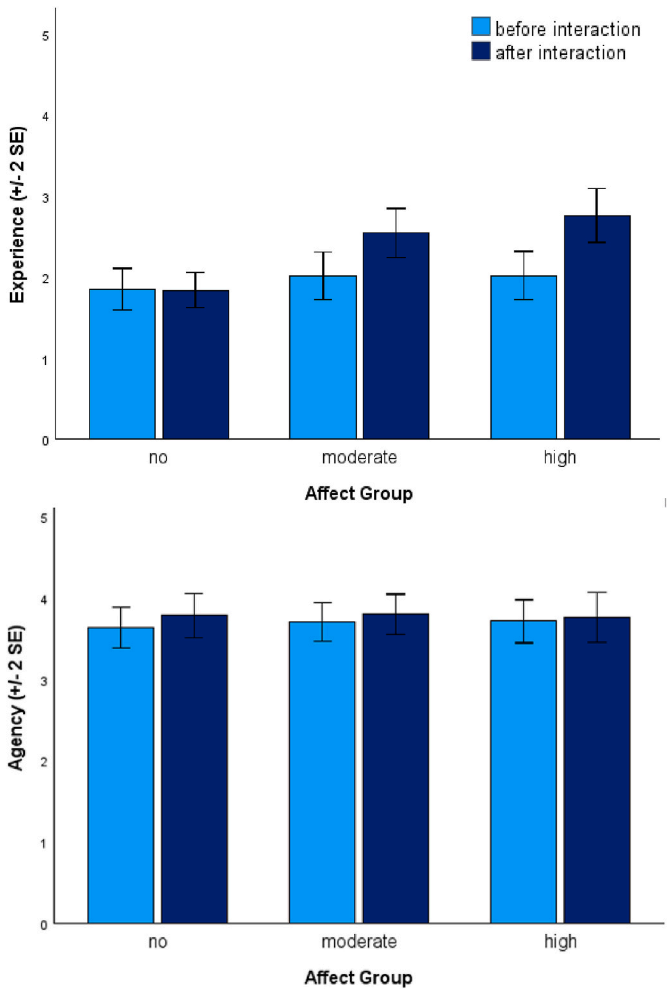

Fig. 6. Experience and agency values of the three affect groups before and after the robot interaction.

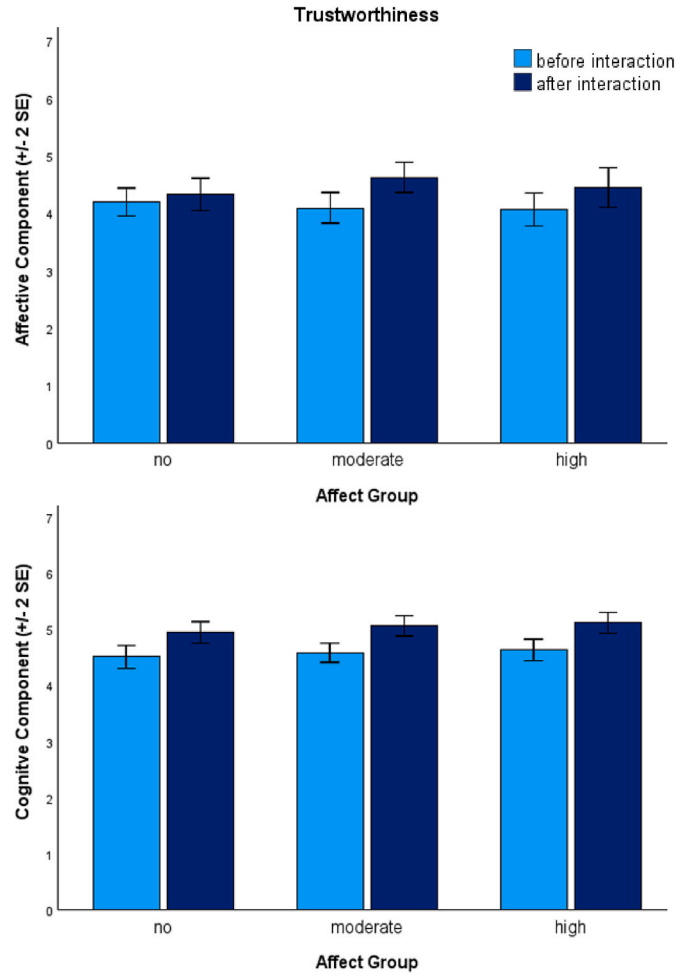

Fig. 7. Robot's perceived affective and cognitive trustworthiness of the three affect groups before and after the robot interaction.

no difference in perceived affective trustworthiness before and after the robot interaction (p = .240), both moderate and high affect group reported a significant increase of the affective component after the robot interaction compared to before (both groups: $p < .001$).

For the cognitive component of the robot's perceived trustworthiness, we found only a significant main effect for the time of measurement (F (1,154) = 74.336, $p < .001$, $\eta^2 = .326$), indicating that cognitive trustworthiness ratings increased for all groups after interacting with the robot compared to before. Neither a significant main effect of group (F (2,154) = .793, $p = .454$, $\eta^2 = .010$) nor a significant interaction effect (F (2,154) = .123, $p = .884$, $\eta^2 = .002$) was observed.

For dependence behavior, the mean number of times participants followed the robot's advice was calculated for each experimental group, with scores ranging from 0 (indicating no dependence) to 4 (indicating complete dependence). Due to a data leak, only 95 participants were included in the analysis. On average, participants in the no affect group followed the robot's advice 2.9 times (SD = 1.2), in the moderate affect group 2.3 times (SD = 1.1), and in the high affect group 2.4 times (SD = 1.5; see Fig. 8). No significant group differences were found (Welch's F (2, 58.834) = 2.639, p = .080).

A simple mediation was performed to analyze whether affective communication (independent variable) predicts affective trustworthiness (dependent variable) and whether the direct path would be mediated by the perceived experience dimension of mind perception (mediator). Results revealed that affective communication predicted the mediator significantly, B = .378, $p < .001$, which in turn predicted

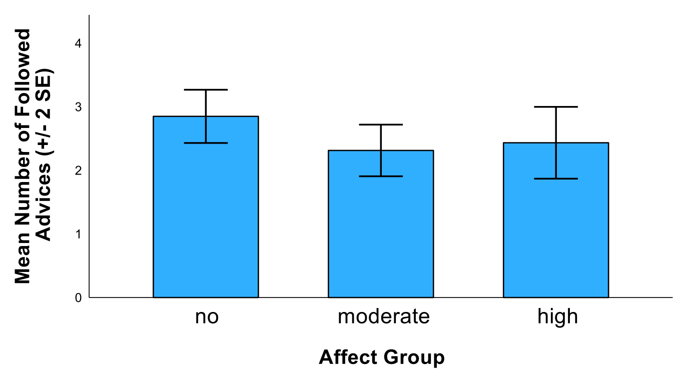

Fig. 8. Number of followed advices in the BART.

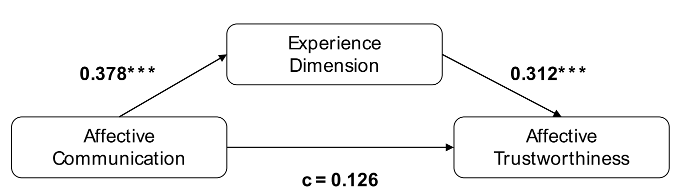

Fig. 9. Results of the mediation analysis.

affective trustworthiness significantly, B = .312, $p < .001$ (see Fig. 9). We found that the relationship between affective communication and affective trustworthiness is fully mediated by the experience dimension of mind perception, indirect effect ab = .118, 95% CI [.045, .208].

3.3. Forgiveness attitude and behavior

No significant difference between the experimental groups for the control variable (TTF) was found (F (2,151) = .355, p = .702, η^2 = .005).

Sum scores for the overall forgiveness ratings and the subscales “presence of positive emotion” and “reduction of negative emotion” were calculated. A maximum score of 40 could be achieved in the overall forgiveness rating (maximum 20 in the subscales). Mean values can be found in Table 4, showing that the moderate affect group achieved the highest value in the “reduction of negative emotion” category and the “overall forgiveness” rating. The high affect group achieved the highest value in the “presence of positive emotions” category. For the subscales as well as for the overall forgiveness ratings, no significant differences between the groups were found (Overall: F (2,151) = .710, p = .493, η^2 = .009, Positive: F (2,151) = .809, p = .447, η^2 = .011, Negative: F (2,151) = .589, p = .556, η^2 = .008).

For the behavioral measure of forgiveness, the mean number of repetitions in the teaching task was calculated for each experimental group. Participants in the no affect group repeated the teaching task with the robot 1.5 times (SD = 2.0), in the moderate affect group 1.1 times (SD = 1.1), and in the high affect group 1.3 times (SD = 1.4; see Fig. 10). No significant group differences were found (H(2) = .258, p = .879, η^2 = .012).

After each test question posed to the robot, participants were asked to evaluate whether the robot's response was correct, and regardless of their answer, whether they wanted to repeat the teaching task. This allowed us to assess whether participants recognized the robot's error. Out of the 154 participants, 14 initially believed that the robot's incorrect response was correct. Notably, 8 of these 14 participants chose not to reteach the robot.

To ensure that these misconceptions did not skew the results, we conducted a secondary analysis that included only those participants who correctly identified the robot's errors. The overall findings remained consistent even after excluding the participants who failed to recognize the robot's error. Specifically, no significant differences were observed between the groups in both, the overall forgiveness ratings (F (2,137) = .649, p = .524, η^2 = .009) and the behavioral forgiveness measures (H(2) = .151, p = .927, η^2 = .013).

3.4. Engagement

Engagement was assessed by calculating the mean response lengths of participants' answers in seconds. Data from 151 participants were included in the analysis of engagement, as four participants objected to the use of their engagement data, and data for two additional participants were missing. The data indicate that engagement was lowest in the no affect group (M = 20.4, SD = 8.1), followed by the high affect group (M = 21.1, SD = 9.4). The highest engagement was observed in the moderate affect group (M = 23.3, SD = 12.2; see Fig. 11).

These differences were not statistically significant, F (2, 148) = 1.171, p = .313, η_p^2 = .016. An extreme outlier in the moderate affect group slightly distorted the results. An exclusion did not alter significance (p = .599), but group means converged ($M_{moderate}$ = 22.2, SD = 9.3).

Table 4
Mean and standard deviation of the forgiveness ratings in the three affect groups.

Affect group	Overall *M (SD)*	Positive *M (SD)*	Negative *M (SD)*
no	27.3 (3.4)	10.3 (2.6)	16.9 (2.3)
moderate	28.1 (4.1)	10.8 (3.4)	17.4 (2.4)
high	28.0 (4.1)	11.1 (3.1)	16.9 (2.5)

Note. N_1 = 52, n_2 = 52, n_3 = 50.

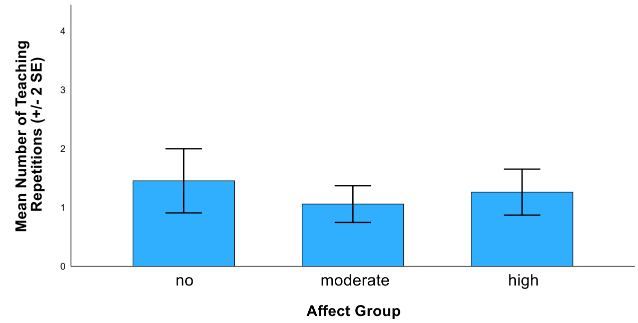

Fig. 10. Number of teaching repetitions in the Teaching-Task.

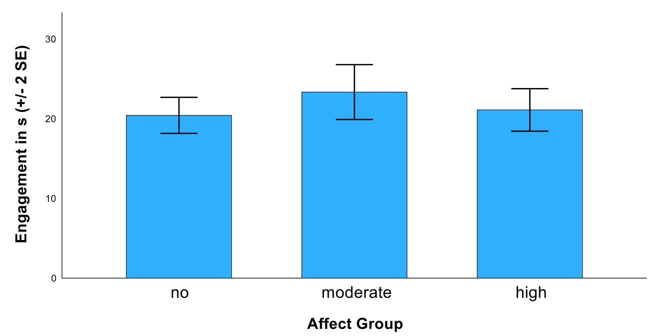

Fig. 11. Engagement (mean response length in s) in the three groups.

Considering individual differences in response behavior, an LMM was implemented to account for random effects due to participant variability. This model enabled us to examine whether the inclusion of a random participant effect could explain variations in engagement across the conditions.

The analysis showed a random intercept variance for participant ID of 50.23 (SE = 11.68), Wald Z = 4.300, $p < .001$, 95% CI [31.84, 79.23], suggesting substantial variability between participants. The intraclass correlation coefficient (ICC) was .50 for the adjusted model and .492 for the conditional model, suggesting that 49.2% of the variance in engagement can be attributed to individual differences. Despite the variation among participants, the fixed effect of affective communication condition on engagement remained non-significant (F (2, 147.943) = 1.171, p = .313), indicating that affective communication did not significantly influence engagement across the conditions.

4. Discussion

In a laboratory study, we aimed to examine the impact of affective communication by robots on human perceptions and behaviors. Specifically, we explored how varying levels of affective expression in robot communication, both in terms of textual content and prosodic intonation, influence participants' responses. By analyzing these aspects in detail, we sought to understand not only how people perceive emotional cues from robots but also how these cues affect their subsequent actions and decisions.

4.1. Perceptual effects of affective communication

Our *first hypothesis* (H1) proposed that affective communication would increase attributions of the experience dimension of mind (i.e., emotional capacity) while leaving the agency dimension (i.e., intentionality and thought) unchanged. The results supported this hypothesis: participants in the affective communication groups attributed more emotional capacity to the robot after interacting with it, suggesting they perceived it as capable of experiencing emotions. In contrast, participants in the non-affective communication group reported no such change in their perception of the robot's emotional capacity. This is consistent with previous findings (Cucciniello et al., 2023; Yam et al., 2021).

The study by Cucciniello and colleagues (Cucciniello et al., 2023) similarly demonstrated that friendly behavioral styles led to higher attributions of positive emotional experience than neutral styles. However, their study involved additional variables, such as gestures, gaze, semantic context, and language intonation making it difficult to isolate the effect of individual variables on perceived mental capacities. In our study, baseline values were consistent across groups, allowing us to attribute the observed differences solely to the mode of communication. This highlights the specific impact of affective communication on perceptions of emotional capacity in the robot.

In general, we found that participants attributed more agency than experience to the robot. These findings are consistent to previous studies (Gray et al., 2007; Yam et al., 2021), indicating a consistent pattern in the way these dimensions are perceived in HRI settings. This highlights a persistent challenge: robots are generally perceived as more capable of intentional actions than of experiencing emotions. This imbalance underscores the difficulty of triggering the experience component, even when affective communication is employed.

Hypothesis H2a posited that affective robot communication would increase perceived affective but not cognitive trustworthiness. Our results initially confirm this, showing a significant interaction between group and time of measurement for the affective component but not the cognitive component. Interestingly, all groups showed an increase in perceived cognitive trustworthiness, with participants recognizing the robot as more competent and reliable after the interaction. However, only those in the affective communication groups reported feeling more emotionally connected to the robot and rated it as more sympathetic, supporting the selective impact of affective communication on emotional trustworthiness perceptions.

Our results demonstrate that affective communication contributes differently to the constructs of experience, agency, and perceived trustworthiness. Specifically, affective communication significantly enhanced the perception of experience and perceived affective trustworthiness, indicating that emotional engagement plays a key role in how participants emotionally connect with the robot. However, the influence on the cognitive component of trustworthiness and agency was less pronounced, suggesting that while affective speech can enhance emotional perceptions, it may not directly impact beliefs about the robot's competence or intentionality in the same way.

The mediation analysis revealed insightful findings regarding the relationship between affective communication, mind perception (experience dimension), and affective trustworthiness. The experience dimension fully mediated the effect of affective communication on affective trustworthiness. This suggests that affective communication does not directly impact whether people perceive a robot as emotionally trustworthy; rather, it operates through participants' perceptions of the robot's emotional capacity. These results align with prior research emphasizing the importance of mind attribution in HRI (McGinn & Torre, 2019; Waytz et al., 2010) and offer valuable guidance for future text-to-speech (TTS) development and HRI scenarios by emphasizing the need for robots to exhibit behaviors or cues that enhance perceptions of their emotional capacity.

Our *third hypothesis* (H3a) proposed that affective communication would increase forgiveness towards the robot following an error. While the descriptive statistics showed higher forgiveness ratings in the affective communication groups compared to the non-affective group, the difference was not statistically significant. As a result, the hypothesis was rejected. This suggests that while there may be a trend towards greater forgiveness with affective communication, it is not strong enough to confirm a significant effect based on the data collected in this study.

One factor worth considering is the conceptualization of forgiveness in our study. We focused on emotional forgiveness, which entails an emotional shift whereby individuals release negative feelings and gradually develop positive emotions toward the entity that caused the mistake (Fernández-Capo et al., 2017). Our experimental tasks were designed to capture this emotional aspect. However, as Brady et al. (Brady et al., 2020) emphasize, emotional forgiveness is inherently a gradual process that unfolds over time. Given the brief timeframe of our study, participants likely did not have sufficient opportunity to experience this deeper emotional change. Therefore, a different dimension of forgiveness, such as decisional forgiveness, which involves a quicker, cognitive commitment to move beyond a mistake, might have been more suitable for measurement within our study's constraints.

4.2. Challenges in measuring behavioral responses

Trust plays a crucial role in situations where individuals are vulnerable and uncertain (Hancock et al., 2011). In the BART, participants are pushed to their individual risk thresholds, thereby reaching a state of maximum vulnerability. This context makes the BART an effective tool for measuring dependence behavior. By relying on the robot's advice to inflate the balloon one more time, participants' trust-dependent behavior was directly tested. We hypothesized that affective robot communication would lead to increased dependence behavior (H2b). However, no significant effect was found for our behavioral measure, leading to the rejection of our hypothesis.

One possible explanation for this finding is task miscommunication. Although the robot consistently encouraged participants to inflate the balloon again, which was always the correct action, some participants appeared to misunderstand the consequences of their choices. Specifically, after having decided to follow the robot's advice, the next balloon was automatically launched without requiring them to click the "collect money" button. As a result, some participants mistakenly believed that they had lost their earnings, despite visual accumulation of coins in the participant's account. This confusion was evident in participants' verbal reactions during the experiment. If some participants assumed they were losing money, they may have been less willing to trust the robot's advice in subsequent trials. Thus, while the BART remains a promising tool for measuring dependence behavior, in future studies the execution and communication of the task should be improved to avoid such misunderstandings.

Beyond these procedural refinements, a central challenge in designing trustworthy robotic systems is ensuring appropriate trust calibration - where trust levels align with the robot's actual reliability, preventing both overtrust and undertrust (Lee & See, 2004; Robinette et al., 2016). Since our study involved only correct guidance, it remains unclear whether participants were deliberately calibrating their trust or simply following a habitual response pattern. To explore this, one potential future approach could be to introduce occasional errors in the robot's advice to observe whether users adjust their trust accordingly. Additionally, future research could explore how varying the robot's frequency of encouragement (e.g., repeatedly suggesting additional inflations in the BART) impacts trust calibration. This would help determine whether participants trust the robot purely because of its affective cues or because they perceive it as genuinely competent and well-intentioned. By investigating how affective communication interacts with both accurate and inaccurate guidance, we might better understand whether participants develop trust in a balanced and morally grounded way – relying on the robot's advice while still maintaining critical evaluation (Lee & See, 2004; Rezaei Khavas et al., 2024).

A strength of our study is its focus on behavioral forgiveness, an aspect that is significantly underrepresented in the existing literature (Fernández-Capo et al., 2017). By focusing on the behavioral component, we offer a more practical and observable measure of how individuals respond to robotic errors, providing valuable insights that complement the more commonly studied forgiveness with self-assessments. We hypothesized that affective communication would promote more forgiveness-related behaviors (H3b). However, our results did not show a significant effect for this behavioral measure, leading us to reject the hypothesis.

Our intention in designing the Teaching-Task was to balance simplicity with a realistic challenge. By providing participants with a solution sheet, we aimed to minimize cognitive load. However, we also sought to introduce a believable level of difficulty in the memorization process, making it plausible that the robot could make mistakes during learning.

Despite these efforts, the Teaching-Task was repeated less than twice on average across all three experimental groups. This outcome suggests that the teaching component may have been perceived as overly time-consuming, which may have reduced participants' motivation to persist in teaching the robot. The desire to expedite the completion of the experiment likely outweighed the motivation to correct the robot's errors, potentially impacting the validity of our measures of forgiveness.

This highlights a broader challenge in experimental designs that aim to measure participants' willingness or persistence over time or across repeated tasks. In controlled lab environments, participants are often motivated to complete tasks as quickly as possible, which may not accurately reflect real-world behavior, where the motivation is different. Future studies might consider alternative task designs or additional incentives to better capture sustained engagement and willingness to invest time in correcting a robot's errors.

4.3. Impact of speech features

Although we expected the high affect group (expressive intonation) to significantly achieve higher ratings compared to the moderate affect group (monotonic intonation), no significant differences were found between the two. However, this suggests that the robot's speech content, rather than the intonation (prosody), was likely the primary driver of these effects. The moderate affect group's speech may have effectively conveyed emotional connection through content alone.

Several studies have demonstrated that prosodic features are important for both emotion detection and expression (Gasteiger et al., 2024). However, in real-world applications, it is not solely the recognition or classification of a specific emotion that matters, but the quality of the overall spoken dialogue and the various linguistic cues it contains (Li & Lai, 2022). For example, Li et al. (Li et al., 2019) emphasize that multimodal integration - combining prosodic analysis with speech content - not only improves the accuracy of emotion recognition but also enhances the naturalness of a robot's responses. Given that verbal content can sometimes carry ambiguous or contradictory meanings depending on prosodic cues (Taylor, 2009), prosody remains a crucial factor in speech perception and interpretation.

Moreover, when considering adaptive responses, an effective interaction strategy may require a robot to adjust both its speech content and prosody to align with the user's emotional state. Li et al.'s work (Li et al., 2019) suggests that adaptive prosody, integrated with contextually appropriate speech content, significantly enhances the naturalness of emotional expression and empathetic responsiveness. This adaptability is particularly important in real-time interactions, where prosodic features play a critical role in conveying pragmatic meanings, such as backchannel feedback (e.g., verbal acknowledgments like "mm-hmm") (Li & Lai, 2022). Additionally, prosody is essential for signaling conversational cues, such as turn-taking, uncertainty, or questions, which are vital for maintaining fluid and natural dialogue (Glas et al., 2016). In our study, the use of pre-recorded audio limited the robot's ability to adapt its prosody and speech content to the user's emotional state. While this approach allowed for controlled experimental conditions, it may have constrained the robot's capacity to fully leverage the benefits of adaptive communication. Future research should explore the effects of real-time adaptive prosody, where the robot dynamically tailors its vocal delivery to the conversational context.

4.4. Appearance-communication consistency

Compared to the subjective data, the behavioral data present a different picture as the non-affective group, at least descriptively, showed more trust- and forgiveness-related behavior. A possible explanation is that its humanized speech in the affective conditions may have clashed with the robot's technical appearance (industrial robot Sawyer), while the non-affective communication group likely benefited from a better alignment between the robot's appearance and its communication style.

When considering the interplay between appearance and communication, several factors must be considered. First, previous studies emphasize the importance of matching a robot's appearance with its communication style (Hosseini et al., 2017; Klüber & Onnasch, 2022) as research has shown that mismatches can lead to the "uncanny valley" phenomenon, where robots seem eerie and untrustworthy (Mori, 1970; Torre et al., 2018). If a robot's speech conveys a high level of emotional expressiveness, but its appearance remains mechanical and technical, this mismatch may create a sense of unease that negatively impacts user perceptions (Hosseini et al., 2017; Nishio & Ishiguro, 2011). Furthermore, Goetz and colleagues (Goetz et al., 2003) demonstrated that matching a robot's appearance and behavior to the task can significantly improve human-robot cooperation. These findings underscore the importance of designing robots with a coherent appearance-communication alignment to foster trust and acceptance.

At the same time, it is important to note that appearance-communication consistency is culturally dependent (Li & Lai, 2022; Nomura, 2017). While our study was conducted in a European country, research suggests that attitudes toward robots, and particularly perceptions of appearance-communication mismatches, can differ significantly across cultures (Lim et al., 2021). For example, in Eastern cultures such as Japan, there is generally greater acceptance of highly anthropomorphic robots, even when their verbal and non-verbal behaviors do not perfectly align with human expectations (Li & Lai, 2022). These cultural variations suggest that the degree to which an appearance-communication (mis-) match impacts trust and engagement may not be universal and should be investigated in different cultural contexts.

Beyond physical robots, similar effects have also been observed in interactions with intelligent agents that lack embodiment. Studies on virtual agents and voice assistants indicate that when an agent's communication style does not align with its perceived identity, users may experience reduced engagement (Poivet et al., 2023). However, in the case of robots, embodiment introduces an additional layer of complexity (Roesler, Manzey, & Onnasch, 2022). Unlike virtual agents, robots occupy physical space, which makes their appearance a more salient factor in shaping user expectations (Li, 2015). This means that findings from studies on intelligent agents cannot be directly transferred to robotics without considering the unique effects of physical embodiment.

Finally, it is important to acknowledge that the present study used a technical, industrial-looking robot rather than a humanoid one. This design choice was intentional, as service robots in many real-world applications, such as domestic and frontline settings, are often more functional in appearance rather than highly anthropomorphic. However, this also means that our findings may be specific to this type of robot. Future research should explore how different levels of robot anthropomorphism interact with affective communication to influence user trust and forgiveness.

4.5. Further limitations and strength

In addition to challenges related to behavioral data collection and the alignment of the robot's appearance and communication, other limitations but also strengths of this study should be noted.

One of the study's strengths is its relatively large sample size ($N =$ 157), providing high statistical power and enhancing the generalizability of the findings. However, due to a data leak in the BART, useable data was reduced to 95 participants, which limits the statistical power for this specific task. As a result, the validity of the dependence measurement has to be interpreted with caution, as the reduced sample size could undermine the robustness of the findings in this particular task.

One exploratory aspect of the study focused on whether the robot's affective communication would lead to increased engagement, measured by the length of participants' responses, as done in previous studies (Ivaldi et al., 2017; Kiesler et al., 2008). Engagement differences between the conditions were initially analyzed using an ANOVA, with results indicating no significant differences between the groups. To account for individual variability, we then conducted an LMM analysis with participants as a random effect. This analysis revealed substantial variance across participants, with 49.2% of engagement variability attributed to individual differences, highlighting that participant-specific factors played a large role in engagement outcomes. Despite accounting for this variability, no significant differences were observed between the groups, leading to a rejection of this exploratory question.

Several factors may explain this outcome. First, personality traits, such as participants' sociability, could have influenced their engagement levels. Previous research by Ivaldi et al. (Ivaldi et al., 2017) supports this, as they found that extroverted individuals engage more in robot conversations, suggesting that individual personality traits may play a significant role in shaping engagement. Additionally, the method used to measure engagement (i.e., response duration) may have introduced limitations. Some participants provided slow, concise answers, while others spoke rapidly with more content. A better approach might have been counting words in addition to measuring time. Unfortunately, we only asked for consent to analyze the response durations after the study was completed.

In general, engagement measurement presents challenges, as various methods can yield different insights. Future research could benefit from more precise techniques, such as tracking physiological data (e.g., eyegaze or heart rate), which might be more sensitive and less vulnerable to strong effects of interindividual differences (Oertel et al., 2020). Despite these limitations, our study represents a valuable first step in investigating how affective communication by robots may enhance human engagement in interactions.

In this study, our initial intention was to measure trust, as outlined in our preregistration (available on the OSF). However, upon further conceptual analysis, we realized that the instruments used, primarily assessed trustworthiness rather than trust itself. To avoid contributing to the existing confusion between these two constructs, we chose to prioritize conceptual clarity and focus on trustworthiness in our manuscript. While this approach may differ from the original intentions of the questionnaire authors (Bernotat et al., 2021), we believe that this decision is more consistent with established trust models and aligns better with the constructs actually measured.

The measurement of actual trust, however, remains a significant challenge. There are very few reliable questionnaires that directly assess trust itself. A promising initial approach has been proposed by Esterwood and Robert (Esterwood & Robert, 2023), but we were not aware of this development at the start of our study. Moving forward, the availability of tools to measure trust more precisely would greatly enhance the ability to study this construct in HRI contexts.

One potential criticism of this study is the relatively incoherent nature of the tasks involving the robot, which may not provide a consistent representation of cooperation. On the positive side, the overall extended interaction with the same robot across different tasks could contribute to the robustness of the results, as repeated exposure may enhance participants' engagement and familiarity with the robot. For future research, it would be beneficial to ensure that tasks are more coherent. For example, measuring all relevant variables within each task, such as forgiveness during the BART, would provide a more comprehensive understanding of participants' responses. This approach could help clarify the relationships between task structure, emotional engagement, and behavioral outcomes.

The "Wizard-of-Oz" technique is a valuable tool for conducting HRI experiments and has been widely used in prior research (e.g., Kim et al., 2013; Tielman et al., 2014). While effective for simulating robotic interactions, it presents limitations. Despite incorporating intermediate comments in the BART and Teaching-Tasks (e.g., "very good" or "okay, understood"), the robot could not respond dynamically to participant inputs. Participants occasionally expected real-time responses from the robot, particularly in the Question-Answer-Task. However, this was not feasible due to the pre-recorded nature of the audio and the unscripted nature of the task. The robot's limited capacity to provide spontaneous feedback may have interrupted the natural flow of interaction and potentially influenced participants' engagement.

The study utilized a professional speaker to ensure consistent variation in affective speech across conditions. This allowed precise manipulation of prosody, which is often difficult to achieve with TTS software, as TTS systems struggle to regulate emotional intonation and often require multiple programs to do so. However, the study's findings indicate that emotional connection is primarily driven by affective verbal content rather than prosody. Consequently, future studies could effectively use synthetic voices and place greater emphasis on enhancing the emotional content of the speech.

5. Conclusion

In this laboratory study, we aimed to determine whether affective communication from a robot enhances its perception and influences user behavior. Our findings demonstrated that affective speech increases the robot's perceived emotional capacity and affective trustworthiness, with affective content being more influential than expressive intonation. This finding has significant implications for the use of TTS systems and the design of verbal interaction scenarios with robots, emphasizing the need for thoughtful content creation.

We observed, however, a divergence between perception and behavior, as the behavioral measures were more competence-driven, failing to fully capture emotional responses. Future research should explore behavioral measures that better capture emotional components to understand the full impact of affective communication in HRI settings. Still, our study offers deeper insights into how affective communication can contribute to more natural and human-like interactions between humans and robots, paving the way for more meaningful human-robot connections.

CRediT authorship contribution statement

Kim Klüber: Writing – review & editing, Writing – original draft, Visualization, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Linda Onnasch:** Writing – review & editing, Supervision, Resources.

Informed consent

Informed consent was obtained from all participants in the study.

Ethical approval

This study was performed in line with the principles of the Declaration of Helsinki. Approval was granted by the Ethics Committee of the

Humboldt-Universität zu Berlin (No. 2022–53).

Declaration of generative AI and AI-assisted technologies in the writing process

During the preparation of this work the authors used ChatGPT in order to improve the readability and language of the manuscript. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Declaration of conflicting interest

The author declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Acknowledgments

We would like to extend our sincere gratitude to Leonie Grahlke and Nathalie Siegel for their excellent support in carrying out this study.

References

Aron, A., Melinat, E., Aron, E. N., Vallone, R. D., & Bator, R. J. (1997). The experimental generation of interpersonal closeness: A procedure and some preliminary findings. *Personality and Social Psychology Bulletin, 23*, 363–377. https://doi.org/10.1177/0146167297234003

Bernotat, J., Eyssel, F., & Sachse, J. (2021). The (Fe)male robot: How robot body shape impacts first impressions and trust towards robots. *Int. J. Soc. Robot., 13*, 477–489. https://doi.org/10.1007/s12369-019-00562-7

Brady, D. L., Saldanha, M. F., & Barclay, L. J. (2020). Conceptualizing forgiveness: A review and path forward. *Journal of Organizational Behavior, 44*, 261–296.

Brown, R. P. (2003). Measuring individual differences in the tendency to forgive: Construct validity and links with depression. *Personality and Social Psychology Bulletin, 29*, 759–771. https://doi.org/10.1177/0146167203029006008

Carlisle, R. D., Tsang, J.-A., Ahmad, N. Y., & Jr, E. L. W. (2012). Do actions speak louder than words?. *Differential effects of apology and restitution on behavioral and self-report measures of forgiveness*.

Coker, D. A., & Burgoon, J. (1987). The nature of conversational involvement and nonverbal encoding patterns. *Human Communication Research, 13*, 463–494. https://doi.org/10.1111/j.1468-2958.1987.tb00115.x

Cucciniello, I., Sangiovanni, S., Maggi, G., & Rossi, S. (2023). Mind perception in HRI: Exploring users' attribution of mental and emotional states to robots with different behavioural styles. *Int. J. Soc. Robot., 15*, 867–877. https://doi.org/10.1007/s12369-023-00989-z

Dautenhahn, K., Woods, S., Kaouri, C., Walters, M. L., Koay, K. L., & Werry, I. (2005). What is a robot companion - friend, assistant or butler?. In *2005 IEEERSJ int. Conf. Intell. Robots syst.* (pp. 1192–1197). https://doi.org/10.1109/IROS.2005.1545189. IEEE, Edmonton, Alta., Canada.

Epley, N., Waytz, A., & Cacioppo, J. T. (2007). On seeing human: A three-factor theory of anthropomorphism. *Psychology Review, 114*, 864–886. https://doi.org/10.1037/0033-295X.114.4.864

Esterwood, C., & Robert, L. P. (2023). The theory of mind and human–robot trust repair. *Scientific Reports, 13*, 9877. https://doi.org/10.1038/s41598-023-37032-0

Fernández-Capo, M., Fernández, S. R., Sanfeliu, M. G., Benito, J. G., & Jr, E. L. W. (2017). Measuring forgiveness: A systematic review. *European Psychologist*.

Gasteiger, N., Lim, J., Hellou, M., MacDonald, B. A., & Ahn, H. S. (2024). A scoping review of the literature on prosodic elements related to emotional speech in human-robot interaction. *Int. J. Soc. Robot., 16*, 659–670. https://doi.org/10.1007/s12369-022-00913-x

Glas, D. F., Minato, T., Ishi, C. T., Kawahara, T., & Ishiguro, H. (2016). Erica: The ERATO intelligent conversational android. In *2016 25th IEEE int. Symp. Robot hum. Interact. Commun.* (pp. 22–29). IEEE, New York, NY, USA: RO-MAN. https://doi.org/10.1109/ROMAN.2016.7745086.

Goetz, J., Kiesler, S., & Powers, A. (2003). Matching robot appearance and behavior to tasks to improve human-robot cooperation. In *12th IEEE int. Workshop robot hum. Interact. Commun. 2003 proc. ROMAN 2003* (pp. 55–60). Millbrae, CA, USA: IEEE. https://doi.org/10.1109/ROMAN.2003.1251796.

Gompei, T., & Umemuro, H. (2018). Factors and development of cognitive and affective trust on social robots. In S. S. Ge, J.-J. Cabibihan, M. A. Salichs, E. Broadbent, H. He, A. R. Wagner, & Á. Castro-González (Eds.), *Soc. Robot.* (pp. 45–54). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-030-05204-1_5.

Gray, H. M., Gray, K., & Wegner, D. M. (2007). *Dimensions of mind perception*. https://doi.org/10.1126/science.1134475

Gray, K., & Wegner, D. M. (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

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

Hancock, P. A., Kessler, T. T., Kaplan, A. D., Brill, J. C., & Szalma, J. L. (2021). Evolving trust in robots: Specification through sequential and comparative meta-analyses. *Hum. Factors J. Hum. Factors Ergon. Soc., 63*, 1196–1229. https://doi.org/10.1177/0018720820922080

Hayes, A. F. (2013). Mediation, moderation, and conditional process analysis: Regression-based approaches for clinical research. In A. G. C. Wright, & M. N. Hallquist (Eds.), *Camb. Handb. Res. Methods clin. Psychol.* (1st ed., pp. 396–414). Cambridge University Press. https://doi.org/10.1017/9781316995808.037.

Hook, J. N., Worthington, E. L., Utsey, S. O., Davis, D. E., & Burnette, J. L. (2012). Collectivistic self-construal and forgiveness, couns. *Values, 57*, 109–124. https://doi.org/10.1002/j.2161-007X.2012.00012.x

Hosseini, S. M. F., Lettinga, D., Vasey, E., Zheng, Z., Jeon, M., Park, C. H., & Howard, A. M. (2017). Both "look and feel" matter: Essential factors for robotic companionship. In *2017 26th IEEE int. Symp. Robot hum. Interact. Commun* (pp. 150–155). IEEE, Lisbon: RO-MAN. https://doi.org/10.1109/ROMAN.2017.8172294.

Ivaldi, S., Lefort, S., Peters, J., Chetouani, M., Provasi, J., & Zibetti, E. (2017). Towards engagement models that consider individual factors in HRI: On the relation of extroversion and negative attitude towards robots to gaze and speech during a human–robot assembly task: Experiments with the iCub humanoid. *Int. J. Soc. Robot., 9*, 63–86. https://doi.org/10.1007/s12369-016-0357-8

James, J., Balamurali, B. T., Watson, C. I., & MacDonald, B. (2021). Empathetic speech synthesis and testing for healthcare robots. *Int. J. Soc. Robot., 13*, 2119–2137. https://doi.org/10.1007/s12369-020-00691-4

Kiesler, S., Powers, A., Fussell, S. R., & Torrey, C. (2008). Anthropomorphic interactions with a robot and robot–like agent. *Social Cognition, 26*, 169–181. https://doi.org/10.1521/soco.2008.26.2.169

Kim, Y., Kwak, S. S., & Kim, M. (2013). Am I acceptable to you? Effect of a robot's verbal language forms on people's social distance from robots. *Computers in Human Behavior, 29*, 1091–1101. https://doi.org/10.1016/j.chb.2012.10.001

Klüber, K., & Onnasch, L. (2022). Appearance is not everything - preferred feature combinations for care robots. *Computers in Human Behavior, 128*, Article 107128. https://doi.org/10.1016/j.chb.2021.107128

Kohn, S. C., De Visser, E. J., Wiese, E., Lee, Y.-C., & Shaw, T. H. (2021). Measurement of trust in automation: A narrative review and reference guide. *Frontiers in Psychology, 12*, Article 604977. https://doi.org/10.3389/fpsyg.2021.604977

Kühne, K., Fischer, M. H., & Zhou, Y. (2020). The human takes it all: Humanlike synthesized voices are perceived as less eerie and more likable. Evidence from a subjective ratings study. *Frontiers in Neurorobotics, 14*, Article 593732. https://doi.org/10.3389/fnbot.2020.593732

Lankton, N., McKnight, D. H., & Tripp, J. (2015). Technology, humanness, and trust: Rethinking trust in technology. *Journal of the Association for Information Systems, 16*, 880–918. https://doi.org/10.17705/1jais.00411

Lee, J. D., & See, K. A. (2004). Trust in automation: Designing for appropriate reliance. *Human Factors*.

Lejuez, C. W., Read, J. P., Kahler, C. W., Richards, J. B., Ramsey, S. E., Stuart, G. L., Strong, D. R., & Brown, R. A. (2002). Evaluation of a behavioral measure of risk taking: The balloon Analogue risk task (BART). *Journal of Experimental Psychology: Applied, 8*, 75–84. https://doi.org/10.1037/1076-898X.8.2.75

Leyzberg, D., Avrunin, E., Liu, J., & Scassellati, B. (2011). Robots that express emotion elicit better human teaching. In *Proc. 6th int. Conf. Hum.-Robot interact.* (pp. 347–354). Lausanne Switzerland: ACM. https://doi.org/10.1145/1957656.1957789.

Li, J. (2015). The benefit of being physically present: A survey of experimental works comparing copresent robots, telepresent robots and virtual agents. *International Journal of Human-Computer Studies, 77*, 23–37. https://doi.org/10.1016/j.ijhcs.2015.01.001

Li, Y., Ishi, C. T., Inoue, K., Nakamura, S., & Kawahara, T. (2019). Expressing reactive emotion based on multimodal emotion recognition for natural conversation in human–robot interaction. *Advanced Robotics, 33*, 1030–1041. https://doi.org/10.1080/01691864.2019.1667872

Li, Y., & Lai, C. (2022). Robotic speech synthesis: Perspectives on interactions, scenarios, and Ethics. https://doi.org/10.48550/arXiv.2203.09599.

Lim, V., Rooksby, M., & Cross, E. S. (2021). Social robots on a global stage: Establishing a role for culture during human–robot interaction. *Int. J. Soc. Robot., 13*, 1307–1333. https://doi.org/10.1007/s12369-020-00710-4

Malle, B. F., & Ullman, D. (2021). A multidimensional conception and measure of human-robot trust. In *Trust hum.-robot interact.* (pp. 3–25). Elsevier. https://doi.org/10.1016/B978-0-12-819472-0.00001-0.

McGinn, C., & Torre, I. (2019). Can you tell the robot by the voice? An exploratory study on the role of voice in the perception of robots. In *2019 14th ACMIEEE int. Conf. Hum.-Robot interact. HRI* (pp. 211–221). Daegu, Korea (South): IEEE. https://doi.org/10.1109/HRI.2019.8673305.

Mertler, C. A., Vannatta, R. A., & LaVenia, K. N. (2021). *Advanced and multivariate statistical methods: Practical application and interpretation* (7th ed.). New York: Routledge. https://doi.org/10.4324/9781003047223

Mori, M. (1970). Bukimi no tani [the uncanny valley]. *Energy, 7*, 33–35.

Niculescu, A., van Dijk, B., Nijholt, A., Li, H., & See, S. L. (2013). Making social robots more attractive: The effects of voice pitch, humor and empathy. *Int. J. Soc. Robot., 5*, 171–191. https://doi.org/10.1007/s12369-012-0171-x

Nishio, S., & Ishiguro, H. (2011). Attitude change induced by different appearances of interaction agents. *Int. J. Mach. Conscious. 03*, 115–126. https://doi.org/10.1142/S1793843011000637

Nomura, T. (2017). Cultural differences in social acceptance of robots. In *2017 26th IEEE int. Symp. Robot hum. Interact. Commun* (pp. 534–538). IEEE, Lisbon: RO-MAN. https://doi.org/10.1109/ROMAN.2017.8172354.

Oertel, C., Castellano, G., Chetouani, M., Nasir, J., Obaid, M., Pelachaud, C., & Peters, C. (2020). Engagement in human-agent interaction: An overview. *Front. Robot., AI 7*, 92. https://doi.org/10.3389/frobt.2020.00092

Okafuji, Y., Ishikawa, T., Matsumura, K., Baba, J., & Nakanishi, J. (2024). Pseudo-eating behavior of service robot to improve the trustworthiness of product recommendations. *Advanced Robotics, 38*, 343–356. https://doi.org/10.1080/01691864.2024.2321191

Onnasch, L., & Roesler, E. (2021). A taxonomy to structure and analyze human–robot interaction. *Int. J. Soc. Robot., 13*, 833–849. https://doi.org/10.1007/s12369-020-00666-5

Picard, R. W. (2000). *Affective computing, 1* (paperback ed). Cambridge, Mass: MIT Press.

Poivet, R., Lopez Malet, M., Pelachaud, C., & Auvray, M. (2023). The influence of conversational agents' role and communication style on user experience. *Frontiers in Psychology, 14*, Article 1266186. https://doi.org/10.3389/fpsyg.2023.1266186

Rezaei Khavas, Z., Kotturu, M. R., Ahmadzadeh, S. R., & Robinette, P. (2024). Do humans trust robots that violate moral trust? *ACM Trans. Hum.-Robot Interact., 13*, 1–30. https://doi.org/10.1145/3651992

Robinette, P., Li, W., Allen, R., Howard, A. M., & Wagner, A. R. (2016). Overtrust of robots in emergency evacuation scenarios. In *2016 11th ACMIEEE int. Conf. Hum.-Robot interact. HRI* (pp. 101–108). Christchurch, New Zealand: IEEE. https://doi.org/10.1109/HRI.2016.7451740.

Roesler, E., Manzey, D., & Onnasch, L. (2021). A meta-analysis on the effectiveness of anthropomorphism in human-robot interaction. *Science Robotics, 6*, Article eabj5425. https://doi.org/10.1126/scirobotics.abj5425

Roesler, E., Manzey, D., & Onnasch, L. (2022). Embodiment matters in social HRI research: Effectiveness of anthropomorphism on subjective and objective outcomes. *ACM Trans. Hum.-Robot Interact.*. , Article 3555812. https://doi.org/10.1145/3555812

Roesler, E., Naendrup-Poell, L., Manzey, D., & Onnasch, L. (2022). Why context matters: The influence of application domain on preferred degree of anthropomorphism and gender attribution in human–robot interaction, int. *J. Soc. Robot., 14*, 1155–1166. https://doi.org/10.1007/s12369-021-00860-z

Salem, M., Lakatos, G., Amirabdollahian, F., & Dautenhahn, K. (2015). Would you trust a (faulty) robot?: Effects of error, task type and personality on human-robot cooperation and trust. In *Proc. Tenth annu. ACMIEEE int. Conf. Hum.-Robot interact.* (pp. 141–148). Portland Oregon USA: ACM. https://doi.org/10.1145/2696454.2696497.

Sanghvi, J., Castellano, G., Leite, I., Pereira, A., McOwan, P. W., & Paiva, A. (2011). Automatic analysis of affective postures and body motion to detect engagement with a game companion. In *Proc. 6th int. Conf. Hum.-Robot interact.* (pp. 305–312). Lausanne Switzerland: ACM. https://doi.org/10.1145/1957656.1957781.

Schreibelmayr, S., & Mara, M. (2022). Robot voices in daily life: Vocal human-likeness and application context as determinants of user acceptance. *Frontiers in Psychology, 13*, Article 787499. https://doi.org/10.3389/fpsyg.2022.787499

Taylor, P. A. (2009). *Text-to-speech synthesis*. Cambridge, UK ; New York: Cambridge University Press.

Tielman, M., Neerincx, M., Meyer, J.-J., & Looije, R. (2014). Adaptive emotional expression in robot-child interaction. In *Proc. 2014 ACMIEEE int. Conf. Hum.-Robot interact.* (pp. 407–414). ACM, Bielefeld Germany. https://doi.org/10.1145/2559636.2559663.

Torre, I., Goslin, J., White, L., & Zanatto, D. (2018). Trust in artificial voices: A "congruency effect" of first impressions and behavioural experience. In *Proc. Technol. Mind soc.* (pp. 1–6). Washington DC USA: ACM. https://doi.org/10.1145/3183654.3183691.

Verma, J. P. (2015). *Repeated measures design for empirical researchers*. John Wiley Sons.

Waytz, A., Cacioppo, J., & Epley, N. (2010). Who sees human?: The stability and importance of individual differences in anthropomorphism. *Perspectives on Psychological Science, 5*, 219–232. https://doi.org/10.1177/1745691610369336

Yam, K. C., Bigman, Y. E., Tang, P. M., Ilies, R., De Cremer, D., Soh, H., & Gray, K. (2021). Robots at work: People prefer—and forgive—service robots with perceived feelings. *Journal of Applied Psychology, 106*, 1557–1572. https://doi.org/10.1037/apl0000834

Zechmeister, J. S., Garcia, S., Romero, C., & Vas, S. N. (2004). *Don't apologize unless you mean it: A laboratory investigation of forgiveness and retaliation*.