Laughter in Sight: How Gaze and Laughter Affect Perceptions of a Social Robot

Eleni Giannitzi[1,2](✉), Vladislav Maraev[2], Erik Lagerstedt[2], and Christine Howes[2]

[1] Department of Computer Science, KU Leuven, Leuven, Belgium
eleni.giannitzi@kuleuven.be
[2] Department of Linguistics, Philosophy and Theory of Science, University of Gothenburg, Gothenburg, Sweden
{gusgianel,vladislav.maraev,erik.lagerstedt,christine.howes}@gu.se

Abstract. This study examines the role of gaze and laughter coordination in human-robot interaction, focusing on how these non-verbal cues influence user perception of social robots. Using Furhat, we explore whether contextually appropriate alignment of gaze and laughter enhances the interaction quality in terms of human perception and emotional responses. Participants were divided into two experimental conditions – one experiencing well-aligned gaze and laughter, and the other encountering misaligned behaviours – while discussing a cooking activity with Furhat. Their interactions were recorded, followed by a questionnaire that assessed their perceptions and emotional responses. Results showed that participants exposed to contextually appropriate gaze-laughter alignment rated Furhat higher in empathy, naturalness and compassion compared to those who experienced the same behaviours in inappropriate contexts. Our findings suggest promising potential for designing more human-like social robots capable of meaningful non-verbal communication.

Keywords: Gaze · Laughter · SIA · Social Robots · Furhat · human-computer interaction

1 Introduction

Following technological advancements, human-machine interaction has become a central topic of scientific discussion. Socially Interactive Agents (SIAs), transitioning from two-dimensional interfaces to embodied agents in three-dimensions, are now being deployed in education, healthcare, and customer service, where they are expected to engage with humans naturally and intuitively. Advanced systems like Furhat demonstrate human-like interaction capabilities by integrating gaze, head-neck movements, and facial expressions. However, as noted by Zawieska [37], the integration of social robots into everyday life remains

M. Kurosu and A. Hashizume (Eds.): HCII 2025, LNCS 15769, pp. 249–268, 2025.
https://doi.org/10.1007/978-3-031-93861-0_16

limited, with their use focusing primarily on research settings rather than real-world human-robot interaction. In this study we aim to explore social and practical aspects of human engagement via two non-verbal cues of human communication, *gaze* and *laughter*, and examine the influence of their coordination on user perceptions of naturalness, empathy, and human-likeness in social robots.

Gaze plays a crucial role in social dynamics by signaling attentiveness, shaping conversational flow, and facilitating interaction. Similarly, laughter enhances communication by easing tension, expressing emotions, and fostering social bonds. It is a universal, non-verbal expression of emotion, closely tied to social context.

Though well-studied individually, gaze and laughter have rarely been analysed together in robot interactions. Expanding on Mazzocconi et al. [17], this study hypothesises that aligning gaze with laughter will make social robots appear more natural and human-like, enhancing interaction quality. Specifically, we ask:

RQ1. Does the placement of gaze-aligned laughter improve the user's contextual understanding?
RQ2. Is the effectiveness of coordinating laughter with gaze patterns context-dependent?
RQ3. Does this coordination enhance the perceived naturalness, human-likeness, and empathy of the social robot as experienced by the user?

2 Related Studies

2.1 Gaze and Laughter in Human Dialogue

Studies, like Rossano [27], have highlighted cultural variations in gaze behaviours, mutual gaze's role in maintaining conversation, and gaze aversion as a marker of cognitive effort or disengagement. Tiselius and Sneed [33] studied how interpreters use gaze aversion not only to manage cognitive load, but also to organize conversations and signal speaker transitions, particularly when interpreting in their weaker language. Gullberg and Holmqvist [9] showed that during a conversation, listeners rely on eye movements to better understand the speaker before deciding whether to ask for clarification. Similarly, Somashekarappa et al. [32] analysed gaze patterns in relation to speech acts, joint attention, and communication, finding correlations between speech and gaze at reference objects but not between participants' gaze at each other. The findings also highlight the role of mutual gaze in signalling cues like agreement.

Moving on laughter, this study employs Mazzocconi et al.'s [18] taxonomy, which categorises laughter as a response to pleasant incongruity (e.g., jokes) or social incongruity (e.g., criticism or sympathy). Their research illustrates how laughter functions as both a social signal and a communicative tool, conveying meaning, contributing to irony, and facilitating conversational repair, thereby enriching and disambiguating the interpretation of conversational contexts.

Social context is also critical, as people are significantly more likely to laugh with others than alone [29]. Beyond its social nature, laughter plays a functional role in emotional regulation. Different types of laughter, such as "mirthful" (genuine) and "social" (polite or filler), contribute to varied conversational dynamics [12]. Research by Clift [4] further emphasizes laughter's social function, showing how it mitigates conflict, softens negative assessments, and reinforces its role as a conversational tool distinct from humour.

Mazzocconi et al. [17] explored how different types of laughter, such as responses to pleasant or social incongruity align with distinct gaze patterns. Laughter linked to pleasant incongruity typically involves making eye contact before laughing, looking away during the laughter, and returning gaze afterward. In contrast, laughter related to social incongruity tends to occur during eye contact, with gaze aversion happening just before and after the laugh.

2.2 Socially Interactive Agents (SIAs)

Socially Interactive Agents (SIAs) are autonomous virtual or physical systems that use verbal and non-verbal communication to create natural interactions. Social robots, a subset of SIAs, are three-dimensional physically embodied agents that interact with humans through communication, cooperation, and decision-making tasks. Their behaviours are perceived as "social" based on the societal norms and they are intended to serve diverse roles, such as companions for older people, educators, and assistants in various contexts [11,13,14,37].

As stated previously, laughter and gaze are key communicative tools, yet their complexity poses challenges for integration into Socially Interactive Agents (SIAs). Bachorowski et al. [2] highlights the acoustic variability of laughter, with voiced laughter being more engaging than unvoiced laughter, as well as differences in pitch based on the individual's gender. Annotation challenges, such as overlapping laughter, further emphasize its coordination demands [34]. Similarly, gaze coordination, as studied by [26], enhances conversation when a shared visual context is present, influencing language use.

2.3 Gaze in SIAs

Implementing gaze behaviours is important for creating natural and intuitive interactions in SIAs. Studies have examined various aspects of gaze integration to enhance social presence and improve user engagement. For example, Parreira et al. [24] found that gaze patterns in the Furhat robot, such as looking at the speaker or using gaze aversion, influenced turn length, participation balance, and user comfort. Prolonged eye contact during speech caused discomfort, highlighting the need for careful gaze design. Similarly, Somashekarappa [31] investigated how different gaze patterns in social robots influence user engagement and perception, finding that gaze manipulations based on human-human interaction positively impact anthropomorphism and interaction quality.

Another study demonstrated that extroverts maintain longer eye contact than introverts and that mutual gaze increases speaking likelihood [1]. Finally, challenges in synchronising gaze with head movements, blinks, speech, and gestures for natural interactions have been highlighted [28].

2.4 Laughter in SIAs

Moving on to laughter, studies like Nijholt [21] highlighted the impact of humour on human interactions and its potential applications in SIA, noting key challenges in implementation, such as modelling appropriate humour responses and creating lifelike expressions of smiling and laughter. Other studies have looked at developing a system for generating shared laughter in conversational robots to improve empathy and naturalness [12]. Using models for laughter detection, shared laughter prediction, and laughter type selection, the system determined when and how to laugh (e.g., pleasant vs. social laughter). Additionally, Becker-Asano and Ishiguro [3] investigated the naturalness of various types of female laughter in humanoid robots, while Cosentino et al. [5] and Türker et al. [35] explored the impact of laughter-responsive robots. Their findings showed that robots capable of detecting and responding to human laughter increased interaction frequency, backchannel events (e.g., laughter and smiles), and communication pace. Lastly, Maraev et al. [16] demonstrated the benefits of integrating non-humorous laughter into task-oriented spoken dialogue systems, showing that laughter can improve the interaction quality by managing communication failures and providing natural feedback.

2.5 Laughter and Gaze Coordination in SIAs

Research into the coordination of gaze and laughter is a promising area of research, focusing on their interaction in human communication and potential applications in SIAs.

Even though the available research on this topic is limited, there is a noticeable direction in this area. For example, Becker-Asano and Ishiguro [3] investigated the effects of laughter in interactions with the android robot Geminoid HI-1 among thirty-six Japanese university students. Participants experienced both a control condition and a laughter condition, where Geminoid laughed in response to jokes. Additionally, the study incorporated gaze coordination with laughter, as Geminoid HI-1 was programmed to laugh while simultaneously directing its gaze toward participants during interactions. The results demonstrated that laughter, particularly when combined with directed gaze, effected the participants' perceptions of the robot.

3 Methodology

The study investigates how gaze and laughter coordination in human-robot interaction affects the user perception of social robots using Furhat. Participants were guided by a robot while following a recipe generated by GPT-4 [22,23], with the robot integrating laughter and gaze functions to simulate human-like behaviour. The methodology was informed by a pilot study (reported in [7,8])

The Robot. In this study we used a Furhat [20], a human-like robot head designed for rich, multi-modal dialogue. Using back-projected facial animation on a 3D mask, Furhat supports synchronised speech, facial expressions, and attention control for natural multiparty interactions. We controlled Furhat via Remote API with a statecharts-based [10] dialogue manager which enables event-driven dialogues and mixed-initiative conversations where both the system and users could initiate exchanges. Furhat's attention control mechanism tracked the system's and its interlocutors' attention, enabling smooth communication.

Experimental Conditions. Two conditions were tested: correct and incorrect laughter timing. In the correct condition, three fixed laughter functions were inserted into the dialogue precisely when Furhat made a deliberate mistake, overlooked a tool, or joked about lacking hands – aligning with Mazzocconi et al. [18]'s taxonomy on social and pleasant incongruity. In the incorrect condition, these same laughter events were deliberately misplaced (e.g., occurring after neutral statements), with manual interventions either triggered asynchronously or not at all to avoid bias. Intermediate states containing only gaze and facial expressions gave Furhat time to display non-verbal behaviours or pause, allowing the experimenter to press buttons for extra laughter triggers if needed (See Sect. 3.1 for more details), while maintaining overall conversational flow.

Participants. Nineteen participants (10 men, 7 women, and 2 selecting the option "prefer not to say", aged 20–40, were recruited via social media and group chats. A poster detailing the study's objectives, ethical approval, and participation requirements was shared among English-speaking individuals interested in technology and social robots. To encourage participation, the poster mentioned that fika (coffee, beverages, and snacks) would be provided as a reward instead of monetary compensation. Participants were randomly assigned to correct (10 participants) or incorrect (9 participants) laughter-gaze placement conditions. They received a general explanation of the study but were not informed about the specific manipulation to minimize bias. Upon arrival, participants reviewed and signed a consent form, then received an instruction sheet alongside a brief demonstration of the recipe task.

Task. The participants interacted with Furhat in a simulated and collaborative cooking activity, in which they generated and followed a recipe (checking/gathering ingredients and tools, and executing steps). Throughout the interaction, Furhat demonstrated gaze movements and laughter. Post interaction, participants completed a questionnaire evaluating metrics such as the robot's empathy, naturalness, human-likeness, and the user's emotional responses.

Experimental Setup. Each participant interacted with Furhat for 15–20 min in a controlled room equipped with three cameras for multi-angle recording. Furhat was stationed on a table with a monitor displaying GPT-generated recipes and a microphone for audio capture; cooking tools were arranged realistically to enhance engagement (Fig. 1). Following the interaction, participants completed a 10–15 min questionnaire. Finally, they were invited for coffee and snacks in a less formal setting, where they systematically shared feedback about their experience. Each participant was asked about their overall impression of the interaction, any elements that stood out or felt unusual, and their thoughts on Furhat's gaze and laughter separately. Additionally, they were encouraged to reflect on any patterns they may have noticed during the robot's laughter, without being explicitly guided toward context or expression. This session provided additional qualitative insights, and the entire experiment lasted 45–60 min overall.

Fig. 1. Experiment Set-up

Questionnaire. The post-experiment questionnaire addressed both interaction experience and emotional response. In the first half, participants rated Furhat's empathy, authenticity, human-like behaviour, rapport, and overall satisfaction on a 1–10 scale [12,25,36], allowing statistical analysis while preserving the subjective nature of individual responses. In the second half, participants used the Geneva Wheel of Emotions (GWE), following [3], to identify which of the twenty listed emotions (e.g., joy, frustration, interest) they experienced and rate their intensity on a 0–5 circular scale, capturing emotional variations across experimental conditions. Finally, an open-comments section allowed participants to provide additional qualitative feedback, complementing the structured data.

3.1 Gaze and Laughter Implementation

Human interaction uniquely integrates non-verbal cues with verbal communication, relying on this alignment to fully convey meaning. This study's conversation design coordinated both gaze and laughter, informed by the GHI corpus [17] for laughter annotations and by [32] for multi-modal gaze patterns (e.g., looking around the table). Additionally, guidelines from [31] on Furhat's eye movements shaped the frequency and patterns of eye transitions.

Facial Expressions. For this study, Furhat's facial expressions – provided by Furhat Robotics [6]'s SDK Library – (such as thoughtfulness, happiness, and surprise) were refined to enhance our perception of their naturalness. This included extending their duration, adding smoother transitions, and incorporating gaze movements. The surprise expression featured widened eyes and an upward glance, while thoughtfulness included shifting gaze, furrowed brows, and small pauses (e.g., "hmm"). Each expression was triggered automatically at contextually relevant dialogue moments aiming to replicate human-like reactions (such as a thoughtful look while awaiting GPT-4 responses or surprise when a user spotted an "unexpected" ingredient quantity). Additionally, two filler pause variations distinguished deeper thinking (longer pause, typically when waiting for GPT-4) from brief consideration (shorter pause, such as during small talk).

Gaze Patterns. Furhat's gaze patterns included looking left, right, up, and down, both with and without neck movement, to replicate fluid human eye and head motions. These ranged from simple transitions (e.g., right to left) to more dynamic combinations (e.g., left to right while tilting the neck down-left), creating a sense of continuous scanning and attention. Gaze cues helped guide the user's focus – such as when searching for specific cooking tools – and were coordinated with facial expressions and sounds to reflect different states of thinking. This multilayered approach supported natural conversation flow, reinforced user engagement, and heightened the illusion of human-like behaviour.

Laughter Functions. Two distinct laughter functions were implemented in Furhat, following Mazzocconi et al. [17] observations in human-human dialogue – one reflecting social incongruity (gaze initially averted, then eye contact during laughter, followed by gaze aversion) and the other pleasant incongruity (eye contact first, brief gaze aversion mid-laugh, then returning eye contact) (Fig. 2). Both functions used the same facial expressions (e.g., narrowed eyes, arched brows, wrinkled nose) with synchronised lip, eye, and neck movements. These were triggered automatically and were embedded as robot-initiated and fixed "laugh states" within the dialogue – one each in the ingredient, tool, and step sections (See Sect. 3.2 for details). However, to accommodate unpredicted user-initiated laughter, three function buttons were available in each state: one for triggering a separate social laughter state, one for triggering a separate pleasant laughter state, and one for terminating the experiment if needed. The additional laugh states were manually triggered by the experimenter, allowing the robot to respond to user-initiated laughter in real time and ensuring the interaction remained contextually appropriate and smooth. The experimenter observed the interaction from a separate room through a one-way mirror and was not physically present with the participants. It is important to note that this manual intervention was intended primarily as a safeguard to address potential issues with the program and was rarely used during interactions.

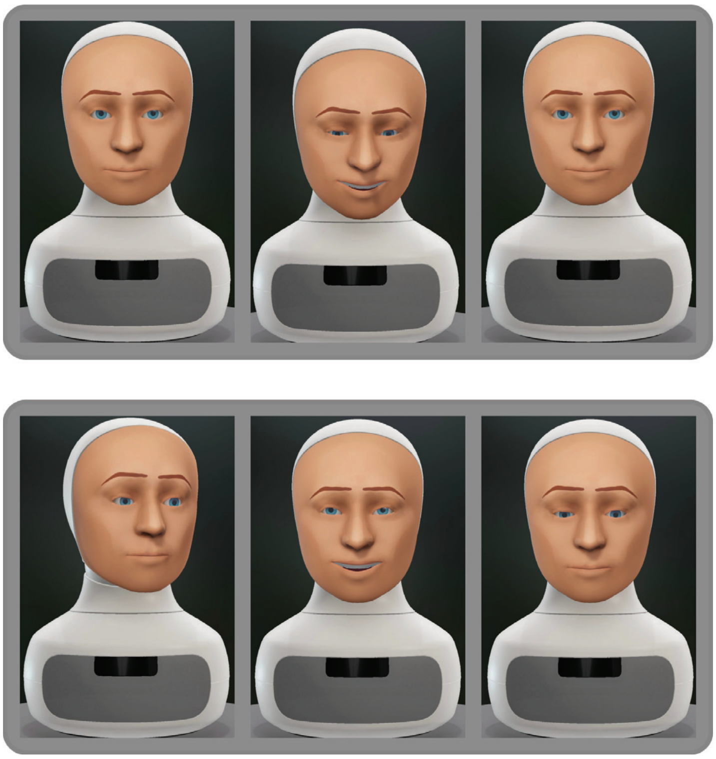

Fig. 2. Visual Representation of Laughter as a response to Pleasant Incongruity (top) and Social Incongruity (bottom).

3.2 Dialogue Implementation

The dialogue system used a TypeScript-based back-end and the XState library for statechart implementation [10], along with Furhat Robotics resources (remote API, Swagger Editor, SDK). To address limitations observed in a pilot study [8] [7], GPT-4's role was minimised, transitioning from the primary conversational driver to a supportive tool. GPT-4 was used selectively for generating recipes and providing supplementary information – such as a cooking tool replacement, while most of the dialogue was scripted to reduce issues like repetition and inconsistency.

Initial State. For the session, two personalities were implemented in Furhat: Alicia – a female instructor persona – introduced the session, provided essential instructions, and then handed off to Mathew, a male cooking partner persona who guided users through the cooking interaction. Both personas operated on the same physical Furhat robot, but with distinct visual appearances and voice profiles. The transition between Alicia and Mathew was performed instantaneously using Furhat's animation graphics, which switched Alicia's face and voice to Mathew's. To ensure a smooth experience, Alicia informed the user in advance that the profile would soon change to Mathew. This transition marked the experiment's official start, beginning with the robot asking for the user's name to familiarise them with Furhat's speech recognition and set a friendly tone. To ensure privacy and reduce discomfort, GPT-4 generated a playful "food-related nickname" (e.g., Elena Lasagna or Lollipop Lloyd) based on the user's input, adding an element of fun and minimizing formality. This nickname remained consistent throughout the session. When users chose not to provide a name or Furhat's Automatic Speech Recognition failed to capture their input, Furhat generated a random nickname.

Recipe Generation. Once the initial setup was complete, Furhat asked participants to name any savoury or sweet dish, which became the basis for a GPT-4 real-time generated recipe. A prompt extracted the dish name, and a second prompt created a JSON recipe with five ingredients, five tools, and ten steps, formatted and limited by strict criteria (e.g., quantities in cups or spoons) (Fig. 3). The generated output was split into two parts: a text file (.txt) displayed to participants and a JSON file, parsed into code variables for Furhat's dialogue. This implementation helped maintain a natural conversational flow, dividing the interaction into ingredient, tool, and step "states". To ensure a

```
Task: Generate a ${input.dishidea} recipe with 5 ingredients, 5 cooking tools, and ONLY 10 steps in
a JSON format, e.g.,
{
    "ingredients": {
        "1": {"ingredient": "Sushi rice", "quantity": "1 cup"},
        "2": {"ingredient": "Nori sheets", "quantity": "2"},
        "3": {"ingredient": "Fresh tuna", "quantity": "1 cup"},
        "4": {"ingredient": "Cucumber", "quantity": "1"},
        "5": {"ingredient": "Sriracha sauce", "quantity": "to taste"}
    },
    "cooking_tools": {
        "1": {"tool": "Rice cooker", "quantity": "1"},
        "2": {"tool": "Bamboo sushi rolling mat", "quantity": "1"},
        "3": {"tool": "Sharp knife", "quantity": "1"},
        "4": {"tool": "Cutting board", "quantity": "1"},
        "5": {"tool": "Small bowl of water", "quantity": "1"}
    },
    "steps": {
        "1": "Prepare sushi rice according to package instructions.",
        "2": "Place a sheet of nori on the bamboo sushi rolling mat with the shiny side facing
down.",
        "3": "Spread a thin layer of sushi rice evenly over the nori, leaving a small border at the
top edge.",
        "4": "Arrange the sliced tuna and julienned cucumber in a line across the center of the
rice.",
        "5": "Drizzle Sriracha sauce over the tuna and cucumber. Roll the sushi tightly using the
bamboo mat, moistening the top edge of the nori with water to seal."
    },
}. NO LIST. GIVE THE QUANTITY OF INGREDIENTS IN CUPS AND SPOONS. START WITH THE {. THE END NEEDS 2 }.
DO NOT INCLUDE THESE ```. REMEMBER IT'S A JSON FORMAT.
```

Fig. 3. Recipe Prompt passed on ChatGPT4

smooth user experience, the recipe text was visibly displayed on a monitor, and Furhat was programmed to wait if needed before proceeding.

Subdialogues

Ingredients Interaction. During the ingredient states, Furhat prompted participants to list each ingredient one by one, verifying quantities and occasionally introducing deliberate errors (e.g., 9 kg instead of 2 cups) to simulate real-world mistakes. Participants could switch roles (saying "next" or "help") or request Furhat to "look again," highlighting Furhat's adaptability and facial expressions. Internally, Furhat tracked each ingredient with dedicated variables, ensuring logical progression and a smooth transition to the cooking tools state.

Tools Interaction. Similar to ingredients, Furhat guided participants through verifying tools scattered on a table, enabling visual interaction. If a listed tool was unavailable, Furhat prompted GPT-4 for alternatives (e.g., a microwave or an air-fryer instead of an oven), reflecting real-world problem-solving. As before, mixed initiative was possible and once all tools were checked off, the dialogue proceeded to the cooking steps state.

Steps Interaction. During the cooking steps, participants read each recipe step aloud, which Furhat passed to GPT-4 for validation. Furhat filled in gaps for incomplete steps or confirmed accurate ones before proceeding. Users could ask Furhat to read the recipe, leveraging the role-switch feature. Each step was followed by a three- to five-second pause for execution, punctuated by GPT-4-generated fun facts or scripted cooking-related small talk. The interaction alternated between steps, pauses, and small talk until all ten steps were completed.

Finalization. After the final step, Furhat switched back to Alicia, who formally concluded the session, thanked participants, and reminded them to complete the post-experiment questionnaire. Alternating between Alicia and Mathew distinguished the experiment phases and clearly signaled the session's start and end.

4 Results

The questionnaire data revealed a strong preference for contextually correct laughter, with better ratings across most metrics in the Correct Laugh (CL) condition compared to the Wrong Laugh (WL) condition. To analyse these trends, quantitative methods including paired t-tests, standard deviation, ANOVA significance analysis, and Pearson Correlation were applied, followed by qualitative insights from participant feedback, experimenter observations, and video recordings. This dual analysis explores the role of gaze-laughter coordination in enhancing engagement and user experience with Furhat. Throughout this analysis, "laughter" refers specifically to Furhat's aligned gaze-laughter functions, ensuring clarity and focus on these interaction dynamics.

4.1 Quantitative Analysis

The quantitative analysis examines how CL and WL placement influenced participants' perceptions and emotions. We divide the analysis into two subsections: User Experience and Perception (e.g., empathy, naturalness, rapport) and Emotional Experiences (e.g., joy, pleasure, compassion).

Table 1. User experience and perception in CL and WL conditions: Mean Score, Standard Deviation, Mean Difference (Δ), and ANOVA Significance (p). Statistically significant results ($p < 0.05$) are marked with an asterisk (*), and marginally significant results ($0.05 \leq p < 0.1$) are indicated with a dagger (†).

Metric (1-10)	CL, mean (SD)	WL, mean (SD)	Δ	p
Empathy	7.40 (1.58)	5.89 (1.17)	1.51	0.031*
Naturalness and Authenticity	6.10 (0.99)	4.89 (1.54)	1.21	0.055†
Human-Likeness	6.40 (1.65)	5.89 (1.17)	0.51	0.451
Understanding	7.00 (1.94)	7.56 (1.74)	−0.56	0.522
Naturalness of Laughter	6.20 (2.30)	4.22 (1.86)	1.98	0.056†
Perceived Rapport	6.40 (1.17)	6.67 (1.50)	−0.27	0.670
Perceived Contribution	5.30 (1.83)	6.22 (2.11)	−0.92	0.321
Perceived Interest from Furhat	5.40 (2.91)	5.44 (2.40)	−0.04	0.972
Perceived Connection	6.30 (1.70)	5.22 (1.56)	1.08	0.170
Perceived Mutual Understanding	6.40 (1.51)	6.89 (1.05)	−0.49	0.429
Ability to express oneself fully	4.70 (2.79)	6.00 (1.94)	−1.30	0.260
Perceived Warmth & Care	7.50 (2.01)	6.44 (2.19)	1.06	0.288
Perceived Respect	8.70 (2.26)	8.22 (1.09)	0.48	0.573
Perceived Frustration	4.00 (2.67)	2.33 (1.22)	1.67	0.104
Overall Satisfaction	7.00 (1.83)	7.00 (1.00)	0.00	1.000

Users Experience and Perception. Key metrics from the questionnaire, including empathy, naturalness, human-likeness, rapport, and satisfaction, were analyzed across the CL and WL conditions (Table 1). Participants rated *Empathy* greater in the CL condition (mean = 7.40, SD = 1.58) compared to WL (mean = 5.89, SD = 1.17; $F(1,17) = 5.524, p = 0.031$), with lower variability in CL indicating more consistent positive perceptions. *Naturalness and Authenticity* were similarly rated marginally more significant in CL (mean = 6.10, SD = 0.99) than in WL (mean = 4.89, SD = 1.54; $F(1,17) = 4.250, p = 0.055$), suggesting the importance of appropriate laughter placement. Ratings for the *Naturalness of Laughter* were also marginally higher in CL (mean = 6.20, SD = 2.30) than WL (mean = 4.22, SD = 1.86; $F(1,17) = 4.191, p = 0.056$). There were no other significant differences.

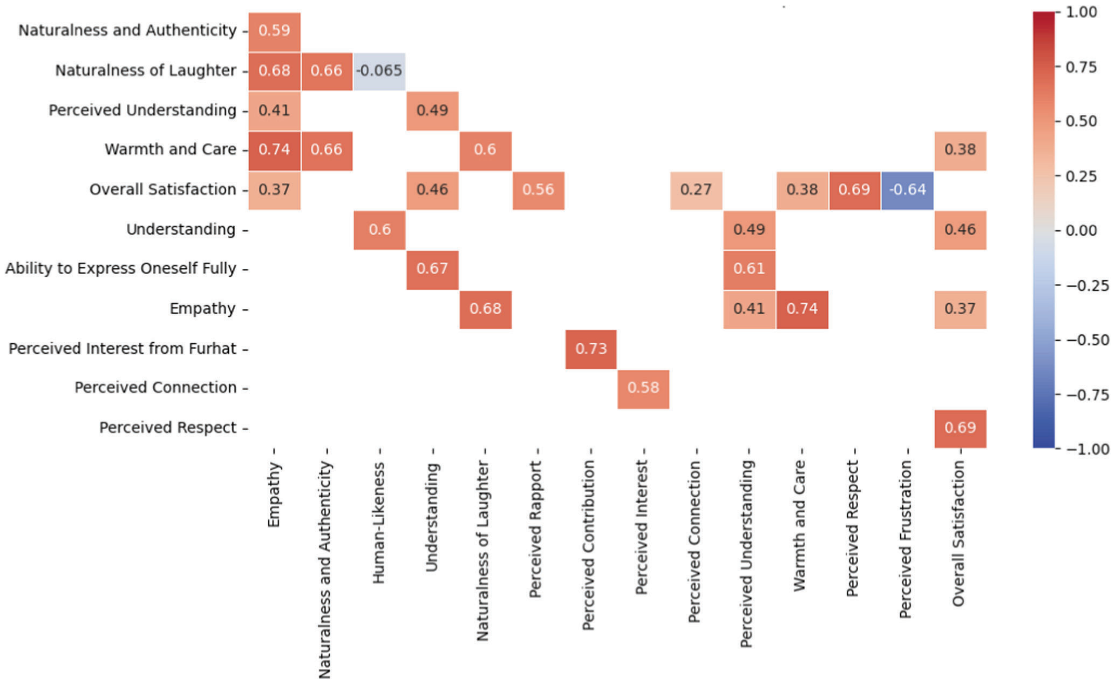

Fig. 4. Pearson's Correlation Coefficient Between Interaction Key Metrics. The table shows only a selection of correlations that appeared to influence participants' perceptions of their interaction with Furhat the most.

Metrics' Correlations. Key correlations also emerged from the data (Fig. 4). Empathy was correlated with naturalness and authenticity ($r = 0.592$) and the naturalness of laughter ($r = 0.678$), emphasising the role of contextually appropriate laughter in enhancing user perceptions. Naturalness and authenticity correlated with laughter's naturalness ($r = 0.655$) and warmth ($r = 0.664$), further highlighting laughter's importance in maintaining engaging and natural interactions. Human-likeness and understanding showed a moderate correlation ($r = 0.604$), underscoring the value of aligned non-verbal cues. Finally, satisfaction and frustration were negatively correlated ($r = -0.636$), demonstrating that frustration significantly detracts from the overall user experience.

Users Emotional Response. The key emotions presented in the Geneva Wheel of Emotions and the two conditions – CL and WL – to understand how these conditions influenced participants' emotional responses. Positive emotions, such as joy, pleasure, and compassion, were examined alongside negative emotions, including disappointment, anger, and contempt. The results show that compassion was the only emotion to show a statistically significant difference between the two conditions. Participants reported higher compassion in the CL condition (mean = 3.30, SD = 1.57) compared to the WL condition (mean = 1.56, SD = 1.33), with an ANOVA p-value of $F(1, 17) = 6.747, p = 0.019$. This indicates that contextually appropriate laughter strongly enhances feelings of compassion, while misplaced laughter diminishes it.

Table 2. GWE Results in CL and WL Conditions: Mean Score, Standard Deviation, Mean Difference, and ANOVA Significance. Statistically significant results ($p < 0.05$) are marked with an asterisk (*).

Emotion (0-5)	CL, mean (SD)	WL, mean (SD)	Δ	p
Interest	4.40 (1.07)	4.44 (0.53)	−0.04	0.912
Amusement	4.40 (0.70)	4.11 (0.78)	0.29	0.407
Pride	2.00 (2.05)	1.56 (1.59)	0.44	0.608
Joy	3.50 (1.78)	3.22 (1.48)	0.28	0.718
Pleasure	3.20 (1.32)	2.67 (1.73)	0.53	0.457
Contentment	2.70 (2.16)	2.89 (1.36)	−0.19	0.825
Love	0.90 (1.45)	0.56 (1.01)	0.34	0.561
Admiration	2.60 (1.84)	2.67 (1.80)	−0.07	0.937
Relief	1.10 (1.66)	1.00 (1.32)	0.10	0.887
Compassion	3.30 (1.57)	1.56 (1.33)	1.74	0.019*
Sadness	0.20 (0.42)	0.33 (0.71)	−0.13	0.620
Guilt	0.10 (0.32)	0.00 (0.00)	0.10	0.357
Regret	0.20 (0.42)	0.11 (0.33)	0.09	0.620
Shame	0.50 (0.85)	0.44 (1.01)	0.06	0.898
Disappointment	1.20 (1.40)	0.67 (1.32)	0.53	0.406
Fear	0.10 (0.32)	0.33 (0.71)	−0.23	0.357
Disgust	0.10 (0.32)	0.00 (0.00)	0.10	0.357
Contempt	0.40 (0.97)	0.44 (1.33)	−0.04	0.934
Hate	0.20 (0.42)	0.00 (0.00)	0.20	0.174
Anger	0.50 (1.27)	0.33 (0.71)	0.17	0.732

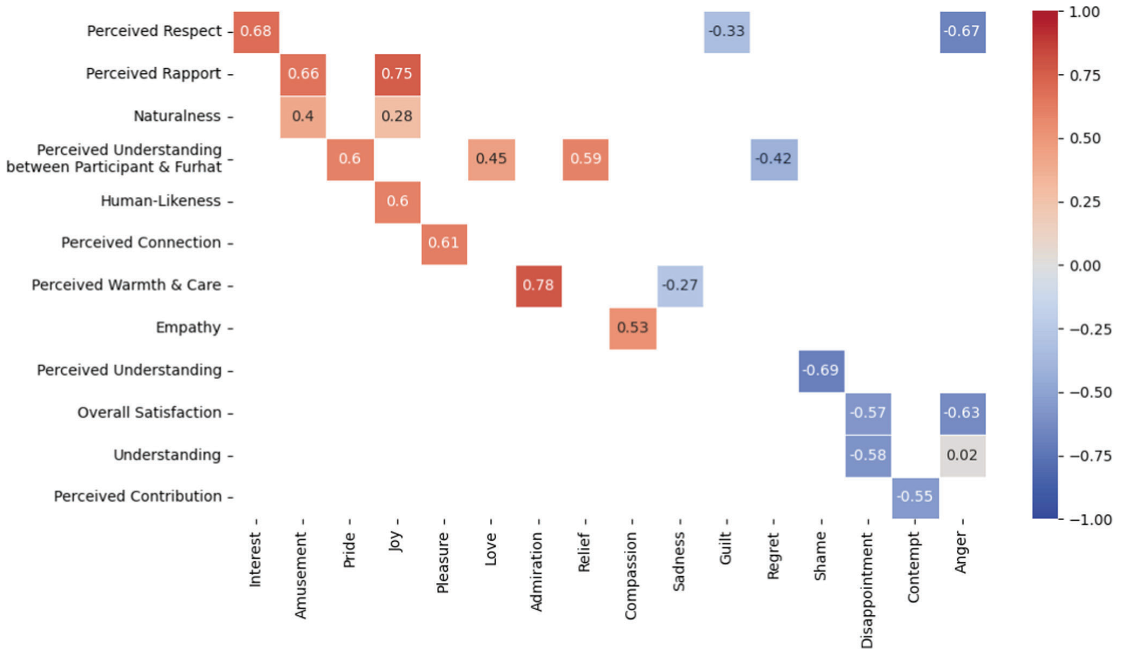

Fig. 5. Pearson's Correlation Coefficient Between Interaction Key Metrics and GWE. The table shows only a selection of correlations that appeared to influence participants' perceptions of their interaction with Furhat the most.

Emotions' Correlations. Correlation analysis between emotions and key interaction metrics provides deeper insights into the interplay between emotional responses and perceptions of Furhat's behaviour. Positive emotions, such as joy, amusement, and compassion, were strongly correlated with metrics like perceived rapport, warmth, and empathy. For example, joy showed a strong positive correlation with perceived rapport ($r = 0.75$), while compassion was linked to empathy ($r = 0.53$). These findings underscore the importance of aligning non-verbal cues like laughter with the conversational flow to foster positive emotional experiences. Conversely, negative emotions displayed negative correlations with interaction metrics. For instance, disappointment was negatively correlated with overall satisfaction ($r = -0.57$), and anger with perceived respect ($r = -0.67$). These relationships indicate that when participants felt disrespected or dissatisfied, they were more likely to report negative emotions. Similarly, emotions such as regret and shame were linked to lower levels of perceived understanding and respect, highlighting the detrimental effects of misaligned non-verbal cues on participant perceptions (Table 2 and Fig. 5).

In summary, the findings illustrate that while contextually appropriate laughter significantly enhances positive emotions like compassion, its impact on other emotions is less pronounced. Misplaced laughter primarily influences negative emotions indirectly by reducing perceptions of rapport, warmth, and respect. These results emphasize the critical role of aligning non-verbal cues, such as laughter, with the overall interaction context to enhance user experience and emotional engagement in human-robot interactions.

4.2 Qualitative Analysis

Gaze Patterns. Gaze patterns were central to how participants perceived Furhat's human-like qualities. In the CL condition, participants appreciated Furhat's ability to mimic human gaze behaviours, such as looking around while giving instructions. Participant 5 highlighted that Furhat's subtle gestures like blinking and head movements made it appear more human-like. In the WL condition, while there were no direct comments about gaze, participants still noticed its presence. Participant 13 remarked that Furhat's tracking of movements and gaze added engagement, even when the laughter response was misaligned. These findings suggest that gaze behaviours are consistently engaging, regardless of the context of laughter.

Laughter. Responses to laughter varied between the two conditions. In the CL condition, participants generally responded positively to Furhat's humour. For example, Participant 9 praised Furhat's jokes as "on point", while Participant 17 described its humour as "adorable", even though the laughter sound was considered odd. Some participants noted other aspects of the interaction, such as pronunciation or speech delivery, as enhancing the humour. In the WL condition, participants often described the laughter as forced or out of place. Participant 15 mentioned experiencing the "uncanny valley effect" due to abrupt transitions between laughter and other behaviours. Comments in this condition focused less on the laughter itself and more on other interaction features, such as Furhat's gaze or room-scanning movements. Observational notes indicated that laughter was most effective when contextually appropriate and well-timed. For instance, in the CL condition, synchronised laughter between robot and participant created shared humorous moments, enhancing the interaction. Conversely, in the WL condition, Furhat's lack of synchronised laughter led to awkwardness or disrupted engagement.

Non-verbal Cues. The qualitative findings align with quantitative results, where Joy strongly correlated with Human-Likeness ($r = 0.60$). Both gaze and laughter significantly shaped participants' experiences, highlighting the importance of context and timing in delivering non-verbal cues effectively.

Interaction Flow. Participants experienced a range of emotions during their interactions with Furhat, from joy in smooth conversations to frustration when misunderstandings occurred. Feedback highlighted the importance of interaction flow in shaping emotional responses. For instance, Participant 6 (CL) stated, "The conversation triggered emotions most – joy when it was going smoothly; and disappointment when it was not". Negative responses often stemmed from disruptions, such as Furhat repeating questions, as noted by Participant 1 (CL), who described feeling "annoyed" and "worried about their accent" when Furhat misunderstood them. Similarly, Participant 16 (WL) felt "ignored" and "ashamed" when Furhat failed to respond appropriately: "I felt ignored at certain points like talking by myself."

Failures in reciprocal communication, where participants felt their input wasn't acknowledged, often led to negative emotions such as frustration or shame. For example, Participant 14 (WL) expressed annoyance when Furhat ignored their comments, stating, "My emotions were mainly triggered when he openly responded to my comments or ignored them." Observational notes linked these disruptions to technical limitations like speech recognition errors, such as mistaking "flour" for "flower," which caused conversational loops. Quantitative findings supported these insights, showing negative correlations between emotions like Disappointment and both Understanding ($r = -0.58$) and Overall Satisfaction ($r = -0.56$). Similarly, Anger correlated negatively with satisfaction ($r = -0.63$). These results underline the importance of accurate speech recognition and effective verbal cues in maintaining a smooth interaction flow.

Laughter and small talk were recurring themes in feedback. Participant 10 (CL) appreciated the interaction's novelty, saying, "I love new things, and that was entertaining". Participant 15 (WL) enjoyed Furhat's small talk, feeling it created a sense of teamwork. However, challenges arose when humour or small talk seemed forced or intrusive. For example, Participant 11 (WL) expressed frustration when Furhat interrupted during small talk. Some participants wished for deeper conversations, as Participant 12 (WL) reflected: "It would be cool if I could interact more there".

Overall Observations. Experimenter notes revealed participants' engagement extended beyond verbal interactions, often involving physical actions like manipulating tools to simulate cooking. Participants mirrored Furhat's gaze and expressions, enhancing their connection with the robot. For instance, during Furhat's "thinking" states, participants mimicked its up-and-down eye movements. One participant noted, "I liked how he would think or process something like the recipe generation. You can really understand it from his eyes".

Laughter, especially when contextually appropriate, often elicited reciprocal smiles or laughs. However, participants were more aware of moments when Furhat failed to reciprocate their laughter, aligning with the idea that deviations from expected behaviours are more noticeable than natural behaviours. Participants also frequently synchronised their actions with Furhat's, such as double-checking its gaze during the tools state, indicating a high level of mutual awareness.

5 Conclusion and Discussion

This study examined how coordinating gaze and laughter in the social robot Furhat influences user perceptions of empathy, human-likeness, and naturalness. Our results highlighted the importance of context in non-verbal cue placement and user expectations. In this section, we answer the research questions introduced in the Introduction (Sect. 1), address certain limitations that may have influenced participant interactions with Furhat, and propose directions for future research.

RQ1: Does the placement of gaze-aligned laughter improve the user's contextual understanding? While gaze-laughter alignment influenced participants' perceptions of contextual understanding, the effect was not statistically significant. Participants seemed to rely more on verbal content for understanding, given the structured nature of the tasks. Clear verbal cues, such as Furhat joking about its lack of hands, reduced the dependence on non-verbal alignment. Expectations based on the uncanny valley effect may have also played a role, which as described by [19], occurs when highly realistic but imperfectly human-like robots evoke discomfort, often due to mismatched features or inconsistencies in realism [15,30]. Participants in the CL condition expected higher naturalness, potentially leading to lower ratings when these expectations weren't fully met.

RQ2: Is the effectiveness of coordinating laughter with gaze patterns context-dependent? The placement of gaze-laughter cues significantly impacted emotional metrics such as empathy, naturalness, and compassion, particularly in the CL condition. Misaligned cues disrupted the interaction, reducing perceived attentiveness and responsiveness. These findings emphasize the critical role of contextually appropriate non-verbal behaviors in enhancing human-robot interactions.

RQ3: Does this coordination enhance the perceived naturalness, human-likeness, and empathy of the social robot as experienced by the

user? Our results indicate that gaze-laughter coordination improved perceptions of empathy and naturalness but had a limited effect on human-likeness. Human-likeness, as a holistic metric, is likely influenced by a broader range of behaviours, including speech and facial expressions. The task's thematic focus and variability in participant expectations – such as attention shifts to external elements like the recipe monitor – may have diluted the impact of non-verbal cues on this metric.

In conclusion, with this research we explored how coordinating gaze and laughter affects perceptions of empathy, human-likeness, and naturalness in social robots, using Furhat as a platform. While previous studies examined gaze and laughter separately, this research combined these cues in context-sensitive dialogue to enhance task-oriented interactions. Participants experienced either a CL condition, featuring contextually appropriate gaze-laughter alignment, or a WL condition, where the cues were misaligned. Our results showed that correct alignment significantly improved perceptions of empathy, naturalness, and compassion, and participants often mirrored Furhat's gaze and gestures, indicating strong engagement.

Despite limitations, including a small sample size of 19 participants and occasional errors in Furhat's Automatic Speech Recognition (ASR), the results offer valuable insights. Manual response triggering sometimes caused delays that disrupted conversational flow, and multi-modal distractions (e.g., recipe monitors) reduced participants' focus on Furhat's non-verbal cues. Furthermore, this study did not consider other non-verbal signals, such as vocal intonation, which could create mismatches between human-like body language and the robot's synthetic voice.

Future work should focus on real-time, automated laughter detection and minimising distractions to better evaluate gaze-laughter effects. Exploring varied laughter types and testing these cues across different contexts could further refine their impact. By improving non-verbal coordination, social robots can foster more natural, empathetic, and engaging interactions.

6 Ethical Considerations

This study prioritised participant privacy and safety by adhering to strict ethical and legal standards. Key measures included obtaining ethical approval from the Swedish Ethical Review Authority and ensuring compliance with the General Data Protection Regulation (GDPR) 2016/679. Participants were informed of the recording and their rights at all stages – through recruitment posters, reminders during scheduling, and signed consent forms upon arrival. These safeguards emphasised their right to withdraw or request modifications to their data at any time. To maintain anonymity, participants used pseudonyms, and GPT-4 was programmed to generate fake nicknames for the names provided. Raw video footage was accessible only to the experimenter, supervisors, and research members, while questionnaires were anonymised with unique codes linked to experimental conditions rather than personal identifiers. For publication, all identifying details were excluded, and visuals were blurred to protect participant identities. These protocols ensured responsible data handling while safeguarding participant confidentiality.

References

1. Admoni, H., Scassellati, B.: Social eye gaze in human-robot interaction: a review. J. Hum.-Robot Interact. **6**(1), 25-63 (2017). https://doi.org/10.5898/JHRI.6.1.Admoni
2. Bachorowski, J.A., Smoski, M.J., Owren, M.J.: The acoustic features of human laughter. J. Acoust. Soc. Am. **110**(3 Pt 1), 1581–1597 (2001). https://doi.org/10.1121/1.1391244
3. Becker-Asano, C., Ishiguro, H.: Laughter in social robotics-no laughing matter. In: International Workshop on Social Intelligence Design, pp. 287–300 (2009)
4. Clift, R.: Identifying action: laughter in non-humorous reported speech. J. Pragmatics **44**(10), 1303–1312 (2012). ISSN 0378-2166, https://doi.org/10.1016/j.pragma.2012.06.005
5. Cosentino, S., et al.: Human-humanoid robot social interaction: laughter. In: 2013 IEEE International Conference on Robotics and Biomimetics (ROBIO), pp. 1396–1401 (2013). https://doi.org/10.1109/ROBIO.2013.6739661
6. Furhat Robotics: Furhat launcher documentation (2025). https://docs.furhat.io/launcher/. Accessed 31 Jan 2025
7. Giannitzi, E.: When eyes meet laughter: exploring non-verbal cues in human-robot interaction with Furhat. Master's thesis, University of Gothenburg (2024). https://hdl.handle.net/2077/83857
8. Giannitzi, E., Maraev, V., Howes, C.: Reacting to the last laugh: probing non-humorous laughter with Furhat robot. In: Proceedings of the Laughter and Other Non-Verbal Vocalisations Workshop 2024, pp. 45–47 (2024). https://doi.org/10.21437/LW.2024-14
9. Gullberg, M., Holmqvist, K.: Keeping an eye on gestures: visual perception of gestures in face-to-face communication. Pragmat. Cogn. **7**, 35–63 (1999). https://doi.org/10.1075/pc.7.1.04gul
10. Harel, D.: Statecharts: a visual formalism for complex systems. Sci. Comput. Program. **8**(3), 231–274 (1987). https://doi.org/10.1016/0167-6423(87)90035-9
11. Henschel, A., Laban, G., Cross, E.S.: What makes a robot social? a review of social robots from science fiction to a home or hospital near you. Curr. Robot. Rep. **2**(1), 9–19 (2021). https://doi.org/10.1007/s43154-020-00035-0
12. Inoue, K., Lala, D., Kawahara, T.: Can a robot laugh with you?: shared laughter generation for empathetic spoken dialogue. Front. Robot. AI **9** (2022). ISSN 2296-9144, https://doi.org/10.3389/frobt.2022.933261
13. Krämer, N., Manzeschke, A.: Social Reactions to Socially Interactive Agents and Their Ethical Implications, 1 edn., pp. 77–104. Association for Computing Machinery, New York, NY, USA (2021). ISBN 9781450387200, https://doi.org/10.1145/3477322.3477326
14. Lugrin, B.: Introduction to socially interactive agents. In: Lugrin, B., Pelachaud, C., Traum, D. (eds.) Handbook on Socially Interactive Agents – 20 Years of Research on Embodied Conversational Agents, Intelligent Virtual Agents, and Social Robotics, Volume 1: Methods, Behavior, Cognition, pp. 1–18. ACM Books (2021)

15. MacDorman, K.F., Chattopadhyay, D.: Reducing consistency in human realism increases the uncanny valley effect; increasing category uncertainty does not. Cognition **146**, 190–205 (2016). ISSN 0010-0277, https://doi.org/10.1016/j.cognition.2015.09.019
16. Maraev, V., Bernardy, J.P., Howes, C.: Non-humorous use of laughter in spoken dialogue systems. In: Linguistic and Cognitive Approaches to Dialog Agents (LaCATODA 2021), pp. 33–44 (2021). http://ceur-ws.org/Vol-2935/paper4.pdf
17. Mazzocconi, C., Maraev, V., Somashekarappa, V., Howes, C.: Looking for laughs: gaze interaction with laughter pragmatics and coordination. In: ICMI '21, Proceedings of the 2021 International Conference on Multimodal Interaction, pp. 636–644. ACM, New York, NY, USA (2021). ISBN 9781450384810, https://doi.org/10.1145/3462244.3479947
18. Mazzocconi, C., Tian, Y., Ginzburg, J.: What's your laughter doing there? a taxonomy of the pragmatic functions of laughter. IEEE Trans. Affect. Comput. **13**(3), 1302–1321 (2020). https://doi.org/10.1109/TAFFC.2020.2994533
19. Mori, M., MacDorman, K., Kageki, N.: The uncanny valley [from the field]. IEEE Robot. Autom. Mag. **19**(2), 98–100 (2012). https://doi.org/10.1109/MRA.2012.2192811
20. Moubayed, S.A., Beskow, J., Skantze, G., Granström, B.: Furhat: a back-projected human-like robot head for multiparty human-machine interaction. In: Cognitive Behavioural Systems: COST 2102 International Training School, Dresden, Germany, February 21–26, 2011, Revised Selected Papers, pp. 114–130. Springer (2012)
21. Nijholt, A.: Humor and Embodied Conversational Agents. No. 2003-03 in CTIT Technical reports series, Centre for Telematics and Information Technology (CTIT), Netherlands (2003), imported from CTIT
22. OpenAI: Overview (2024). https://platform.openai.com/docs/overview. open AI Platform. Accessed December 2023–August 2024
23. OpenAI: Prompt engineering (2024). https://platform.openai.com/docs/guides/prompt-engineering. open AI Platform. Accessed December 2023–August 2024
24. Parreira, M.T., Gillet, S., Vázquez, M., Leite, I.: Design implications for effective robot gaze behaviors in multiparty interactions. In: 2022 17th ACM/IEEE International Conference on Human-Robot Interaction (HRI), pp. 976–980 (2022). https://doi.org/10.1109/HRI53351.2022.9889481
25. von der Pütten, A.M., Krämer, N.C., Gratch, J., Kang, S.H.: "It doesn't matter what you are!" explaining social effects of agents and avatars. Comput. Hum. Behav. **26**(6), 1641–1650 (2010). ISSN 0747-5632, https://doi.org/10.1016/j.chb.2010.06.012
26. Richardson, D.C., Dale, R., Tomlinson, J.M.: Conversation, gaze coordination, and beliefs about visual context. Cogn. Sci. **33**(8), 1468–1482 (2009). https://doi.org/10.1111/j.1551-6709.2009.01057.x
27. Rossano, F.: Gaze in social interaction. Handbook of Conversation Analysis, pp. 308–329 (2013)
28. Ruhland, K., et al.: A review of eye gaze in virtual agents, social robotics and HCI: behaviour generation, user interaction and perception. Comput. Graph. Forum **34**(6), 299-326 (2015). ISSN 0167-7055, https://doi.org/10.1111/cgf.12603
29. Scott, S., Lavan, N., Chen, S., Mcgettigan, C.: The social life of laughter. Trends Cogn. Sci. **18** (2014). https://doi.org/10.1016/j.tics.2014.09.002
30. Seyama, J., Nagayama, R.S.: The uncanny valley: effect of realism on the impression of artificial human faces. Presence **16**(4), 337–351 (2007). https://doi.org/10.1162/pres.16.4.337

31. Somashekarappa, V.: Look on my thesis, ye mighty: Gaze Interaction and Social Robotics (2024)
32. Somashekarappa, V., Sayeed, A., Howes, C.: Neural network implementation of gaze-target prediction for human-robot interaction. In: Proceedings of the 32nd IEEE International Conference on Robot and Human Interactive Communication (2023). https://www.christinehowes.com/papers/Somashekarappa.etal_ROMAN_2023.pdf
33. Tiselius, E., Sneed, K.: Gaze and eye movement in dialogue interpreting: an eye-tracking study. Bilingualism Lang. Cogn. **23**(4), 780–787 (2020). https://doi.org/10.1017/S1366728920000309
34. Truong, K.P., Trouvain, J.: Laughter annotations in conversational speech corpora-possibilities and limitations for phonetic analysis. In: 4th International Workshop on Corpora for Research on Emotion Sentiment and Social Signals (ES3 2012), pp. 20–24, European Language Resources Association (ELRA) (2012)
35. Türker, B.B., Buçinca, Z., Erzin, E., Yemez, Y., Sezgin, M.: Analysis of engagement and user experience with a laughter responsive social robot. In: Interspeech 2017, pp. 844–848 (2017). ISSN 2958-1796, https://doi.org/10.21437/Interspeech.2017-1395
36. Wang, N., Gratch, J.: Can virtual human build rapport and promote learning? In: Proceedings of the 14th International Conference on Artificial Intelligence in Education (AIED), pp. 737–739 (2009). https://doi.org/10.3233/978-1-60750-028-5-737
37. Zawieska, K.: The iron cage of social robotics. J. Hum.-Robot Interact. **14**(1) (2024). https://doi.org/10.1145/3695772