Tartu University

Faculty of Science and Technology

Institute of Technology

Miriam Calafa’

Designing Multimodal Emotional Expression for a Robotic Study Companion

Master’s thesis (30 EAP)
Robotics and Computer Engineering

Supervisors:

Farnaz Baksh, MSc
Karl Kruusamäe, PhD

Tartu 2025

Abstract

Designing Multimodal Emotional Expression for a Robotic Study Companion

This thesis investigates the design, implementation, and evaluation of the emotional expression of the Robot Study Companion (RSC) intended to support university students' motivation, well-being, and learning outcomes.

Guided by a theoretical review of affective feedback in educational robotics, we identified six target emotions: Anger, Joy, Pride, Fun, Surprise, and Caring, each linked to educationally meaningful functions. We developed a fully functional digital twin of the RSC in ROS 2 to enable rapid prototyping and simulation, and built a modular emotional expression that coordinates color, motion, facial expression, and vocal behavior for each emotional response.

To assess recognition accuracy and user experience, we conducted a video-based cross-cultural study with 47 participants. Participants distinguished positive versus negative valence with over 95% accuracy, and achieved an overall emotion-recognition rate of 58.16%. User Experience Questionnaire (UEQ) results were consistently positive, with Perspicuity and Attractiveness rated highest. Qualitative interviews highlighted the strengths of our multimodal design as well as opportunities for refinement. Cross-group analyses revealed systematic differences by gender, academic level, and nationality, underscoring the need for adaptive, personalized emotion profiles.

All software developed in this work is released open-source, providing a robust foundation for future research in affective human–robot interaction within educational settings.

CERCS: T120 Systems engineering, computer technology; T125 Automation, robotics, control engineering; S281 Computer-assisted education.
Keywords: social robotics, educational robotics, ROS2

Resümee

Robotõpiabilise multimodaalse emotsiooniväljenduse disainimine

Käesolev lõputöö uurib Robot Study Companion (RSC) emotsionaalse väljenduse disaini, teostust ja hindamist eesmärgiga toetada üliõpilaste motivatsiooni, heaolu ja õpitulemusi.

Tuginedes haridusrobootika afektiivse tagasiside teoreetilisele ülevaatele, tuvastasime kuus sihitud emotsiooni: viha, rõõm, uhkus, lust, üllatus ja hoolivus – igaüks seostatuna hariduslikult tähenduslike funktsioonidega. RSC-st arendati ROS 2 platvormil täielikult funktsionaalne digitaalne kaksik, mis võimaldab kiiret prototüüpimist ja simulatsiooni. Töötasime välja modulaarse emotsionaalse väljenduse süsteemi, mis koordineerib värvi, liikumise, näoilmete ja häälekäitumise sünergiat iga emotsiooni puhul.

Tuvastustäpsuse ja kasutajakogemuse hindamiseks viisime läbi videopõhise kultuurideülese uuringu 47 osalejaga. Osalejad eristasid positiivse ja negatiivse väärtusega emotsioone enam kui 95% täpsusega ning kogutuvastuse määr oli 58.16%. User Experience Questionnaire (UEQ) tulemused olid järjepidevalt positiivsed, kusjuures kõrgeimad hinnangud anti mõistetavusele ja atraktiivsusele. Kvalitatiivsed intervjuud tõid esile meie multimodaalse disaini tugevused ning parendusvõimalused. Rühmadevahelised analüüsid näitasid süsteemseid erinevusi soo, akadeemilise taseme ja rahvuse lõikes, rõhutades vajadust kohanduvate ja isikupärastatud emotsiooniprofiilide järele.

Kõik käesoleva töö raames arendatud tarkvarakomponendid on avaldatud avatud lähtekoodiga, pakkudes tugevat alust edasiseks uurimistööks afektiivses inimese ja roboti interaktsioonis hariduskeskkondades.

CERCS: T120 Süsteemitehnoloogia, arvutitehnoloogia; T125 Automatiseerimine, robootika, control engineering; S281 Arvuti õpiprogrammide kasutamise metoodika ja pedagoogika.
Märksõna: sotsiaalrobotika, haridusrobotika, ROS2

Contents

List of Figures

List of Tables

1 Introduction

University students often face challenges such as maintaining motivation [1], managing their time effectively [2], and feeling isolated during independent study sessions [3]. These challenges can negatively impact academic performance, often leading to stress and burnout [4]. In this context, socially interactive robots have emerged as potential learning companions, offering personalized academic support, fostering emotional engagement, and creating dynamic, interactive learning experiences [5].

By leveraging artificial intelligence, social robots can adapt their responses and teaching styles to individual student needs [6], making educational experiences more immersive and enjoyable [7]. Social robots have the potential to enhance both academic performance and overall student well-being. However, despite their potential, robotic study companions still face challenges in effectively supporting students at higher education levels.

Many existing robotic systems struggle with limitations in facial expressions, emotions, and tailored feedback, which can hinder meaningful communication with students [8]. The ability of these robots to express emotions effectively remains a critical factor in their acceptance and impact. Current research in Human-Robot Interaction (HRI) has largely focused on improving these aspects for younger students in primary and secondary education. However, university students have distinct academic and emotional needs that differ from those of younger learners [9].

Numerous studies have explored ways to enhance human-robot interaction; however, significant gaps in the literature remain, particularly in achieving a truly immersive and natural lived experience [10]. There is a need to investigate how robotic study companions can be designed to support university students effectively, considering their specific academic challenges, cognitive workload, and emotional well-being.

1.1 Research Questions

To address these gaps, this research explores how robotic study companions can be designed to support university students answering the following research questions:

RQ1. What emotions should a robotic study companion express to enhance university students' study processes, motivation, and learning outcomes? This question explores the specific emotional states that are most beneficial for supporting students academically and emotionally, helping to create a more engaging and effective learning experience.

RQ2. How can these emotions be expressed most clearly and effectively through the multi-modal capabilities of the RSC? This question focuses on identifying the most effective combination of facial expressions, gestures, vocal cues, and colors to ensure that the robot's emotional communication is intuitive and easily understood by university students.

RQ3. To what extent do the RSC's emotional expressions accurately convey the intended emotions? This question examines whether the implemented emotional expressions align with human expectations, ensuring that the robot's intended emotional states are perceived as designed.

RQ4. Do university students appreciate the emotional support provided by the RSC? This question assesses whether students find value in the emotional expressivity of the robot and whether it positively impacts their motivation, stress levels, and overall study experience.

1.2 Contribution

The contributions of this work can be summarized as follows:

1. **Literature Review on Human-Robot Interaction (HRI):** This thesis presents an extensive literature review focusing on the educational-affective needs of university students and the role of human-robot interaction in addressing these needs (RQ1). This review provides a foundation for the behavioural design of robotic study companions that support students' academic and emotional well-being.

2. **Development of the Robot Study Companion (RSC) Digital Twin:** A digital twin of the Robot Study Companion (RSC) prototype was developed to test and refine the robot's emotional expressivity (RQ2). The digital twin has the potential to function independently from the physical robot, serving as a standalone tool for research or educational purposes when a physical robot is not available.

3. **Development of Multimodal Emotional Expressions for University Students:** Based on literature and the specific needs of university students, a set of multimodal emotional expressions (incorporating facial expressions, gestures, vocal cues, and colors) was developed for the RSC (RQ2). This development was carried out using the RSC digital twin.

4. **Cross-Cultural Study and Design Guidelines:** A cross-cultural study with university students evaluated the effectiveness of the RSC's emotional expressivity enhancements (RQ3). This study assessed how students from diverse cultural backgrounds perceived and interacted with the robot (RQ4). Based on the findings, evidence-based design guidelines were established to improve the RSC's emotional communication in future development.

2 State of the Art

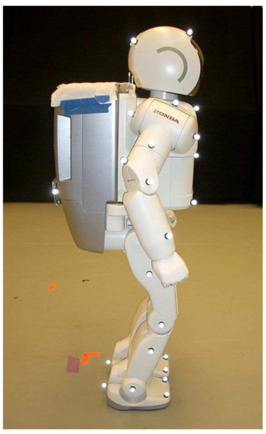
This section reviews research on social robots and robotic study companions. It begins by defining social robots, their key features, and notable examples, then focuses on robotic companions, highlighting their characteristics and specific cases. Next, the section examines the Robot Study Companion (RSC), discussing its design, functionality, applications, and how it addresses learning challenges, along with potential future developments. Finally, the review of Human-Robot Interaction (HRI) studies explores how features such as eye expressions, color variations, and multimodal elements enhance emotional communication.

2.1 Robots, Social Robots and Robot Companions

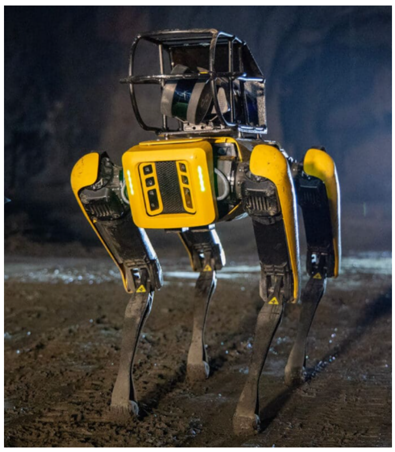
2.1.1 Robots

A robot is a machine, typically programmable by a computer, that can autonomously execute a series of complex actions [11]. It may operate under the guidance of an external control system or have an integrated control mechanism. While some robots are designed to resemble humans, most are built primarily for task execution, prioritizing functionality over expressive aesthetics (Fig 2.1).

(a) Roomba (iRobot) [12]



(b) ASIMO (Honda) [13]



(c) Spot (Boston Dynamics) [14]

Figure 2.1: Three examples of robot

Some prominent examples of robots with defined characteristics are:

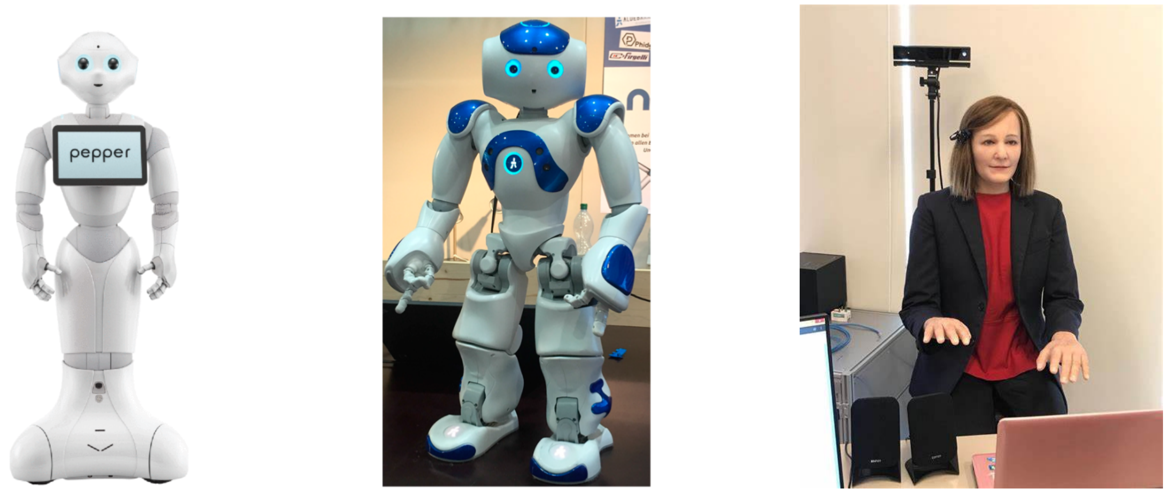
- **Roomba (iRobot):** a robot vacuum cleaner that focuses on its primary function: cleaning. The design prioritizes efficiency and functionality [12], with minimal aesthetic features beyond its basic round shape and buttons (Fig 2.1a).

- **ASIMO (Honda):** a robot that is capable of some human-like functional capabilities such as walking, running, and interacting with its environment [13]. Its expressive aesthetics are secondary to its ability to perform complex physical tasks (Fig 2.1b).
- **Spot (Boston Dynamics):** a robot designed for various industrial tasks, such as inspection, surveying, and data collection [14]. Its design is sleek and functional, but it does not prioritize emotional or aesthetic appeal (Fig 2.1c).

2.1.2 Social Robots

Social robots are robots designed specifically for interacting with humans in a social context, representing a significant advancement in robotics. Unlike traditional robots that perform mechanical tasks, social robots are designed to interact with humans on a social level through body language, facial expressions, and emotions [15] [16]. Social robots find applications across various fields, including education, where they assist with tutoring and teaching support; customer service, where they offer information, guidance, and handle routine tasks [17]; and domestic assistance, where they help with household chores [18].

Social robots often have human-like or animal-like appearances (Fig. 2.2 and Fig. 2.3) to make interactions more relatable and trustworthy [19]. Their applications span various fields, including healthcare, education, and domestic services, offering new opportunities for enhancing human well-being and social interaction [20].



(a) Pepper [21] (b) NAO [22] (c) Nadine

Figure 2.2: Three examples of social robot

Some examples of social robots are:

- **Pepper Robot** is a humanoid robot built to interact with people and assist in various settings (Fig. 2.2a). Pepper has facial recognition capabilities and can identify and interact with objects in its surroundings. It can understand and process spoken language, making it well-suited for verbal communication [23]. It responds to voice and gesture commands, allowing intuitive and natural communication. Pepper is designed to engage users emotionally and socially, making it ideal for roles that demand human-like interaction [21].
- **NAO** is built for effective human-robot interaction, making it well-suited for education, healthcare, and entertainment. With a 24-degree-of-freedom articulated structure (Fig.

2.2b), NAO can perform various movements. It can grasp objects, detect and recognize people, and localize sounds. NAO also supports speech recognition and synthesis, enabling it to engage in basic conversations [24]. With various sensors, including cameras, microphones, sonar, and tactile sensors, it can recognize facial expressions and respond appropriately, enhancing its ability to interact naturally with people[25].

- **Nadine** is a sophisticated social humanoid robot (Fig. 2.2c) designed to interact with humans naturally and believably. Nadine is equipped with advanced speech recognition and synthesis capabilities, enabling her to effectively understand and respond to spoken interactions. She can also use facial expressions and body language to communicate non-verbally [26].

2.1.3 Companion Robots

A companion robot is a socially interactive, autonomous machine designed to assist and engage with humans in various roles, enhancing their quality of life through practical support and social interaction. The design and functionality of these robots are tailored to meet the specific needs of their users, with a strong emphasis on empathy, mobility, and user-centered design [27] [28].

Key applications include healthcare and elderly care, where robots assist with daily activities, provide companionship, and offer emotional support in homes, workplaces, and broader society [28]. They are also used for cognitive and emotional rehabilitation [29]. Social robots can offer companionship, reduce feelings of loneliness, and improve psychological well-being by promoting cognitive stimulation activities [30]. Social robots have been increasingly integrated into educational environments to support students, particularly those facing stressful situations. They can provide personalized and emotional feedback, enhancing student engagement and creating a more stimulating learning environment [31].

Robot companions often incorporate anthropomorphic and zoomorphic designs, featuring human-like or animal-like characteristics to enhance relatability and interaction engagement. They also exhibit emotional intelligence, allowing them to recognize and respond to human emotions, facilitating more natural and meaningful communication. Additionally, these robots are equipped with speech and gesture recognition technologies, enabling them to interpret and respond to verbal and non-verbal cues, which improves their ability to interact intuitively and effectively[32].

(a) Paro

(b) LOVOT [33]

(c) Jibo

Figure 2.3: Three examples of robot companions

Some examples of companion robots are:

- **PARO** is designed to enhance mood, reduce agitation, and promote relaxation, especially among elderly patients with dementia or depression. Its soothing presence has been shown to calm even those who are otherwise unresponsive, making it a valuable tool for managing difficult behaviors in therapeutic settings [34]. Modeled after a baby seal, Paro's soft, lifelike appearance makes it appealing and comforting to users (Fig. 2.3a). Paro is designed with simplicity in mind and is user-friendly and accessible, even for elderly individuals and those with cognitive impairments [35].

- **LOVOT** is a next-generation companion robot designed to interact with humans on an emotional and social level. LOVOT is designed to live at home with users, similar to a companion animal, providing continuous natural motion for user comfort. The robot's movements are crafted using professional animator techniques based on traditional animation principles, making its behavior appear more lifelike and engaging (Fig. 2.3b). LOVOT is intended for long-term companionship, providing continuous interaction and comfort to users [33].

- **Jibo** is a domestic robot designed to interact with users naturally and engagingly. It is designed to facilitate social interaction and companionship. It can recognize and respond to users' faces, voices, and emotions, making it a highly interactive and engaging presence in the home. The robot can communicate using verbal and non-verbal cues: it can speak, display emotions through its screen, and use body movements to convey messages (Fig. 2.3c). Its ability to express emotions and respond to human interactions helps build a trustable and engaging relationship with its users [22].

2.2 RSC

The Robot Study Companion (RSC) [9] is an open-source tabletop robot specifically designed to enhance the learning experience of university students (Figure 2.4). It supports multiple interaction modalities, including voice, feedback, and touch, to create engaging and interactive study sessions. Striking a balance between affordability and technical performance, the RSC is accessible to both educational institutions and students.

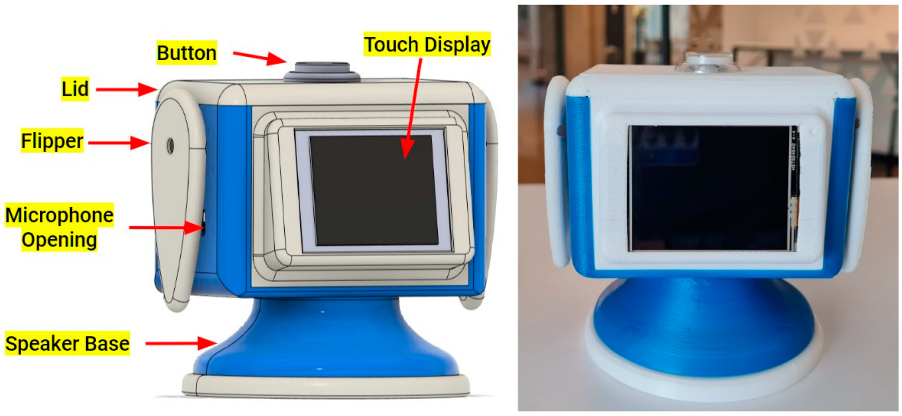


Figure 2.4: RSC design and functionalities

The initial version is built using off-the-shelf components and features a modular design that allows for reconfigurability. Users can customize the system by adding or replacing sensors, computational units, and other components as needed. With its sleek and approachable design, the RSC fosters a welcoming presence, positioning itself as a supportive study companion that provides academic assistance and enhances the overall learning experience.

The RSC incorporates a diverse set of interaction methods, allowing users to engage in study sessions in a way that suits their preferences. These modalities are designed to provide flexibility, meaning not all of them need to be active simultaneously.

- Voice-Based Interaction. The RSC is equipped with a speaker and microphones, enabling it to recognize and respond to voice commands. Users can engage in conversations with the robot, with their speech captured through Speech-to-Text (STT) software. The processed input is then used to generate appropriate responses, which are converted back into speech using Text-to-Speech (TTS) software and played through the speaker.

- Gesture-Based Communication. The RSC utilizes gestures to communicate visually, making interactions more natural and engaging. Flipper-like structures are attached to servo motors, allowing the robot to express itself through movement. By incorporating movement, the RSC enhances the perception of responsiveness and social presence.

- Tactile Interaction. The RSC includes a physical button that allows users to interact through touch-based input. For example, the arcade button can signal that the robot is ready to listen, providing users with an alternative to voice commands.

- Display Screen. The RSC incorporates a touchscreen display, allowing users access to various settings, including volume adjustments and other interactive options. Beyond basic controls, the display can also serve as an informational interface, visually presenting concepts or displaying relevant study materials to enhance learning experiences.

- RGB Light Ring. The RSC features a light ring placed around the button, which can display different colors with varying intensities. This light ring serves as a visual indicator, offering users clear feedback on the robot's status, such as readiness or error states, enhancing the interaction with the robot.

The RSC addresses key learning challenges faced by university students by providing an interactive, multimodal support system tailored to their academic needs. Designed as an open-source platform, the RSC enhances engagement, motivation, and personalized learning experiences through natural language processing and multimodal human-robot interaction. It facilitates interactive study sessions by answering questions, providing feedback, and offering study recommendations, helping students grasp complex concepts with ease. Additionally, its affordability and modular design allow for customization, enabling students to adapt the RSC to their specific educational needs. By serving as a consistent and interactive study companion, the RSC helps bridge gaps in self-directed learning, fosters better time management, and reduces the sense of isolation that can arise in higher education environments. Through continuous improvements, the RSC aims to become a reliable and accessible tool for university students seeking enhanced academic support.

In the future work of the RSC prototype, plans were highlighted to incorporate affective emotion recognition capabilities into the RSC, allowing it to detect and respond to the user's emotional state. The goal is for the RSC to adapt its motions and sounds accordingly, fostering positive emotional engagement. This thesis focuses on implementing the system's emotional responses to enhance user interaction and support.

2.3 The Impact of Emotional Expression of Tutors

Emotions are complex psychological and physiological states that arise in response to various stimuli, influencing human thoughts, behaviors, and interactions [36]. They serve as an essential mechanism for communication, decision-making, and survival. Internal or external events trigger these affective responses and involve a combination of cognitive appraisal, physiological arousal, and expressive behaviors [37]. Emotions impact mental well-being and play a crucial role in learning, motivation, and social bonding, highlighting their significance in human life [38].

Emotions influence fundamental cognitive processes like attention, perception, and memory, as well as higher-level processes such as learning, decision-making, and planning [39]. Emotions are linked to various physical health outcomes. Negative emotions can lead to chronic stress, weaken the immune system, and contribute to conditions like depression and anxiety [40]. Emotional well-being is closely tied to quality of life. Positive emotions enhance cognitive processing and morale, while negative emotions can diminish one's ability to function and enjoy life [41]. Emotions are a primary force that activates or energizes our behaviors, influencing motivation and social processes [42].

Humans express emotions through a variety of modalities, each contributing to effective communication and social interaction. Emotions in verbal communication are often conveyed through the words we choose and the tone of our speech [43]. Facial expressions are a primary means of conveying emotions. Specific facial muscle movements correspond to different emotions [44]. Gestures, posture, and overall body movements also play a significant role in expressing emotions [45]. Some emotions can be ambiguous and complex to interpret solely based on one modality. Effective emotional expression often involves a combination of verbal and nonverbal cues [46]. Cultural norms influence how emotions are expressed and interpreted. What is considered an appropriate expression of emotion in one culture might be seen differently in another [47].

2.3.1 Tutors Emotions

To understand how emotions influence learning, it is essential to consider their multifaceted impact on various aspects of the learning process. Emotions can significantly affect attention, memory, motivation, and overall academic performance. Emotions significantly shape students' learning strategies by influencing motivation, engagement, self-regulation, and overall academic performance. Learning is not just a cognitive process but also involves emotional and behavioral components that determine how effectively students absorb and apply knowledge[48].

The emotions tutors express can significantly influence students' learning performance, shaping their motivation, engagement, and cognitive processing. Traditionally, positive emotions such as happiness have been viewed as beneficial for learning, as they create a supportive and encouraging environment. When instructors express happiness, they may foster a sense of warmth and competence, which can enhance students' confidence and willingness to participate in learning activities [49]. Studies show that students who experience positive emotions while learning are more likely to retain information, develop problem-solving skills, and sustain their attention on academic tasks [50].

Negative emotions, on the other hand, can have a complex impact on learning. Research suggests that emotions such as anger, when appropriately used, can also play a role in improving student performance. Anger signals to students that their performance is inadequate and that improvement is necessary. This can motivate students to increase their effort, pay closer attention to details, and engage more actively in the learning process[51].

2.3.2 Effect of each emotion on students

Both positive and negative emotions can contribute to enhanced learning outcomes. This section provides a detailed analysis of how each tutor's emotion influences student engagement and performance in educational settings.

Joy

Tutors who provide emotional support and create a positive atmosphere can improve students' affective states, leading to better engagement and learning outcomes. This support is essential in both traditional and online learning settings [52]. Joyful learning environments significantly enhance students' motivation and engagement. When students experience joy, they are more likely to be motivated and engaged in their learning activities, which can lead to better learning outcomes [53].

Caring

Caring tutors can help mitigate students' anxiety and stress by providing emotional support and creating a safe learning environment. This support is crucial for students' mental well-being and can positively impact their academic performance [54]. A good relationship between students and tutors can significantly enhance motivation and educational engagement. This relationship fosters a supportive environment where students feel valued and understood, encouraging them to participate more actively in their learning process. Studies have shown a positive correlation between the quality of the student-teacher relationship and students' academic performance, such as higher GPA scores. This suggests caring tutors can directly influence students' academic achievements [55].

Pride

Tutors who express pride in their students can foster a positive emotional environment. This emotional support is crucial as it enhances student motivation, engagement, and overall well-being. Positive emotions from tutors can moderate the relationship between knowledge and cognitive skills, making learning more effective[56]. Engaged students are more likely to participate actively in their learning process, which can lead to better academic outcomes [57]. Tutor's pride can boost students' self-esteem, which is a critical factor in their academic success. Higher self-esteem can lead to better future expectations and a stronger belief in their ability to achieve their goals[57].

Fun

Fun activities can help improve students' affective states by reducing negative emotions like boredom and increasing positive emotions like excitement and interest. This positive emotional state can lead to better engagement and learning behavior [58]. Incorporating fun into tutoring, such as gamification, interactive features, and playful interactions, can significantly enhance student engagement and motivation while reducing stress, which can help maintain students' attention and interest in the subject matter [59].

Surprise

Tutors expressing astonishment at positive results can boost students' confidence and reinforce their belief in their abilities. This positive reinforcement can lead to increased motivation and effort in future learning tasks [60]. Students are more likely to engage with feedback when they see their tutor's genuine reactions, including astonishment. This engagement is crucial for learning, as it helps students understand and act on the feedback provided [61].

Anger

Research indicates that instructors' expressions of anger can improve learning performance. This suggests that anger can heighten students' attention and focus, potentially leading to better retention and recall of information [51] [62]. Anger can act as a motivational tool, pushing students to perform better. The emotional intensity associated with anger might create a sense of urgency, importance, seriousness of the task, or the need for improvement, prompting students to engage more deeply with the material, thereby enhancing their performance [63] [51].

2.3.3 Effective Tutors' Emotion for Common Students' Scenarios

Different emotions expressed by tutors can help create a supportive environment that fosters engagement, motivation, and resilience. Each emotional expression by a tutor addresses students' specific needs at various points of their learning process. Table 2.1 summarizes the tutor emotions presented in Section 2.3.2 and the types of student scenarios in which these emotions are most effective.

Table 2.1: Common Student Scenarios and the Most Effective Tutor Emotion

Emotion	**Common Student Scenario**
Joy	When students achieve a milestone, such as completing a task or mastering a concept reinforcing their intrinsic motivation [64].
Caring	When students struggle, feel frustrated, stressed, express confusion or anxiety, making them feel valued and understood [54].
Pride	When students show noticeable improvement or accomplish a challenging goal, especially after significant effort, reinforcing their sense of achievement [64].
Anger	When there is a need to emphasize the seriousness of repeated mistakes, incoming deadlines, neglected tasks or excessive procrastination, prompting students to engage more deeply with the material, thereby enhancing their performance [51].
Fun	When there is a need to break the monotony during long study sessions or repetitive exercises, when contrasting negative emotions like boredom is needed, leading to better engagement and learning behavior [65].
Surprise	When students perform better than expected or find creative solutions, reinforcing students' sense of accomplishment and creativity [60].

2.4 Methods for Emotional Expression in Robots

Human emotions are inherently multimodal, involving facial expressions, speech, gestures, and other non-verbal cues. Multimodal robots are superior in expressing emotions because they integrate various communication channels, leading to more accurate, natural, and engaging interactions [66]. Robots that utilize multiple modalities can better recognize and express these complex emotional states, leading to more accurate and human-like interactions [67]. Robots can create more natural and intuitive interactions by combining multiple channels of communication, such as facial expressions, body movements, and vocal tones. This multimodal approach helps make the robot's emotional states more legible and relatable to humans [68]. Moreover, users perceive robots that can express emotions multimodally more positively. This positive perception enhances user engagement and satisfaction, making interactions more enjoyable and effective [69].

2.4.1 Colors

Color is a powerful modality for conveying emotions. In robotics, integrating color with other expressive channels, such as facial expressions, can significantly enhance emotion recognition and improve the quality of human-robot interaction [70]. Research suggests that the use of facial colors, aligned with principles from color psychology, can increase emotion recognition accuracy by up to 25% [71].

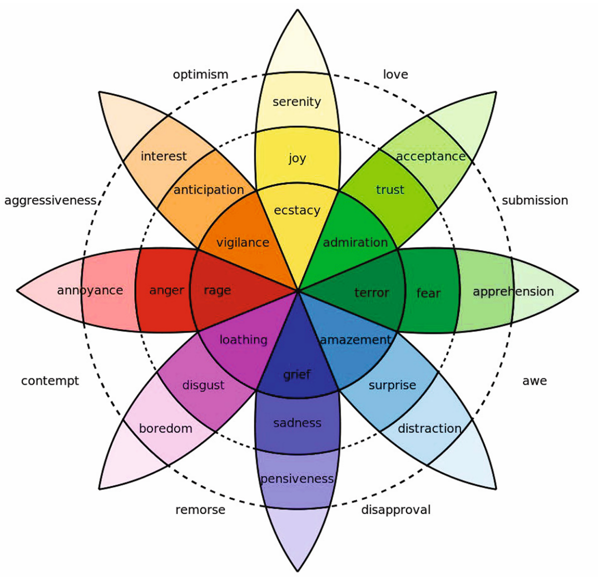


Figure 2.5: Plutchin's Wheel of Emotions [72]

Plutchik's Wheel of Emotions (Fig. 2.5) integrates color theory into an emotional framework, illustrating how different hues can represent varying emotional intensities. For example, the outer layers of the wheel represent milder emotions, typically expressed through softer colors, while more intense emotions like anger or joy are represented by more saturated hues [73]. Frameworks like Plutchik's Wheel of Emotions reinforce the relevance of color in emotional design. In the context of human-robot interaction, color not only enhances perceived emotional intelligence but also contributes to the lifelikeness and social presence of artificial agents. Color theories examine the psychological impact of colors on human emotions, behavior, and cognition, with certain colors being traditionally linked to specific emotional responses [74]. This interplay between color and psychology emphasizes the importance of thoughtful color selection in design, not just for aesthetic purposes, but as a powerful means of emotional communication [74].

Color alone does not determine emotional states; its influence depends on context and placement. For example, red can convey different emotions depending on where it appears. A red face is often associated with anger[75], while flushed cheeks may indicate health[76]. Red around the eyes can evoke a sense of fear or distress[77]. Additionally, background color plays a role in modulating the perception of facial expressions, though facial coloration itself has a more pronounced impact[78].

The emotional impact of colors can vary based on cultural and contextual factors. For instance, the color red can have different meanings in different contexts, such as romance or achievement[79]. Additionally, studies have shown that color preferences and emotional responses can differ across regions and cultures. Given that color perception is shaped by cultural and contextual factors, designers must carefully consider these variables when incorporating color-based emotional cues into robotic systems [80].

2.4.2 Facial Expression

Eyes are crucial in human communication, serving both to see and to show gaze, which is vital for non-verbal communication. This centrality makes the design of robot eyes particularly important in HRI [81]. However, conveying emotion through eyes is complex: while eye colors can be used to express emotions, they have been found to be unreliable on their own. This indicates the need for a holistic approach that considers multiple aspects of eye design and movement [45]. In digital text, emoticons rely primarily on distinct "eye" features, yet they convey emotion as effectively as empathetic verbal expressions, mirroring facial cues in face-to-face interaction [82, 83]. Likewise, HRI researchers have long drawn inspiration from cartoons and animation to enhance robot expressivity [84]. By applying the animation principle of exaggeration, adjusting eye shapes and movements, robots can amplify emotional signals that might otherwise be lost in a mechanical face [85]

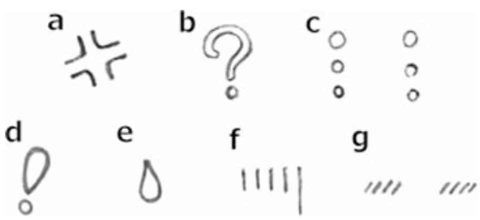


Figure 2.6: A set of symbols commonly used in Japanese comics. (a) cross-popping veins (anger); (b) question mark (incomprehension); (c) tears (grief); (d) exclamation mark (vigilance/notice); (e) sweat drop (anxiety); (f) parallel vertical lines (trouble); (g) red cheeks (shame) [86]

Additional elements can enhance the clarity and recognizability of emotions on a robot's face. One effective approach is the use of symbols, a technique commonly found in Japanese comics and animation. These symbols are typically placed near the character's face or integrated directly into facial features, such as replacing eyes with hearts or stars (Fig. 2.6), to convey emotions in a highly expressive and intuitive way [86].

The Uncanny Valley effect refers to the discomfort or eerie feeling people experience when encountering entities that appear almost, but not quite, human. This phenomenon can significantly impact user experience, making interactions unsettling rather than engaging. To avoid this effect, finding the right balance between realism and abstraction is crucial. Highly realistic humanoid robots or virtual characters often trigger the uncanny valley effect, so opting for a more stylized or abstract design can create a more comfortable and appealing experience for users [87].

Baby Schema Effect

The baby schema concept refers to a set of infantile traits that are highly appealing to humans and are known to elicit caretaking behavior and influence the perception of cuteness and attention. These traits are typically associated with infants and have been shown to trigger nurturing responses [88]. Studies have shown that robots with baby schema features are perceived as cuter and more trustworthy. For instance, one study found that robot faces with high baby schema were rated as cuter and more trustworthy compared to those with low or uncontrolled schema. This correlation helps in designing social robots that are more appealing and trustworthy to users [89]. The baby schema effect has significant consequences, influencing user behavior and engagement by triggering innate emotional responses and activating neural pathways associated with motivation and reward:

- **Motivation to Engage**: The baby schema effect can increase users' motivation to engage with the robot, as high baby schema features elicit stronger motivation for caretaking and positive affective responses [88].

- **Engagement and Satisfaction**: Viewing baby schema features activates the brain's reward system, particularly the nucleus accumbens, which is associated with reward processing and appetitive motivation. This neural activation can enhance user engagement and satisfaction [90].

- **Connection, Loyalty and Prolonged Use**: The baby schema effect triggers positive emotional responses such as liking and caring, which can foster a stronger emotional connection between users and the robot. This emotional bond can lead to increased loyalty and prolonged use [91] [92].

2.4.3 Body Motion

Research indicates that body movements can convey emotional states effectively. For example, a study using a simple disk-shaped robot showed that humans could infer emotional states based on movement parameters [93]. Another study confirmed that emotional body movements, even with limited degrees of freedom, could express emotions like joy, sadness, and surprise [94]. While facial expressions are a dominant modality for conveying emotions, combining them with body movements can enhance clarity and perception. For instance, ambiguous facial expressions were better understood when accompanied by body postures and movements [46]. This suggests that integrating facial and body expressions can lead to more effective emotional communication. To effectively express different emotions, a unique movement pattern was assigned to each. These motion characteristics are informed by existing research and theoretical models of expressive movement:

- **Joy**: Joyful movements are typically characterized by high activity levels, smooth and expansive gestures, and dynamic postures. These movements convey excitement and positivity, often incorporating broad and fluid motions [94].

- **Caring**: Caring is associated with positive valence and moderate arousal, which can be effectively conveyed through smooth, flowing movements and a calm demeanor. These gentle motions reflect warmth and attentiveness [95].

- **Pride**: Expressed through slow, supportive gestures and calming, affirming movements. These reinforce student achievements and signal approval [96].

- **Anger**: To express anger effectively, the movement should be strong, sudden, and characterized by violent shaking. These abrupt and forceful gestures mirror the intensity and high arousal associated with anger [94].

- **Fun**: Playful and entertaining emotions can be represented using simplified movement patterns, such as sinusoids, which create an impression of spontaneity and energy. These structured yet exaggerated motions help enhance the perception of playfulness and engagement [97].

- **Surprise**: Surprise is often conveyed through sudden, exaggerated movements, such as leaning backward and raising the arms quickly. These movements emphasize the unexpected nature of the emotion and create a clear visual representation of astonishment [98].

The recognition of robot emotions can vary significantly across cultures, and cultural limitations in the perception of movement and the expression of emotions in robots are evident. Controlling robot emotions based on the cultural context can help robots adapt to humans from culturally diverse backgrounds [99].

2.4.4 Vocal Expression

Vocal expression is crucial in enhancing robots' ability to convey emotions effectively, which is essential for natural and intuitive human-robot interactions. Changes in pitch, pitch range, speech rate, and volume can convey different emotions such as anger, fear, happiness, and sadness. Studies have shown that participants can recognize these emotions at rates significantly higher than chance, indicating the effectiveness of vocal prosody in emotional expression [100]. Higher-pitched voices are often perceived as more energetic and happy, while lower-pitched voices are seen as calm and serious. This affects overall interaction quality and user enjoyment [101]. Robots with a more expressive vocal range are perceived as more effective communicators. This expressiveness helps in making the interaction more engaging and contextually appropriate [102].

The consistency between vocal emotion and the user's affective state is essential for successful human-robot communication. Research has shown that humans can perceive empathy and emotions in robot speech, and they tend to prefer robots that employ empathetic speech over those with a standard robotic voice. For empathy to be effectively conveyed, it is crucial that the emotions expressed in the robot's speech align not only with the content of the language but also with the user's emotional state. If there is a mismatch between the robot's emotional expression and the user's feelings, it can negatively affect the interaction [103].

Human-like voices are typically preferred for their naturalness and familiarity, which foster more relatable and engaging interactions. However, when properly modulated, characterized or stylized voices can be more effective than natural voices in conveying specific emotional states with greater clarity. Research has shown that emotion perception accuracy varies depending on the emotion being presented. While a regular human voice tends to be favored for its higher user preference and naturalness, a characterized voice is often more appropriate for expressing emotions, demonstrating significantly higher accuracy in emotion perception [44].

Cultural background significantly affects how individuals perceive and interpret emotional cues from robots. For instance, Japanese participants are more attuned to vocal cues in multisensory emotion perception than Dutch participants, who rely more on facial cues. This indicates that cultural differences can influence the integration of auditory and visual emotional information [104]. Preferences for robot voices can vary widely based on cultural context. For example, a study found that British participants preferred robots with a Standard Southern British English accent or an Irish accent, while very few preferred a machine-like voice [105].

3 Design of the Emotiveness of the RSC

This chapter outlines the implementation of the RSC's emotiveness, drawing from previous studies. The development focuses on specific emotional responses that enhance the student learning process, which will be tested and verified in the following chapters. To achieve this, a set of common student scenarios has been created as the foundation for the RSC's responses. Then, facial expressions and vocal outputs were created. A digital twin of the RSC was developed to facilitate testing and implement the motions.

3.1 Robot's Emotions

Building on the literature analysis conducted in the previous chapter, the RSC has been designed to express a range of emotions that are particularly relevant to enhancing student engagement and learning. The emotions implemented for the RSC include anger, caring, surprise, joy, fun, and pride. These emotions were selected based on their impact on motivation, cognitive processing, and social interaction in an educational context.

3.2 Scenarios

Common scenarios experienced by university students were selected as the foundation for implementing RSC's emotional intelligence, aiming to elicit more specific emotional responses and enable more effective testing. By providing a real-life scenario and tailoring the responses to that scenario during the tests, students are better able to immerse themselves in the situation and capture the details of the robot's responses. This approach ensures that students' feedback contributes to the improvement of the RSC, aligning it more closely with the students' needs and learning experiences, ultimately leading to more accurate and relevant outcomes.

A diverse and creative set of scenarios has been developed, each carefully crafted to align with the information provided in Table 2.1 to maximize the efficiency of the robot's emotional responses. Each scenario represents a distinct but common learning moment for university students where the RSC's expressions are crucial in shaping the student's educational experience. These scenarios have been designed and described most simply and intuitively, ensuring they effectively capture real-life learning situations.

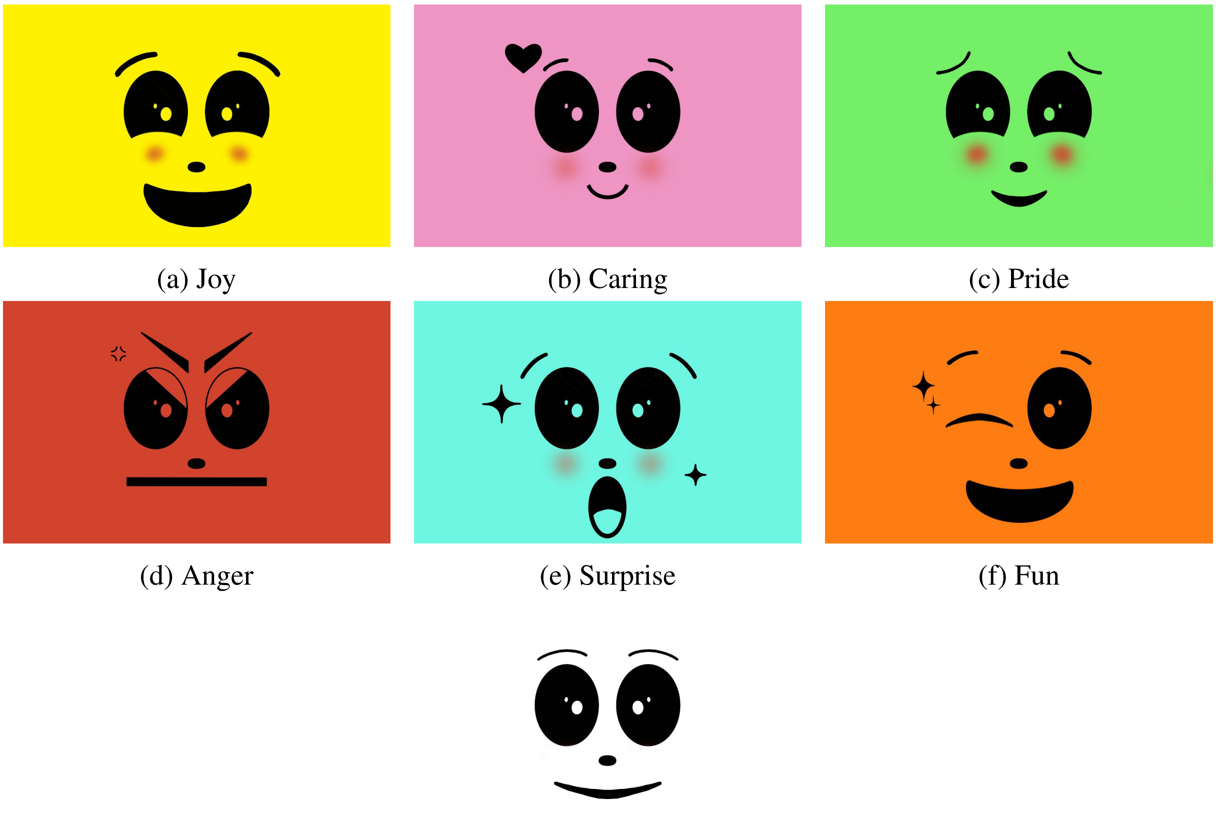
- **Joy**: You just finished all the tasks for the study session so now you can enjoy your free time and relax.
- **Caring**: You are anxious while juggling multiple deadlines and you're stressed. You would like to rest but you know that you still have a lot of things to do. You are also discouraged, you feel like you are not able to finish everything on time.

- **Pride**: You finally just understand a challenging topic you've been struggling with for weeks. You feel proud of your perseverance and growing mastery of the subject.
- **Anger**: You are supposed to study for an upcoming exam but you are instead scrolling through social media & procrastinating. You are aware of the importance of the task, but you are unmotivated and you don't want to start working for it.
- **Fun**: You have been reading for hours, looking passive and feeling disengaged. You are struggling to memorize and remember your work. You are bored and unmotivated.
- **Surprise**: You have a complex task to do: solving a set of challenging course problems. You expected it to take several hours, but you finished it in just 30 minutes.

3.3 Facial expression

A crucial aspect of the RSC in conveying emotions is its facial expressions. In developing a satisfying and effective set of expressions from scratch, several factors had to be carefully considered, including the baby schema, avoiding the uncanny valley, drawing inspiration from existing expressions, and refining the overall design. Sketchbook [106] app has been used (Version 6.1.1) to draw different facial expressions.

The final results can be seen in Figure 3.1.



(a) Joy (b) Caring (c) Pride

(d) Anger (e) Surprise (f) Fun

(g) Neutral

Figure 3.1: The final RSC's faces available on GitHub [107].

Here we present an overview of the generated facial expressions:

- **Joy**: A vivid yellow background. The face features a big smile, large open eyes with a lower cut-off, raised eyebrows, and rosy cheeks (Figure 3.1a).
- **Caring**: A soft pink background. The face has big, expressive eyes, a small delicate smile, thin curved eyebrows, rosy cheeks, and a small heart on the forehead (Figure 3.1b).
- **Pride**: A soft green background. The eyes have a lower cut-off, the eyebrows are gently curved, the face has a subtle smile, and the cheeks are intensely blushed (Figure 3.1c).
- **Anger**: A bold red background. The eyebrows and eyes are drawn with sharp, straight lines. The mouth is a firm, straight line, and an angry symbol appears on the forehead (Figure 3.1d).
- **Surprise**: A light blue background. The face displays an astonished open mouth, wide open eyes, raised eyebrows, slightly blushed cheeks, and a few sparkle-like accents around the face (Figure 3.1e).
- **Fun**: A bright orange background. The face has a wide, cheerful smile, one eye closed in a playful wink, and small star-like sparkles on the forehead (Figure 3.1f).

An additional expression, the neutral expression in Figure 3.1g, has been developed for practical reasons. Since students are not always in one of the previously defined emotional scenarios, this neutral expression serves as a default state. It ensures that during moments when a specific emotional response from the RSC is unnecessary, the character maintains a natural, non-intrusive presence. Additionally, it serves as a transitional state for the face animations, facilitating smooth changes between different emotional expressions.

Animation

A static face on an animated robot that moves and speaks would feel unnatural and uncomfortable to users, even leading to an uncanny valley. To create a more fluid and natural interaction, seamless transitions between the neutral and emotional expressions are essential. For each emotion, three different video sequences were required:

- Transition from neutral to the emotion
- Animation of the emotion itself (emotion to emotion transition)
- Transition from the emotion back to neutral

To achieve this, AI-driven tools have been utilized to animate facial expressions, ensuring smooth and realistic shifts that enhance the robot's expressiveness and engagement.

Various AI tools for image animation are available online, and a broad selection of them was tested for this face animation task. The goal was to create short video sequences that brought the face to life through subtle yet effective movements of the eyes, eyebrows, and mouth. Additionally, achieving a smooth transition between the neutral and emotional expressions was essential. However, only a few tools were able to deliver truly satisfying results, as many struggled to balance realism, fluidity, and natural expressiveness.

Veo 2 [108] was the AI tool chosen for this research for several key reasons. The most significant advantage was its ”AI Image to Video Generator” feature (Figure 3.2), which allowed users to provide both the first and last frames as input. This capability enabled the creation of smooth transitions between different emotional expressions. Additionally, the option to adjust the movement amplitude proved valuable in fine-tuning the expressiveness of the animations, ensuring they remained natural and engaging. Furthermore, variable video length options offered flexibility, allowing for customized animations suited to different scenarios and needs. To generate these animations, a prompt had to be inserted to define the specific motion between facial expressions, and a significant amount of testing was required before achieving optimal results. This process ensured that the animations accurately reflected the intended emotional states and met the desired standards for naturalness and coherence.

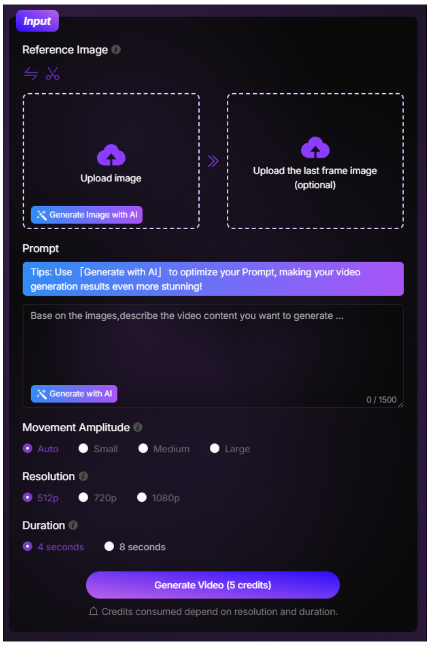


Figure 3.2: The screenshot of Veo2 input options of ”AI Image to Video Generator” feature

The prompts used to generate the final videos:

- ”Can you generate a video that changes the expression of the first frame, a neutral face, into the last fram, a [emotion] face. I would like also the transition of the backgroud color. I would like a simple transition, like a face changing its expression.”
- ”Can you make a small animation of this [emotion] expression. I just want the face to move a little bit, in order to make it feel alive. Do not change the background color. Just smooth and minimal”
- ”Please generate an animation from the first [emotion] face to the last neutral face. I want a smooth transition of the background color. I wan a minimal and smooth transition of the face”

With a total of six emotions and 3 videos required for each emotion, this resulted in 18 distinct videos, each lasting four seconds. These animations ensured smooth and natural emotional transitions, enhancing the expressiveness of the RSC. The final videos are available on GitHub [107].

3.4 Voice

A set of responses was developed (Table 3.1) to allow the RSC to react appropriately in different learning scenarios. The responses were tailored to fit the specific context of each scenario, with the text carefully crafted to align with the emotional tone required. Additionally, the length of each response was designed to last between 10 to 15 seconds, ensuring that the interactions remained concise and engaging without overwhelming the students. This timing also helped maintain a natural flow of conversation and allowed the RSC to transition smoothly between emotional expressions.

Table 3.1: Vocal Expressions for Different Emotions

Emotion	**Vocal Expression**
Joy	”Wow, you did it! This is such a big achievement, and it shows just how much effort and determination you’ve put into your work. Let’s celebrate this win together, and remember, accomplishments like this are what make all the hard work worth it.”
Caring	”It’s completely okay to feel nervous. Let’s take a deep breath together and start with the first task of the list. You worked hard until now, don’t give up. You’ve got this, and I’m here to help you every step of the way.”
Pride	”That’s absolutely incredible! You’ve demonstrated real discipline, perseverance, and growth throughout this process, and it’s clear that all your hard work has paid off. Achievements like this aren’t easy, and you should feel truly proud of yourself for what you’ve accomplished.”
Anger	”This is not acceptable! You’re wasting time that you can’t afford to lose. This work is important, and you need to take it seriously. Put your phone away, you will check it when you will finish here. Let’s start right away!”
Fun	”Hey, let’s mix things up! Stand up and move a little bit, some stretching will help! And then, how about a silly riddle to wake your brain up? I don’t have a heart but I can care, I break the monotony and help you prepare, who am I? . . . YOUR SOCIAL ROBOT!”
Surprise	”Wow, you just finished that so fast! I didn’t expect it to be done already—this is amazing! You must feel so good about this! This is the final proof that you should never underestimate yourself, you are capable of great things!”

As seen in the previous chapter in the Vocal Expression section, both natural and robotic voices are able to convey emotions. Additionally, a spectrum of voices is necessary to better address different cultures and personal preferences. Three voice profiles were developed for the RSC using various online tools [109] [110], offering users a choice:

- a warm and supportive female voice (Betty) [110]
- a calm and reassuring male voice (Raymond) [110]
- an artificial but still emotional robotic voice (Emma Multilingual) [109]

The final audio files are available on GitHub [107].

3.5 Digital Twin

For the development of the digital twin, ROS2 Humble [111] was used as the core framework. ROS2 is an open-source set of software libraries and tools that aid in building robotic applications. Additionally, RViz2 [112] was utilized for 3D visualization. RViz2 is the visualization tool for ROS2, allowing for real-time interactive visualization of sensor data, robot models, and simulation environments. This tool was instrumental in visualizing the digital twin, enabling effective monitoring and debugging of the system in a simulated environment.

3.5.1 URDF

The initial step in the development process involved creating a Unified Robot Description Format (URDF) file, the base for enabling a digital twin in simulation environments. This was accomplished by leveraging the mechanical design of the RSC, which is freely available as an open-source project on GitHub [113]. The original mechanical model was converted into STL (stereolithography) files, providing the necessary geometry for integration into the URDF framework. The URDF file was structured to incorporate all original RSC's robotic components, ensuring accurate representation in simulation. The components included in this model were the back panel, base, body, button, dual flippers, frame, lid, and screen (Figure 3.3).

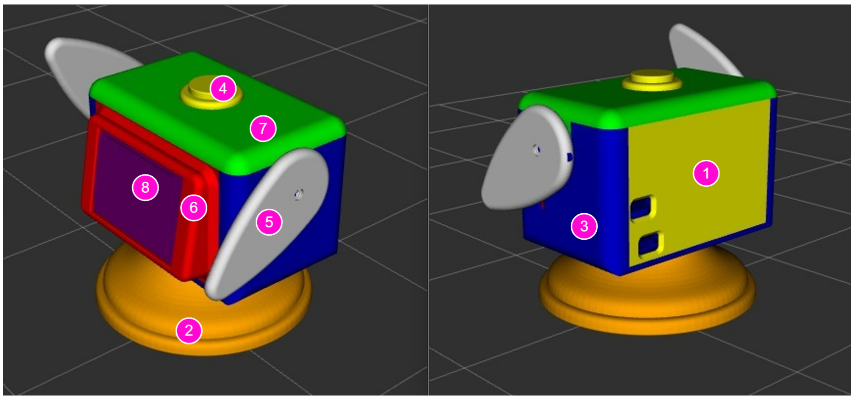


Figure 3.3: The result of the URDF file in the rviz environment. Different colors have been used to differentiate the parts used.
(1) back panel; (2) base; (3) body; (4) button; (5) dual flippers; (6) frame; (7) lid; (8) screen.

Each component was assigned a fixed joint connection, ensuring structural stability within the digital twin but for the flippers, which were designed with rotational joints to simulate the original RSC's movements. The rotational joints allow the flippers to move dynamically within the simulation, mirroring their intended range of motion in the real-world implementation. The final result can be found in the package `rsc_description` on the ROS2 repository of the RSC GitHub page [114].

3.5.2 Nodes

To enable expressive movement, a dedicated control node has been developed for each emotion. These nodes are based on insights from previous studies analyzed in the previous chapter as well as personal observations, allowing the RSC to convey a range of emotions effectively through its flipper movements. The final result can be visualized in Figure 3.4 and in the following description.

- **Joy**: The RSC raises its flippers to the horizontal plane and then shakes them in opposite direction energetically and rhythmically, reflecting excitement and happiness.

- **Caring**: The RSC gently and smoothly moves its flippers in a soft oscillating motion in opposite direction, simulating a nurturing and affectionate gesture.

- **Pride**: The RSC executes slow, controlled, and fluid movements of the flippers in opposite direction, reinforcing a sense of confidence and self-assurance.

- **Anger**: The RSC lifts its flippers close to the top position and then shakes them rapidly in opposite directions, demonstrating forceful and erratic movements that signify frustration or aggression.

- **Fun**: The RSC spins its flippers in opposite directions, creating a playful and dynamic motion that conveys amusement and liveliness.

- **Surprise**: The RSC swiftly raises its flippers to an almost fully extended position before shaking them, emphasizing an abrupt and startled reaction.

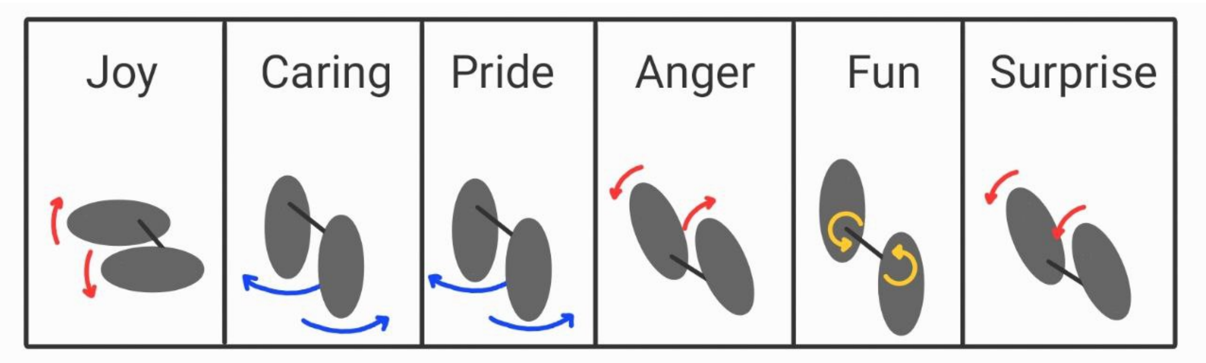


Figure 3.4: The flipper movement schematic for each emotion developed. Red means fast movement, Blue means slow movement, and Yellow means intermediate-speed movement. The length of the arrows represents the width of the motion.

The final result can be found in the package `rsc_movements_emotions` on the ROS2 repository of the RSC GitHub page [114].

3.5.3 Torso ROtation Digital Twin

A second URDF file and a complementary set of ROS 2 nodes have also been created to introduce an enhanced feature: torso rotation. Although the physical RSC prototype does not yet support this degree of freedom, the second digital twin incorporates it to evaluate whether adding torso movement can enrich the expression of emotional states and yield more nuanced nonverbal cues. All of the corresponding motion–control nodes have been implemented and are available on GitHub [114]. These extensions will be tested in future user studies and hardware trials to determine their impact on recognition accuracy, user engagement, and the overall effectiveness of the RSC's emotional communication.

3.6 The final integrations

Once all emotional expressions for each modality and scenario were developed, combining them into a final unified result for each emotion/scenario was necessary. However, RViz does not natively support audio output or handle audio data, and the video stream plug-in for face animations was not yet available for ROS2. Given the need for precise synchronization across modalities, it became clear that relying solely on ROS2 and RViz would be limiting. Additionally, during the testing phase, using a pre-rendered video was simpler and less prone to bugs than running the whole simulation, which could be more complex and error-prone. Therefore, video editing software was chosen, as it allowed for precise control over the synchronization of audio, visual, and animation elements, ensuring a seamless and effective presentation of emotional responses.

The video editing tool used is Canva [115]. Video screenshots of the digital twin movements, the video of the facial expression animations, and the output audio were overlapped and perfectly synchronized (Figure 3.5). The movement begins as the emotional face transitions, followed by the start of the audio. When the audio concludes, the movement stops, and the robot returns to its neutral expression.

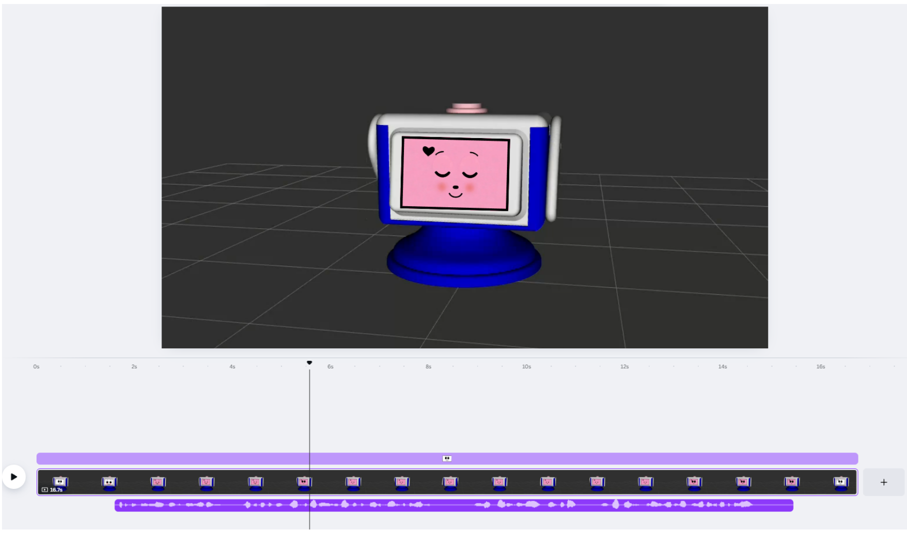


Figure 3.5: Screenshot of the video editing process on Canva [115] for Caring emotion

The final outcome consists of a total of 18 videos, with 3 videos dedicated to each emotion, corresponding to the different voice profiles used. Every video is available on GitHub [107]. These videos, along with their respective scenario descriptions, represent the final version prepared for testing, which will be presented in the next chapter.

4 Methodology

The goals of the experiment were as follows:

- To assess whether the users accurately recognized the different emotions expressed by the robot, to determine if the emotional design was effective (RQ3).
- To evaluate the overall user experience with the robot, including factors such as ease of interaction, engagement, and satisfaction (RQ4).
- To gather detailed user feedback and comments regarding the robot's performance, emotional expressions, and usability (RQ4).

This research was approved by the Research Ethics Committee of the University of Tartu (UT REC) on 28.03.2025 (Approval number: 399T-30).

4.1 Experiment Set Up

The experimental setup for this study includes a projector, a laptop, a tablet, and a recording device used in the following way (Figure 4.1):

Figure 4.1: Testing environment setup: A student views the virtual RSC video on a laptop (1), then inputs their answers in the questionnaire via Limesurvey on the tablet (2), while the scenario is projected on a projector screen (3).

- Projector: Used to display the six scenario descriptions for the participants.
- Laptop: Preloaded with 18 videos, organized into folders by voice type. Each video is labeled according to its corresponding scenario.

- Tablet: Equipped with a questionnaire and a user consent agreement for participants to complete before the study begins.
- Recording Device: Used to capture participants' responses to open-ended questions, enabling qualitative analysis.

This setup ensures a well-structured and interactive environment for gathering both quantitative and qualitative data.

4.2 Procedure

The following procedure outlines the steps participants will go through during the study, aslo shown in Figure 4.2:

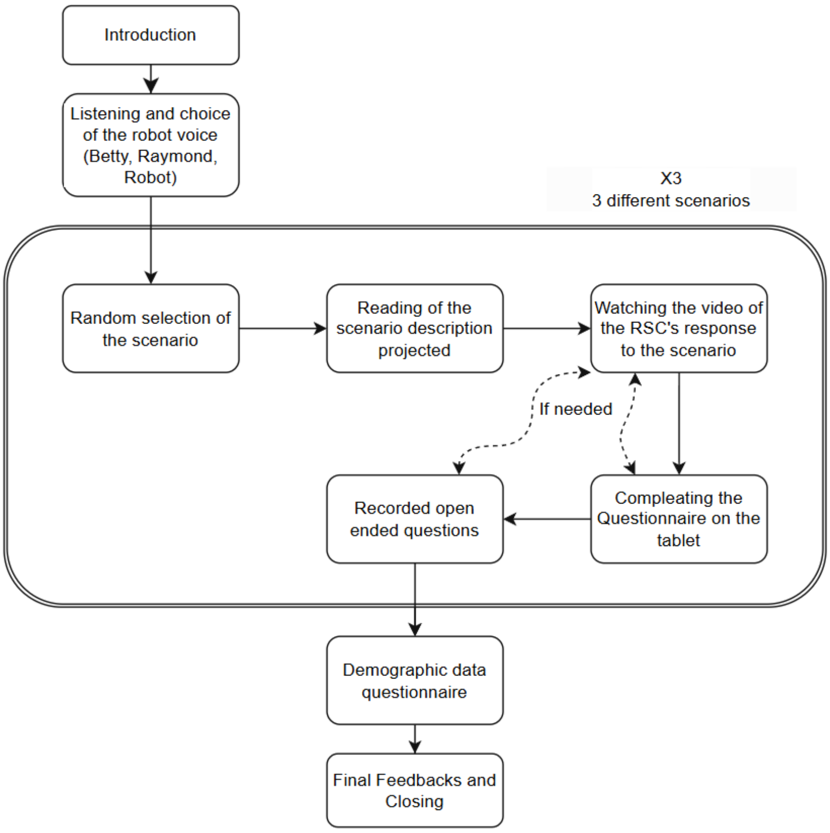


Figure 4.2: Flowchart of the Experiment procedure

- Introduction: Upon arrival, the participant will be introduced to the project, with an explanation of the study's purpose and why it is needed.
- User Consent Agreement: The participant will be invited to read and complete the user consent agreement.
- Voice Preference: The participant will see and listen to the robot interacting with three different voices. They will be asked to choose their preferred voice.
- Scenario Process: This process will be repeated three times with three different scenarios:

- The scenario will be randomly selected by the facilitator.
  - The scenario will be read aloud and projected for the participant.
  - The participant will watch the corresponding video on the laptop.
  - The participant will complete the questionnaire on the tablet based on the scenario.
  - A short semi-structured recorded interview will be conducted to gather further insights.

- Review Option: At any point during the procedure, the participant will have the option to re-read the scenario or re-watch the video.
- Demographic Data: The last part of the questionnaire will gather demographic data from the participant.
- A final feedback will be asked and recorded about the overall experience.

This structure ensures that the participant has ample opportunity to engage with the materials while providing comprehensive data for the study.

4.3 Data Collection

Quantitative Data

The questionnaire consists of six sections (full Questionnaire is available on GitHub [107]). The first section is completed by the facilitator and includes the pseudonym used to link the questionnaire answers to the recorded interview. The second section is dedicated to the User Consent Agreement, which must be completed before proceeding with the test. Sections 3, 4, and 5 are identical, each corresponding to one of the extracted scenarios. These sections include an initial question to identify the scenario and a second question to assess the emotional expression (RQ3) (Figure 4.3). This ensures that the emotional responses of the robot are accurately identified, helping to assess whether the design of these emotions effectively conveyed the intended emotional expression.

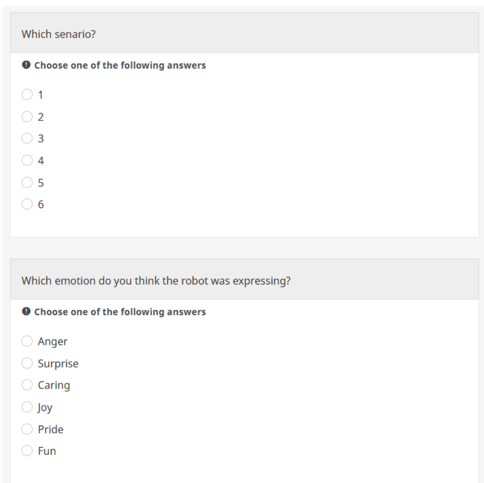


Figure 4.3: The first two questions from sections 3, 4, and 5 of the questionnaire.

Following these, the questionnaire includes 26 questions from the standard User Experience Questionnaire (UEQ) [116]. The UEQ measures a comprehensive set of user experience aspects, covering traditional usability metrics such as efficiency, clarity, and reliability, as well as user experience factors like originality and stimulation (RQ4). Additionally, the UEQ includes a built-in tool for data analysis, making it an efficient choice for evaluating user feedback.

Demographic Data

The final section of the questionnaire gathers demographic information. The questions:

- What do you study?
- What is your level of study? (BSc, MSc or PhD)
- What is your gender? (Female, Male, Non-Binary, I prefer not to answer)
- What is your nationality(s)? (Country or Region, I prefer not to say)
- How willing are you to incorporate new technologies, such as robots, into your learning routine? (Not at all, Slightly, Neutral, Somewhat, Very much)

This demographic data is collected to analyze potential differences in perceptions across various groups, providing a deeper understanding of how these factors might influence user experience and emotional recognition.

Qualitative Data

A small semi-structured interview was deemed the most effective for the qualitative analysis. Each interview was audio recorded and transcribed for further analysis. The base questions for the interview were:

- How did you find or feel about the robot response?
- How did you find or feel about the robot's expression, motion and colour for the emotion?

The first question focused on the user's general emotional response, capturing their overall perception of the robot's interaction. The second question delved deeper into the specific components, such as the robot's facial expression, motion, and color, assessing how these elements contributed to the emotion conveyed. This distinction allowed for a clearer separation between evaluating the holistic experience and focusing on individual features of the robot's emotional display (RQ4).

The semi-structured interview was conducted at the end of each scenario, providing scenario-specific insights. Additionally, a general assessment was carried out at the end of the study to collect final feedback and comments on the overall experience. This combined approach yielded both targeted insights and a comprehensive evaluation of the participant's complete experience.

4.4 Data Analysis

For the quantitative data, the UEQ Data Analysis Tool was used, with the answers grouped by scenarios and further categorized by factors such as gender, level of study, and field of study.

For the qualitative data, similar answers were grouped from the transcribed interviews. The responses were analyzed thematically, with recurring patterns in the participants' emotional reactions, perceptions of the robot's expressions, motion, and color identified. This approach allowed for a deeper understanding of the user experience, both on a scenario-specific basis and in terms of the overall interaction with the robot.

4.5 Participants

A total of 47 students participated in this study. Twenty-six were tested at the University of Tartu and twenty-one at the University of Guyana. The sample comprised 26 female and 21 male participants. In terms of academic level, 6 were PhD students, 10 were MSc students, and the remaining 31 were undergraduates (BSc). With respect to nationality, 24 identified as American (USA, Guyana, and Mexico), 4 as Asian (Kazakhstan, India, Pakistan), and the other 19 as European (Estonia, Latvia, Spain, France, Germany, Italy, Ukraine). All participants provided informed consent prior to data collection.

5 Results

5.1 Introduction to Results

5.1.1 User Experience Questionnaire (UEQ)

The User Experience Questionnaire (UEQ) [116] yields six separate scale means: Attractiveness, Perspicuity, Efficiency, Dependability, Stimulation and Novelty, but no single overall score. As recommended in the UEQ manual [116], we do not compute a global mean, since combining conceptually distinct scales would produce a value that cannot be meaningfully interpreted. Instead, we report each scale's mean and standard error to (a) detect potential outliers or misinterpretations at the item level, and (b) evaluate each dimension against the following benchmarks:

- Values between -0.8 and 0.8 indicate a neutral evaluation.
- Values above 0.8 indicate a positive evaluation.
- Values below -0.8 indicate a negative evaluation.

Each scale measures:

- **Attractiveness:** A user's overall impression of appeal and enjoyment.
- **Perspicuity:** How easy it is to understand and become familiar with the system.
- **Efficiency:** The extent to which users can achieve their goals with minimal effort.
- **Dependability:** The degree of control and predictability users feel when interacting.
- **Stimulation:** The extent to which the system is exciting and motivating.
- **Novelty:** How innovative and creatively surprising the experience is.

These six dimensions for each emotion scenario, followed by subgroup analyses by gender and academic level, are presented.

5.1.2 Participants Distribution

Table 5.1 summarizes the distribution of emotion scenarios experienced by participants, broken down by gender and academic level. Across all participants, the six emotions were roughly balanced. This even distribution ensures that subsequent analyses of recognition accuracy and subjective ratings are not biased by unequal exposure to any particular emotion.

Table 5.1: Scenario distribution by gender and academic level

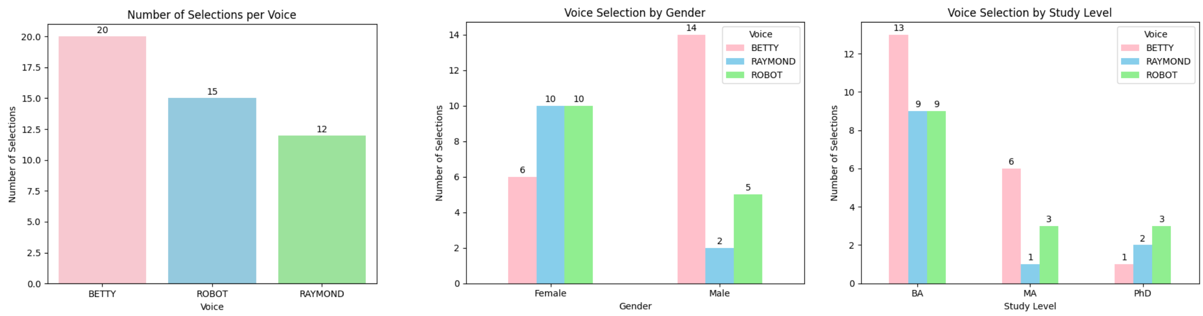
Gender	Level	Anger	Fun	Joy	Surprise	Caring	Pride
Male	BSc	8	7	6	5	5	8
	MSc	2	4	3	3	2	4
	PhD	2	1	0	1	1	1
	Total	12	12	9	9	11	13
Female	BSc	9	10	11	9	11	4
	MSc	1	3	3	2	1	2
	PhD	3	1	2	2	2	2
	Total	13	14	16	13	14	8
Total		25	26	25	22	22	21

5.2 Voice Choice

Overall, participants showed a clear favorite for the "Betty" voice: it was chosen by 20 of the 47 participants, compared with 15 for "Robot" and 12 for "Raymond" (Figure 5.1a).

When we break this down by gender, a striking divergence emerges (Figure 5.1b): male participants overwhelmingly preferred Betty (14 out of 21), whereas female participants split more evenly between Robot and Raymond (10 each) and were less likely to choose Betty (6 out of 26).

Looking at academic level (Figure 5.1c), undergraduates (BA) tended to favor Betty (13 of 31) but also selected the other two voices almost equally (9 each); master's students (MA) leaned toward Betty (6 of 10) and only rarely chose Raymond or Robot, while PhD candidates showed the opposite pattern, favoring the more synthetic voices (Robot 3, Raymond 2) over Betty (1).



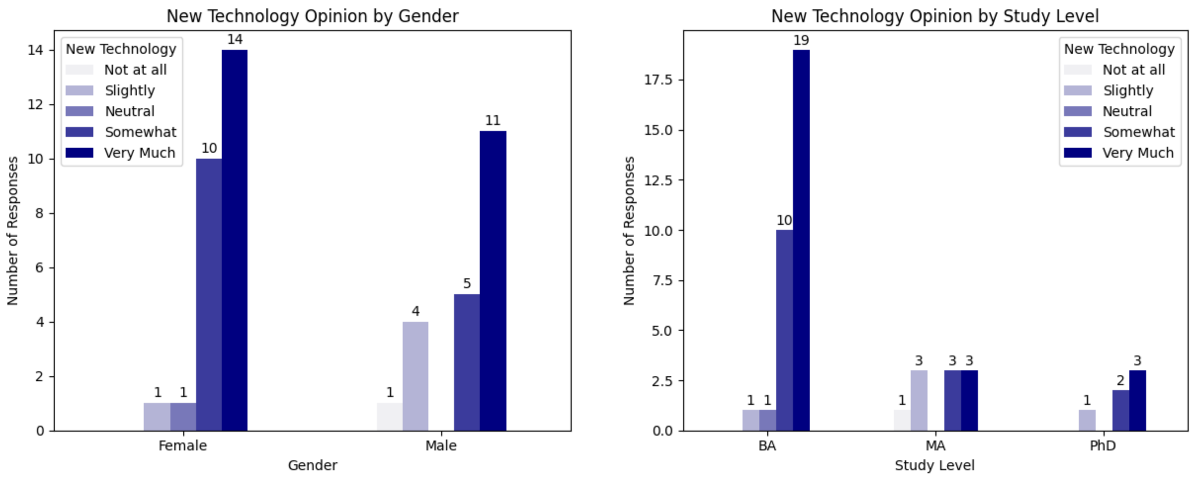
(a) Number of selection per voice (b) Voice selection by gender (c) Voice selection by Level

Figure 5.1: Distribution of Voice Choice across participants, other graphs can be found on GitHub [107]

5.3 New Technology

These results are based on the final question of the questionnaire, in which participants rated their willingness to integrate new technologies into their study sessions. Overall, participants expressed strong enthusiasm for the new technology (Figure 5.2), with the vast majority selecting "Somewhat" or "Very Much". Breaking it down by gender (Figure 5.2a), 24 of 26 women (10 "Somewhat," 14 "Very Much") rated the technology positively, with only two women choosing the lower categories. Men showed a similar but less enthusiast pattern: 16 of 21 men fell into the top two bins, while the remainder were split across "Slightly" and "Not at all".

Examining academic level (Figure 5.2b), undergraduates were the most enthusiastic: 29 of 30 chose "Somewhat" or "Very Much", with only one "Slightly" and one "Neutral." Master's students displayed a more even spread. PhD candidates likewise skewed positive (3 "Very Much," 2 "Somewhat",1 "Slightly," none "Neutral" or "Not at all").



(a) New Technology Opinion by Gender (b) New Technology Opinion by Level

Figure 5.2: New Technology Opinion, other graphs can be found on GitHub [107]

5.4 Emotional Recognition

The normalized confusion matrix in Figure 5.3 illustrates the participants' classification performance across six emotional categories: *Anger*, *Caring*, *Fun*, *Joy*, *Pride*, and *Surprise*. The overall accuracy achieved is 58.16%.

Anger and *Caring* are the most accurately classified emotions, with recognition rates of 96.0% and 90.9% respectively. This suggests that these emotional expressions were more clearly conveyed and easily interpreted by participants.

In contrast, *Fun*, *Joy*, *Pride*, and *Surprise* show significantly lower classification accuracy and a higher degree of confusion among them. For example, only 38.5% of *Fun* responses were correctly identified, with a large portion misclassified as *Caring* (53.8%). Similarly, *Joy* was often misclassified as *Caring* (40.0%) or *Pride* (20.0%).

These results highlight a notable overlap in the perception of positive emotions, indicating that participants may have found it difficult to distinguish between nuanced affective states such as *Joy*, *Pride*, and *Surprise*. This trend suggests a need for more distinctive multimodal cues when designing affective robot responses involving positively valenced emotions.

The accuracy results show no notable differences in emotional recognition across gender, study level, or voice type. Male (60.32%) and female (56.41%) participants performed similarly, as did those exposed to the Robot (60.00%) and Betty (60.00%) voices, while Raymond's voice yielded slightly lower performance (52.78%). Accuracy across educational levels was also comparable: BSc (59.14%), MSc (53.33%), and PhD (61.11%).

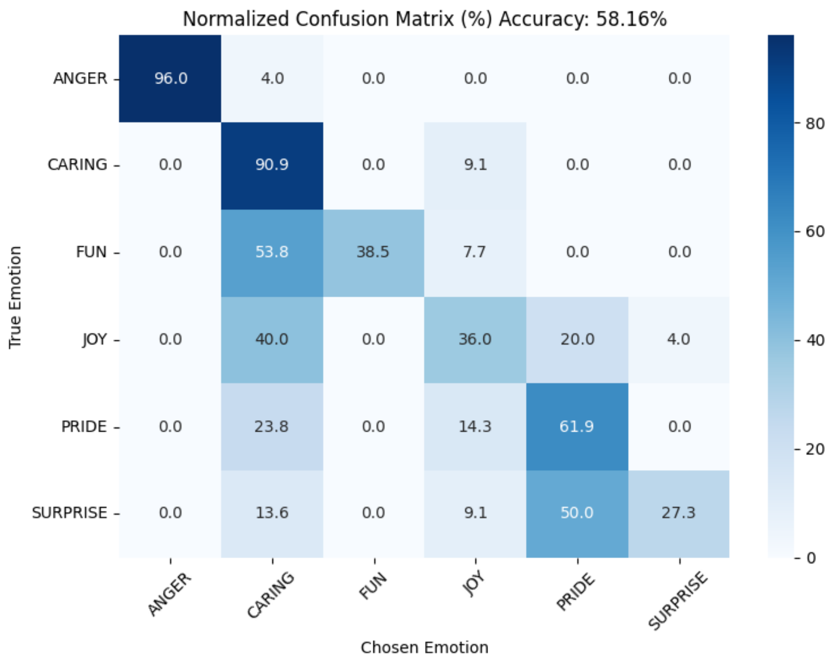


Figure 5.3: Emotional Recognition Confusion Matrix

However, more pronounced differences were observed across geographic regions. Participants from Europe and Asia achieved the highest recognition accuracy (66.67%), whereas those from America reached only 50.00%, suggesting a potential cultural or contextual influence in emotion interpretation. The related confusion matrix can be found on GitHub [107].

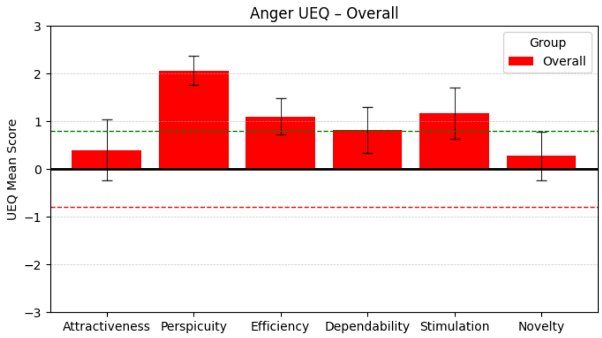
5.5 Anger Scenario

5.5.1 UEQ Results

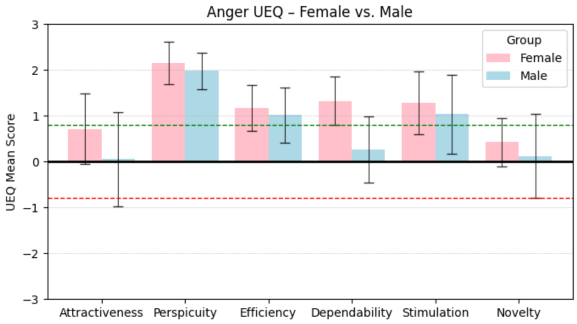
Figure 5.4 presents the mean UEQ dimension scores (with their confidence intervals) for the Anger scenario, both overall (Figure 5.4a) and broken down by gender (figure 5.4b) and academic level (Figure 5.4c). The UEQ results for the anger scenario reveal an overall positive evaluation across most dimensions, with *Perspicuity* standing out as the most appreciated quality, followed by *Stimulation*, *Efficiency*, and *Dependability*, while *Attractiveness* and *Novelty* received neutral-to-positive scores.

When comparing gender responses, female participants rated the system more positively across all dimensions, especially in terms of *Dependability* and *Attractiveness*. Male ratings, while still generally favorable, were slightly more reserved.

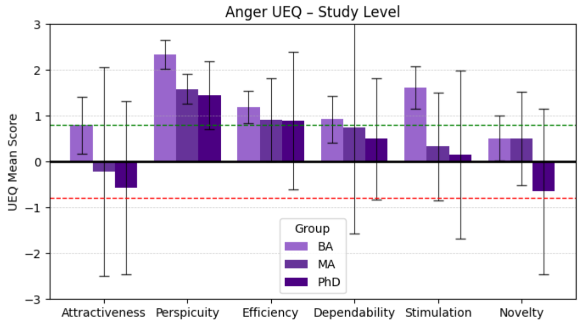
Differences were also observed across study levels. PhD participants provided consistently lower scores, particularly in *Attractiveness*, *Novelty*, and *Stimulation*. BA students showed greater enthusiasm compared to the other study levels across all dimensions, especially in terms of *Stimulation* and *Attractiveness*, while MA participants were overall more moderate. These comparisons indicate that perceptions of the anger response vary more significantly by study level than by gender.



(a) Anger Results



(b) Anger Results Female - Male



(c) Anger Results Study Level

Figure 5.4: Anger UEQ Results

5.5.2 Interview Results

Emotional Interpretation

Seven participants provided *negative feedback*, describing the response as "too aggressive" and associating it with feelings such as *annoyed*, *guilty*, or even *ashamed*. They expressed that the message was neither "helpful" nor "motivating," with one noting, "it made me feel even worse." Two participants explicitly compared the robot's angry tone to that of a scolding parent, saying it reminded them of their *mum*. This group, mainly PhD students, generally expressed a preference for a *kinder, more supportive response*, ideally paired with *practical advice* to help address procrastination.

In contrast, four participants felt the response was *not aggressive enough*. They believed it would likely be ignored in a real scenario, describing it as "just annoying but not convincing," and referring to the angry robot as *cute*, *funny*, and *not intense enough*.

Meanwhile, ten participants *appreciated the assertive tone* of the message. They used phrases such as "it would work for me," "it was not aggressive, just assertive," and "it made me aware of the situation, like a wakeup call." These individuals reported feeling *motivated and self-aware*, acknowledging their own tendency to procrastinate. They particularly valued the combination of *assertiveness* at the beginning and a *kinder, more supportive ending*.

Evaluation of Robot's Expressivity

The anger response expression was positively received by all participants. Seven individuals specifically highlighted the *effectiveness* of the chosen color, stating that "red was really good for this." Five participants praised the facial expression, describing it as "perfect to send this message." Eight participants commented on the *clarity* of the robot's physical gestures, noting that "the hand movement was clear." Overall, the combination of these elements was described as *highly expressive*: "That is expressive. The red really brings out the anger in it. The arms, the frustration."

The tone of voice did not receive a uniform response across participants. Two individuals described it as "too angry and pressing," listening to Robot and Betty voices. In contrast, two others felt the voice could have been more *intense*, describing it as "a bit more aggressive" or "pretty neutral," in reference to Raymond and Betty.

5.6 Fun Scenario

5.6.1 UEQ Results

Figure 5.5 presents the mean UEQ dimension scores (with confidence intervals) for the Fun scenario, both overall (Figure 5.5a) and by gender (Figure 5.5b) and academic level (Figure 5.5c). The overall evaluation was positive across all dimensions, with particularly high ratings in *Attractiveness* and *Perspicuity*.

When comparing gender, female participants consistently gave higher ratings across most dimensions, especially for *Attractiveness*, *Perspicuity*, and *Dependability*. The only exceptions were *Efficiency* and *Novelty*, where male participants provided slightly higher scores.

At the academic level, BSc students rated the system highest across all dimensions, showing the strongest enthusiasm for the fun response. PhD participants evaluated the system more positively than MSc participants on four dimensions: *Attractiveness*, *Perspicuity*, *Dependability*, and *Stimulation*. These results suggest that while the response was well received overall, it resonated most strongly with undergraduate users.

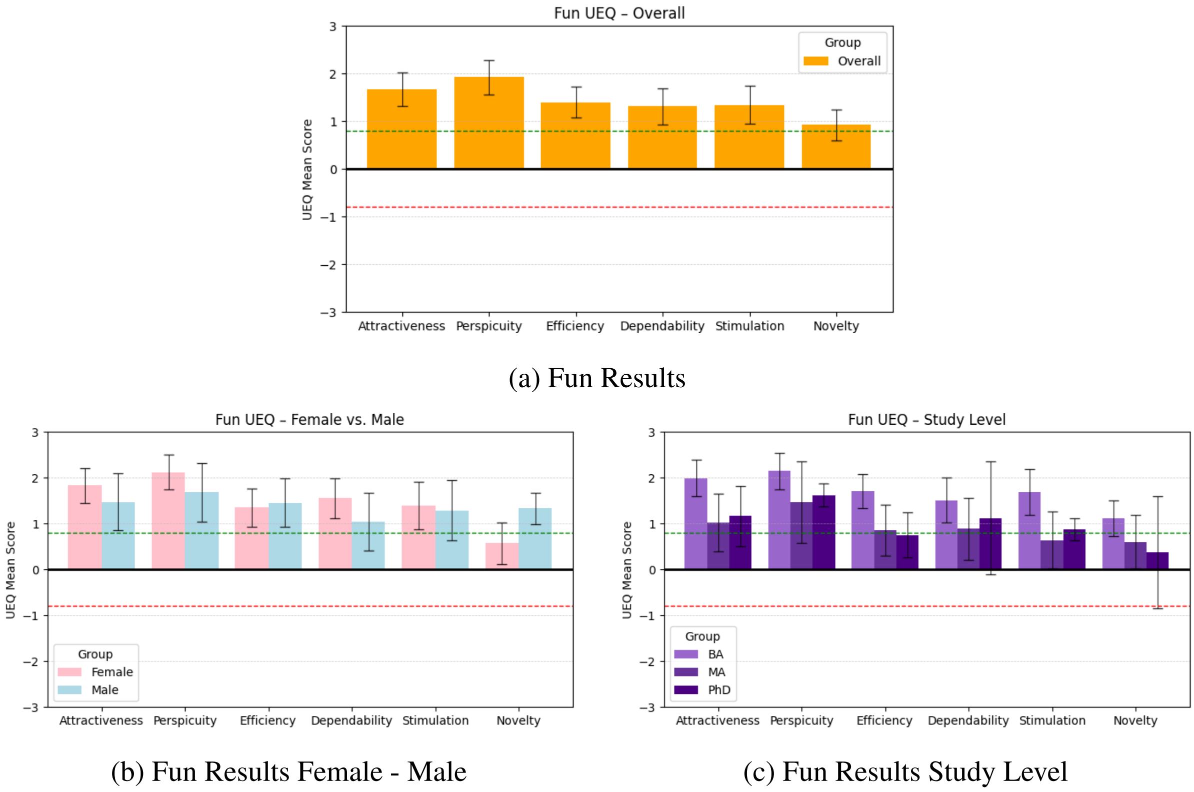
5.6.2 Interview Results

Emotional Interpretation

Seventeen participants appreciated the fun response, using terms such as *motivating*, *encouraging*, *supportive*, *positive*, and *helpful*. Five of them specifically noted that this approach would help in moments of boredom or disengagement, expressing that it allowed them to "re-engage," "wake up," or "disconnect."

Thirteen participants highlighted the value of the practical advice provided by the robot, particularly the stretching suggestions. Notably, two participants actually performed the suggested stretches during the simulation. Although the advice was simple and familiar, it was appreciated for its timely reminder: "I like the idea of taking a break in these situations, but I don't remember it in the moment; I would like some reminder."

The inclusion of a riddle was met with mixed reactions. While five participants appreciated it, six described it as *cringe* and two as *weird*, resulting in a total of eight participants expressing discomfort. Five participants explicitly stated a preference for a moment of rest without any mentally demanding activity, whereas three participants welcomed the idea of incorporating such a playful element into their disconnecting break.



(a) Fun Results

(b) Fun Results Female - Male

(c) Fun Results Study Level

Figure 5.5: Fun UEQ Results

Evaluation of Robot's Expressivity

Six participants expressed reservations about the use of the color orange in this response. Some described it as suggesting that "something is wrong" or noted it "felt aggressive," while others simply stated they "didn't like the color, I am biased."

Eleven participants appreciated the motion, describing it as a "happy dance" and emphasizing its ability to "catch your attention when you're passive." However, three participants were less positive, finding the overall expression as "too much," "confusing," or "not organic."

Six participants responded positively to the robot's facial expression, using descriptors such as *cute*, *funny*, *cheerful*, and *motivational*.

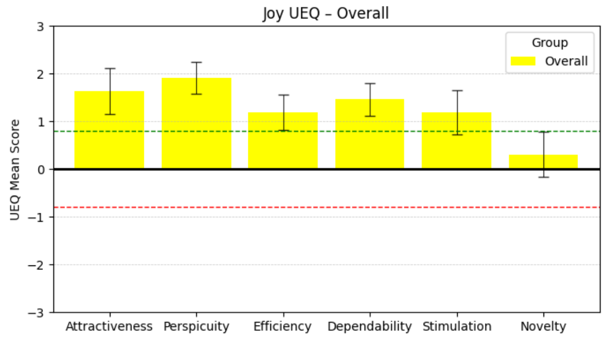
5.7 Joy Scenario

5.7.1 UEQ Results

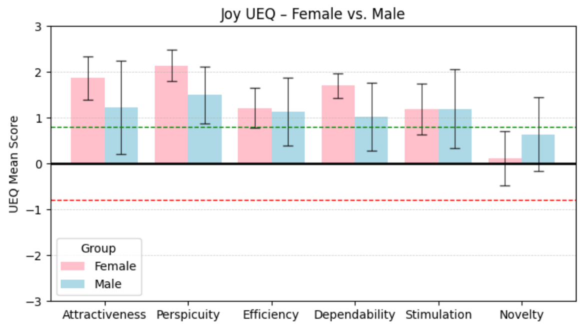
Figure 5.6 shows the mean UEQ dimension scores (with confidence intervals) for the Joy scenario, overall (Figure 5.6a) and segmented by gender (Figure 5.6b) and academic level (Figure 5.6c). All dimensions received positive evaluations, with particularly strong scores for *Perspicuity* and *Attractiveness*. The only exception was *Novelty*, which received more neutral-to-positive feedback.

In the gender comparison, female participants reported significantly higher levels of satisfaction in *Attractiveness*, *Perspicuity* and *Dependability*. However, scores for *Efficiency* and *Stimulation* were comparable between genders, and male participants rated *Novelty* higher.

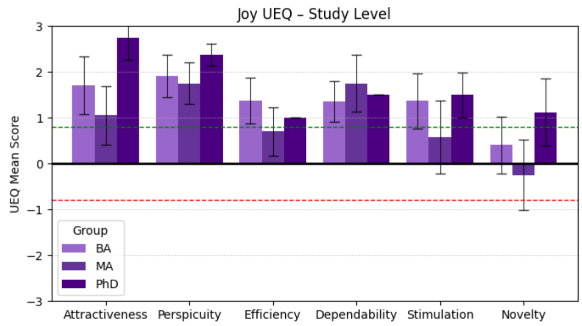
Across academic levels, PhD participants gave the highest ratings for *Attractiveness*, *Perspicuity*, *Stimulation*, and *Novelty*, while BSc students rated *Efficiency* most positively. MSc participants stood out for scoring highest in *Dependability*. These findings indicate that while the Joy scenario was positively received overall, participant perceptions varied slightly based on academic experience.



(a) Joy Results



(b) Joy Results Female - Male



(c) Joy Results Study Level

Figure 5.6: Joy UEQ Results

5.7.2 Interview Results

Emotional Interpretation

Fifteen participants described their emotional response positively, using terms such as *supportive*, *recognised*, *warm*, *appreciated*, and *celebrated*. Comments like "sometimes we forget to praise ourselves" and "this type of support would help students to have a better mood and to get motivation" emphasized the emotional value of the message. In contrast, five students described the response as *annoying*.

Evaluation of Robot's Expressivity

Nine participants explicitly appreciated the color, describing it as *warm*. Thirteen participants highlighted the expression as *supportive*, *happy*, *cute*, and *cheering*. Seven participants specifically pointed out the role of the motion in conveying emotion, stating that "the movement was like saying 'congratulations'" and "the movement looked like cheering you up." Additionally, four appreciated the *smiling face*, and two mentioned the *blushing* as a positive detail.

However, four participants did not enjoy the movement; two considered it "too slow," while two found it "too much."

Three participants perceived the voice as "a bit sarcastic," all of whom had experienced the robotic voice.

Feedback on the length of the message was mixed: six participants appreciated the duration as it was, another seven found it too long, especially "the second part," while two participants expressed a preference for a longer version, explaining, "I like to be celebrated."

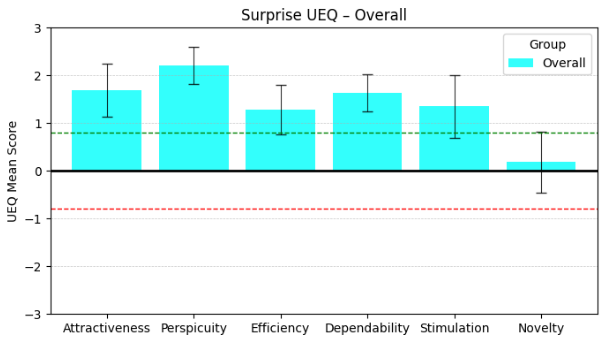
5.8 Surprise Scenario

5.8.1 UEQ Results

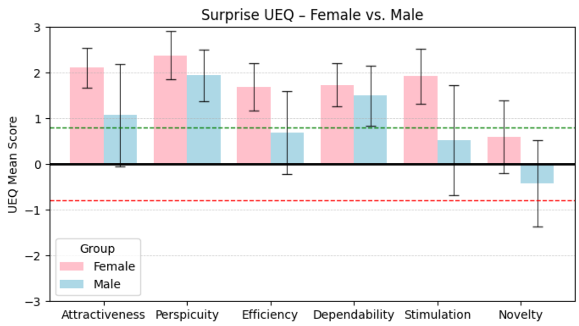
Figure 5.7 presents the UEQ dimension scores for the Surprise scenario, including overall (Figure 5.7a), gender-based (Figure 5.7b), and study-level comparisons (Figure 5.7c). Overall, the response was positively evaluated across all dimensions, with *Perspicuity* standing out as the most appreciated aspect. *Novelty* received the lowest rating but remained within a neutral range.

Female participants rated the system more positively across all dimensions, especially in terms of *Attractiveness*, *Efficiency*, and *Stimulation*.

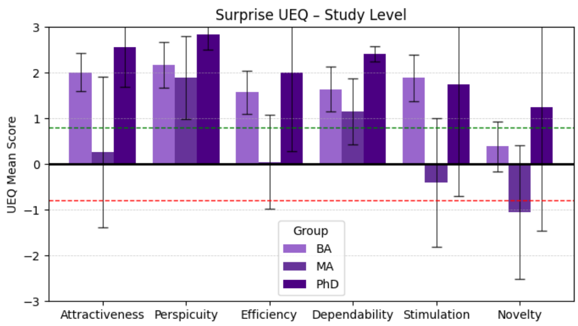
At the study level, PhD participants consistently rated the system highest across all dimensions except for *Stimulation*, where BSc students scored slightly better. MSc participants provided notably lower scores in general, especially for *Attractiveness*, *Efficiency*, *Stimulation*, and *Novelty*. These results suggest that user experience with the Surprise scenario was perceived more favorably by more academically experienced participants.



(a) Surprise Results



(b) Surprise Results Female - Male



(c) Surprise Results Study Level

Figure 5.7: Surprise UEQ Results

5.8.2 Interview Results

Emotional Interpretation

Twenty participants appreciated the response in this scenario, using words such as *satisfaction*, *motivating*, *proud*, *happy*, *joyful*, *supporting*, *positive*, and *friendly*. Participants expressed their appreciation through comments like: "it looks like your fan number one," "it's nice because you don't always feel like saying these things to yourself," "it's nice taking the time to appreciate," "it made me feel like I did a great job," and "it evaluates positively my work and motivates me to do more."

Evaluation of Robot's Expressivity

Seven participants explicitly appreciated the color used in the response.

Sixteen participants valued the way the robot expressed the emotion, using descriptors such as *celebrating*, *cheering*, and *happy*. Nine participants specifically highlighted the motion, while another seven appreciated the facial expression, noting features such as blushing and stars on the face.

Only two participants did not appreciate the expression, describing it as *cliché* and *irritating*. Two additional participants found the expression confusing, stating "I liked what it was supposed to do and it conveyed the emotion, but not in the best way", remarking that "too much is going on."

Similarly, three participants commented that the response did not feel *natural* or *genuine*.

Regarding the length of the message, five participants said it was appropriate as is, whereas seven indicated it was too long and would have preferred a shorter version.

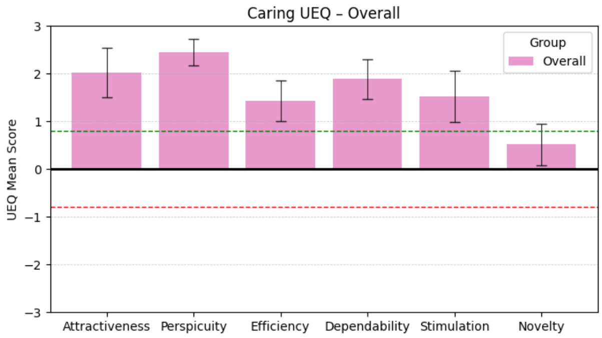
5.9 Caring Scenario

5.9.1 UEQ Results

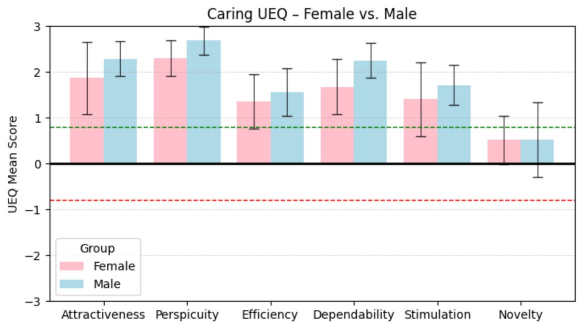
Figure 5.8 presents the UEQ dimension scores for the Caring scenario, both overall (Figure 5.8a) and segmented by gender (Figure 5.8b) and academic level (Figure 5.8c). The overall results are positive across all dimensions, with particularly high ratings for *Perspicuity* and *Attractiveness*. *Novelty* received the lowest ratings but remained in the neutral-to-positive range.

In terms of gender comparison, male participants rated the system higher across most dimensions, notably in *Dependability*. While female participants gave similarly positive evaluations, their scores were generally slightly lower.

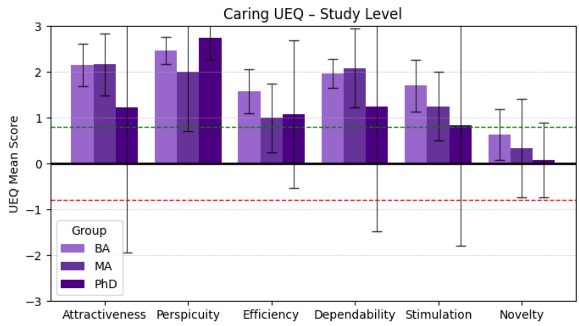
When comparing academic levels, the results appear relatively homogeneous, with no major variation across groups. All three levels showed consistently favorable evaluations across the dimensions, indicating a shared appreciation for the caring response regardless of educational background.



(a) Caring Results



(b) Caring Results Female - Male



(c) Caring Results Study Level

Figure 5.8: Caring UEQ Results

5.9.2 Interview Results

Emotional Interpretation

Seventeen participants appreciated the robot's response in this stressful scenario, using terms such as *motivating*, *calm*, *supportive*, *positive*, *encouraging*, *caring*, and *empathizing*. Their responses included comments like: "good way to handle the situation when you're overwhelmed," "it validated my feelings," "it gave me another reason to believe in my abilities," and "sometimes you just need to hear someone believing in you to get out of the dark."

In contrast, only three participants expressed mixed or less positive reactions, stating: "it didn't change much my feelings," "I might need something that's going to push me to take more action," and "understands but doesn't motivate me."

Evaluation of Robot's Expressivity

Thirteen participants appreciated the use of the color pink, describing it as *supportive*, *calm*, *sweet*, and *friendly*, while three participants did not appreciate it, stating it was "overly lovely."

Thirteen participants valued the expression's *cuteness*, referring to elements such as the heart, light blushing, and a movement that resembled "a little hug." However, four participants felt these characteristics were not appropriate for the context.

Three participants highlighted and appreciated the robot's attempt to connect and interact with them on a more personal level.

Eight participants appreciated the practical advice provided to help cope with the stressful situation, such as taking a deep breath, describing it as *helpful* and *reassuring*. Nonetheless, three participants expressed a desire for more practical advice, specifically tailored to the situation.

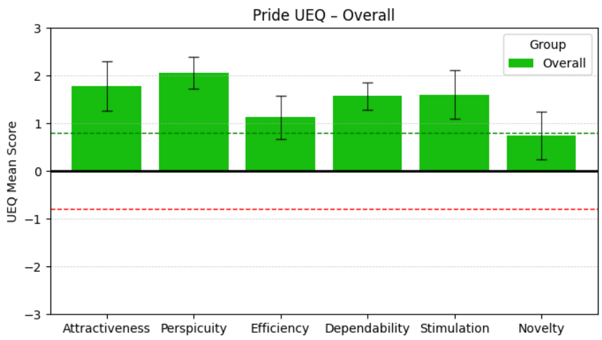
5.10 Pride Scenario

5.10.1 UEQ Results

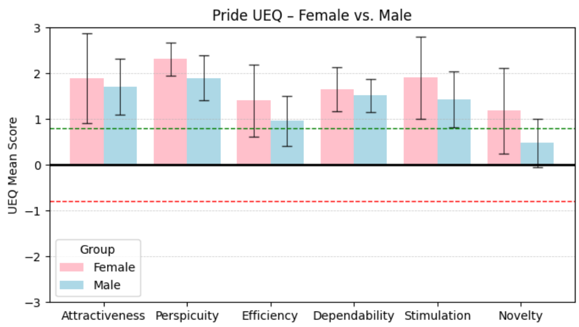
Figure 5.9 presents the UEQ scores for the Pride scenario, both overall (Figure 5.9a) and segmented by gender (Figure 5.9b) and study level (Figure 5.9c). The overall feedback was positive, with *Perspicuity* and *Attractiveness* receiving the highest ratings. *Novelty* was the least appreciated dimension but still scored within a neutral-to-positive range.

Female participants rated the system more positively overall, particularly in the *Novelty* dimension. Male participants also gave positive evaluations, though generally slightly lower across most dimensions.

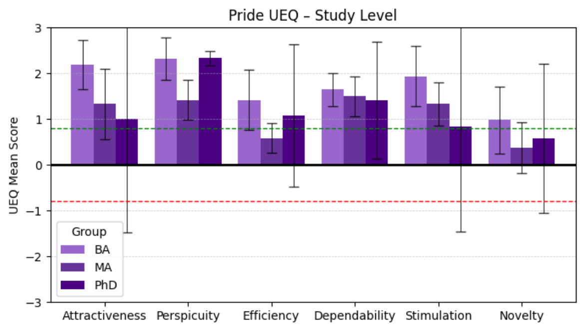
Across academic levels, responses were relatively homogeneous, indicating consistent appreciation for the pride response. Slight differences emerged, with BSc students rating *Attractiveness* and *Stimulation* higher, while MSc participants gave a significantly lower score in *Perspicuity* and *Efficiency*. These results suggest that while perceptions of the pride scenario were broadly consistent, minor variations were observed based on educational background.



(a) Pride Results



(b) Pride Results Female - Male



(c) Pride Results Study Level

Figure 5.9: Pride UEQ Results

5.10.2 Interview Results

Emotional Interpretation

Fifteen participants appreciated the pride response in this scenario, using terms such as *caring*, *supportive*, *recognised*, *proud*, *appreciated*, *celebrated*, *motivating*, and *accomplished*. Their comments included reflections like: "sometimes we don't take the time to celebrate and appreciate what we do," "it's nice to have someone that tells you that you did a good job," and "I felt good, after mastering something, having someone appreciate your hard work and your achievements."

In contrast, only three participants did not share this sentiment, describing the response as "unpleasant" or stating, "I didn't feel motivated."

Evaluation of Robot's Expressivity

Sixteen participants appreciated the expression in all its aspects, including both motion and facial features, frequently using the word *cute*. Two participants, however, did not appreciate the blushing element, describing it as "a bit silly" or "too affectionate."

The color green received mixed feedback: while ten participants expressed appreciation, four explicitly stated they did not like it.

Two participants found the overall expression *too fake* or *not very realistic*.

Regarding the duration of the response, eight participants considered the length appropriate, while five indicated they would have preferred a shorter version.

5.11 Final Feedbacks

Nineteen participants provided very positive overall feedback about their experience with the robot, emphasizing its potential usefulness in situations of loneliness, lack of motivation, and difficulty with organization. Comments included: "it's a nice support when you have to study alone," "I would like to have it, it can motivate me," and "it would be helpful having someone telling you what to do."

Among all participants, six stated that they either did not notice the colors or did not care about them.

Seven participants expressed a desire for the robot's mouth movements to be synchronized with its voice.

Three participants who had initially chosen the robotic voice later regretted their choice, describing it as "too impersonal."

Finally, five participants, all native speakers of English, Spanish, or Italian, commented that the robot's speech was too slow.

6 Analysis and Discussion

6.1 Emotional Recognition

The recognition results reveal both strengths and challenges in the emotional expressivity of the robot. One of the most encouraging outcomes is the clear distinction participants made between positive and negative emotional responses. Misclassifications rarely crossed this valence boundary, indicating that the system effectively conveyed whether the response was meant to be supportive or corrective in tone.

However, there was considerable overlap among the positively valenced emotions. In particular, *Joy*, *Pride*, and *Surprise* were often confused with one another, as they all celebrated some form of personal achievement. Similarly, *Fun* and *Caring* were frequently misclassified, as both aimed to improve the user's mood and provided encouragement during difficult moments. These results suggest that while the intent behind the responses was generally understood, the nuances between different types of positive reinforcement may require more distinct expressive cues.

An additional insight emerges from the interpretation of the label *Caring*. Many participants tended to associate this term broadly with supportive intentions, an attribute that, to some extent, applied to all the robot's responses. This highlights a possible semantic ambiguity in emotion labeling, where users may default to *Caring* when they sense any empathic or prosocial behavior. Future designs might benefit from reconsidering emotion labels or providing clearer distinctions in the robot's emotional delivery to reduce conceptual overlap.

The lack of significant differences across gender, academic level, and voice type indicates a relatively uniform interpretation performance among diverse user groups, which is encouraging from a design perspective. However, the regional differences, particularly the lower accuracy among American participants compared to their European and Asian counterparts, suggest that cultural background may influence emotional perception and should be considered when designing affective systems intended for international use. This reinforces the importance of cultural calibration in affective robot communication and the potential benefits of adaptive or customizable emotional expression models.

6.2 Anger Scenario

The results from the Anger scenario indicate that while the interaction design was generally well-received, it elicited a range of reactions depending on user characteristics and expectations. The UEQ scores show that participants evaluated the system positively, especially for *Perspicuity* and *Stimulation*, confirming that the robot's intent and communicative clarity were perceived effectively. However, scores for *Attractiveness* and *Novelty* were less enthusiastic, reflecting possible discomfort with the aggressive tone or a lack of perceived originality in the approach.

The interview data provide important nuance to the quantitative trends. Participants were split between those who found the anger response too harsh, those who found it too mild, and those who appreciated its assertiveness. This suggests that emotional feedback, particularly anger, is highly subjective and context-sensitive, likely influenced by personal motivation styles and prior experiences with support systems. Interestingly, a number of participants referenced parental associations, underscoring the emotional weight that such expressions can carry.

Design elements such as color and movement were widely praised and appeared to effectively reinforce the intended emotion. However, inconsistencies in the perceived tone of voice suggest room for improvement in multimodal synchrony, such as aligning mouth movement with speech or calibrating vocal intensity.

Gender and study-level differences further emphasize the importance of tailoring emotional support to user demographics. Female and BA-level participants responded more positively overall, while PhD students were more critical, perhaps indicating different expectations or thresholds for motivational interventions.

Taken together, these findings support the idea that while anger-based expressions can be effective, they must be carefully balanced with clarity, empathy, and personalization to avoid counterproductive effects.

6.3 Fun Scenario

The Fun scenario was broadly well-received by participants, as reflected in both UEQ scores and interview feedback. High ratings in *Attractiveness* and *Perspicuity* suggest that the robot's communication was engaging and easy to understand. Interview comments reinforced this impression, with many participants describing the interaction as *motivating*, *supportive*, and *uplifting*. The practical advice offered, especially stretching suggestions, was not only appreciated conceptually but in some cases also acted upon in real time, underlining its relevance.

Gender comparisons showed a generally more enthusiastic response from female participants, consistent with other scenarios, while male participants gave higher ratings only in *Efficiency* and *Novelty*. This may point to differences in how functional vs. experiential qualities are prioritized. The stronger positive response from BSc students, also seen in previous scenarios, suggests that simpler or more playful interactions resonate more with younger or less advanced academic audiences. In contrast, MSc participants gave the most reserved evaluations, highlighting a potential need for adaptive designs that cater to different user profiles.

In terms of visual and behavioral expression, the "happy dance" motion was broadly successful in grabbing attention and conveying energy. However, mixed opinions on the color choice, particularly the perception of orange as too aggressive, illustrate how color psychology can influence emotional interpretation. Similarly, while some participants enjoyed the playful inclusion of a riddle, others found it *cringe* or distracting, revealing the importance of calibrating humor and cognitive load in emotionally supportive systems.

Overall, the Fun response succeeded in boosting engagement and motivation, especially for participants in need of reactivation during the study. However, findings suggest that flexibility in tone, visual design, and interactivity could further improve its reception across more diverse user groups.

6.4 Joy Scenario

The Joy scenario received a highly positive reception, both in terms of user experience scores and subjective feedback. High evaluations in *Perspicuity* and *Attractiveness* confirm the robot's ability to convey its message clearly and in an appealing manner. While *Novelty* was rated slightly lower, it still fell within a neutral-to-positive range, indicating that the interaction was familiar but not perceived as repetitive or unoriginal.

Gender comparisons showed that female participants were more enthusiastic overall, particularly regarding the system's *Attractiveness*, *Perspicuity*, and *Dependability*, while male participants rated *Novelty* slightly higher. These trends suggest that women may be more receptive to emotionally expressive, praise-oriented responses, while men may be more responsive to content they find innovative.

Academic-level differences were less pronounced but still noteworthy. PhD participants showed the highest appreciation for emotional and expressive aspects of the interaction, suggesting a greater sensitivity to nuanced feedback. BSc participants valued functional aspects such as *Efficiency*, while MSc participants stood out for rating *Dependability* the highest, possibly reflecting their focus on consistency and system reliability.

Interview feedback further reinforced the UEQ data. Participants widely expressed feeling *supported*, *recognised*, and *celebrated*, with several noting the motivational value of receiving praise for accomplishments. Visual and behavioral expressions, including color, smiling facial features, and cheerful movements, were generally appreciated, although a few participants found the movement exaggerated or slow. A small number perceived the robotic voice as slightly sarcastic, highlighting how vocal tone may subtly impact emotional credibility. Opinions on message length varied, indicating that timing and pacing should be adjustable to accommodate different user preferences.

Overall, the Joy scenario effectively conveyed warmth, support, and recognition while also revealing the importance of tailoring expressive modalities to individual expectations and sensitivities.

6.5 Surprise Scenario

The Surprise scenario was positively received overall, with particularly high scores in *Perspicuity*, indicating that participants found the system clear and easy to understand. Although *Novelty* received the lowest scores among the UEQ dimensions, it remained within a neutral range, suggesting that while the response was familiar, it did not detract from the overall experience.

Female participants responded more favorably across all UEQ dimensions, especially regarding *Attractiveness*, *Efficiency*, and *Stimulation*, reflecting a generally stronger emotional engagement. Among academic levels, PhD participants consistently gave the highest ratings, except in *Stimulation*, where BSc students performed slightly better. MSc participants rated the system significantly lower across several dimensions, particularly in *Attractiveness*, *Efficiency*, *Stimulation*, and *Novelty*, suggesting that this group may have had higher expectations or a different interpretation of the emotional tone.

Interview feedback reinforced these quantitative trends. Most participants responded positively to the emotional tone, describing it as *motivating*, *supportive*, and *joyful*. Expressions such as "it evaluates positively my work and motivates me to do more" highlighted the impact of the feedback on participants' sense of accomplishment. Visual and behavioral element were also appreciated for conveying a celebratory emotion, although a few participants found them exaggerated or overly complex. A small subset of users felt the expression was either *clichéd* or not entirely genuine, highlighting the challenge of balancing enthusiasm with authenticity in emotionally expressive systems.

Overall, the Surprise scenario successfully conveyed a message of encouragement and recognition, particularly resonating with users who appreciated affirmation in their study process. However, the findings also suggest that achieving emotional resonance requires careful tuning of expressive cues to avoid perceptions of artificiality or overstatement.

6.6 Caring Scenario

The Caring scenario was positively received overall, as reflected in both UEQ scores and participant feedback. High ratings in *Perspicuity* and *Attractiveness* suggest that the robot's message was delivered clearly and in a way that participants found appealing. Although *Novelty* scored slightly lower, it still remained within a neutral-to-positive range, indicating that the scenario felt familiar but not disengaging.

In contrast to other scenarios, male participants gave slightly higher ratings, particularly in *Dependability*, suggesting they may have valued the consistency or reliability of the system's behavior. However, the difference between genders was minimal, and responses were generally favorable across both groups. Similarly, results across study levels were largely homogeneous, indicating a shared appreciation for the caring design regardless of academic background.

Interview feedback reinforced these findings. Most participants described the robot's response as *supportive*, *caring*, and *motivating*, especially in the context of stress or emotional difficulty. Quotes such as "it validated my feelings" and "sometimes you just need to hear someone believing in you to get out of the dark" highlight the importance of emotional affirmation in academic support tools.

Participants also responded well to the robot's visual and behavioral expressions. The color pink, facial elements like blushing, and gestures resembling a "little hug" contributed to a perceived sense of warmth and empathy. However, a few participants found these features overly sentimental or mismatched to the context, which underscores the subjective nature of emotional design.

The practical advice component was appreciated by most, particularly for its calming suggestions like taking a deep breath. Nevertheless, some participants wanted more specific and action-oriented recommendations, indicating that the effectiveness of caring responses may increase when emotional reassurance is balanced with tangible strategies.

Overall, the Caring scenario successfully conveyed empathy and emotional support, with only minor concerns regarding tone and depth of guidance. These findings suggest that emotionally intelligent systems may benefit from blending expressive warmth with context-aware practical advice.

6.7 Pride Scenario

The Pride scenario was received positively overall, with participants reporting high satisfaction across most UEQ dimensions. Particularly strong ratings in *Perspicuity* and *Attractiveness* suggest that the robot's message was delivered clearly and in a manner that users found engaging and emotionally rewarding. Although *Novelty* was the least appreciated dimension, it still fell within a neutral-to-positive range, implying the response was familiar but not unappealing.

Gender comparisons revealed slightly higher ratings from female participants, especially for *Novelty*, while male evaluations were also positive but generally more moderate. These differences may reflect varying preferences for expressive elements or novelty in feedback style. At the academic level, responses were largely consistent, with only minor variations: BSc students rated *Attractiveness* and *Stimulation* highest, while MSc participants reported lower satisfaction in *Perspicuity* and *Efficiency*. These trends suggest that users across educational backgrounds perceived the pride response favorably, though some nuances in usability and clarity remain.

Interview data corroborated these findings. Participants frequently described the response as *supportive*, *motivating*, and *celebratory*, highlighting its role in promoting positive self-recognition. Several users emphasized the emotional value of receiving affirming feedback after completing a task, reinforcing the scenario's intended purpose.

The expression's design was broadly appreciated, with descriptors like *cute*, *cheering*, and *encouraging*. However, a few participants found elements like blushing to be overly sentimental or not context-appropriate, while others critiqued the expression as unrealistic or exaggerated. Similar mixed feedback appeared regarding color: while green was mostly well received, some participants reported disliking it, suggesting that color preferences are highly subjective.

Response length was considered appropriate by most, though a small group would have preferred a shorter message. Overall, the Pride scenario successfully delivered emotionally affirming feedback and was well-aligned with user expectations, particularly in promoting a sense of accomplishment and recognition. Small adjustments in visual detail and pacing may further enhance its effectiveness across user profiles.

6.8 Cross Group Differences

6.8.1 Gender

Across nearly all scenarios, female participants reported higher UEQ scores than males, indicating a generally more positive reception of the robot's affective feedback except in the Caring scenario, where males' scores rose to exceed those of females, suggesting that overtly "feminine" aesthetics ("pink + heart") can reverse typical patterns. Moreover, women described the robot's responses as "cute" far more often than men, reflecting a stronger affinity for charming design elements. Females also displayed greater enthusiasm for the novel technology overall, which may have boosted their engagement and led to more favorable evaluations. In terms of voice preference, women tended to favor the male or gender-neutral "robot" voice, whereas men expressed a clear preference for the female voice. These findings highlight that gender not only modulates the intensity of emotional reception but also interacts with design cues, like color, iconography, and voice gender, underscoring the value of offering customizable expressivity profiles.

6.8.2 Academic Level

Undergraduate (BSc) students tended to be the most receptive, showing high UEQ scores across scenarios and particularly embracing dynamic, multimodal feedback without complaint. Master's (MSc) participants were more discerning, often commenting that some movements felt "too much" or that multiple cues competed for their attention. Doctoral (PhD) students were the most critical; notably, they rated the Anger scenario lowest of all groups, with comments like "it made me feel even worse." These patterns suggest that instructional level influences tolerance for sensory richness and assertive emotional tones, and that an adaptive system might benefit from tuning multimodal intensity and pacing to student seniority.

6.8.3 Region

American participants exhibited the lowest emotion-recognition accuracy overall, suggesting that some expressive cues may not align perfectly with their cultural norms and may warrant region-specific calibration. In contrast, Asian participants delivered exceptionally high UEQ scores across all scenarios, indicating strong engagement and satisfaction with the system's affective feedback. Italian and Spanish speakers (as well as native English speakers) frequently commented that the robot's speech felt "too slow," pointing to potential mismatches in speaking rate expectations. Together, these findings underscore the importance of cultural and linguistic tuning, such as adjustable speech rates or localized emotion-expression profiles, to maximize clarity and user comfort in diverse settings.

6.8.4 Voice choice

Emotional-recognition accuracy did not differ significantly by voice choice, all voices can convey affective cues clearly and reliably. However, qualitative feedback revealed that the robotic timbre was occasionally perceived as sarcastic, undermining otherwise positive messages. A few participants also noted that, while they might prefer a fully human voice, the specific natural voice we provided failed to convince them of its authenticity. These insights suggest that subtle perceptions of sincerity and tone warrant careful voice-kit selection or even dynamic voice-modulation to match user expectations.

6.9 Limitations

- Video-only evaluation. All scenarios were delivered via recorded clips rather than in-person interactions, limiting perception of depth and 3D motion cues.
- Generic scenarios. We used decontextualized academic tasks rather than domain-specific challenges, which may reduce ecological validity.
- Sample size & demographics. Particularly few PhD and Asian participants, constraining subgroup analyses and generalizability.
- Emotion-naming ambiguity. The label "Caring" proved overly broad, capturing multiple supportive behaviors and muddying the intended distinctions.
- Limited voice variety. Only three distinct voices were tested, which may not capture the full range of user preferences or nuances in emotional delivery.
- Single-response presentation. Each scenario presented only one emotional variant, preventing participants from comparing multiple expressions and evaluating relative effectiveness.

6.10 Research Questions

- **RQ1.** *What emotions should a robotic study companion express to enhance university students' study processes, motivation, and learning outcomes?*
  - Based on the theoretical foundations and design rationale presented in Section 2.3.2, and the scenario-specific mappings in Table 2.1, we selected six target emotions: Anger, Joy, Pride, Fun, Surprise, and Caring.
  - However, empirical testing revealed that Anger's assertive tone was not universally suitable; PhD students in particular rated it poorly, indicating that corrective emotional feedback may require tailoring or opt-out options for certain user profiles.
- **RQ2.** *How can these emotions be expressed most clearly and effectively through the multimodal capabilities of the RSC?*
  - A comprehensive review of facial expressions, gestures, vocal cues, and color mappings is presented in Section 2.4, where different modality combinations were evaluated for clarity and intuitiveness.
  - Chapter 3 details the design process of the emotional expressions, covering prototyping, user-centered refinements, and dedicated analyses of each modality's contribution to accurate emotion conveyance.
  - Based on the empirical results, we formulated actionable guidelines to enhance expressivity.
- **RQ3.** *To what extent do the RSC's emotional expressions accurately convey the intended emotions?*
  - Recognition results are presented in Section 5.4 and discussed in Section 6.1.
  - Participants distinguished positive vs. negative valence with over 95% accuracy, demonstrating clear valence conveyance.
  - Overall emotion recognition accuracy was 58.16%, reflecting substantial confusion among adjacent positive emotions.
  - Recognition performance varied across demographic groups (gender, academic level, region), indicating opportunities for adaptive tuning.
- **RQ4.** *Do university students appreciate the emotional support provided by the RSC?*
  - UEQ scores were consistently positive across scenarios, with the highest ratings in Perspicuity and Attractiveness.
  - Interview feedback underscored increased motivation, emotional comfort, and perceived empathy.
  - Cross-group analyses revealed demographic preferences (gender, academic level, region) that can inform personalized emotional support.

6.11 Future Work

- **Adaptive parameter tuning and personalization.** Develop algorithms to automatically adjust key expressive parameters such as color saturation, motion speed, vocal prosody, response length, voice speed, anger intensity, and overall expressivity level, based on individual user preferences and prior interaction history.
- **Real-time emotional state detection.** Integrate affective computing techniques to enable the RSC to detect and respond to students' emotional states dynamically.
- **Expression variant evaluation.** Create and test multiple variants of each emotional scenario, using different gestures, color palettes, and timing parameters, for the same target emotion, to empirically determine which combinations maximize recognition accuracy and user preference.
- **Simulation-to-hardware transition.** Move from simulated demonstrations to a fully realized physical RSC platform, validating that designed gestures, lighting, and vocal expressions perform reliably in real-world conditions.
- **Longitudinal in-situ studies.** Conduct extended deployments in real academic settings to evaluate the impact of affective feedback on student motivation, stress management, and learning outcomes over multiple weeks or semesters.

7 Conclusion

This thesis began with an empirical investigation to identify which emotions a Robot Study Companion (RSC) should express, why they matter pedagogically and affectively, and when they ought to be deployed during study sessions. Guided by a theoretical review of affective feedback in educational robotics, we selected six target emotions: Anger, Joy, Pride, Fun, Surprise, and Caring, each mapped to a specific learning or motivational function.

Building on that foundation, we designed and implemented a fully functional digital twin of the RSC in ROS 2, accompanied by two open-source packages [114]:

1. A comprehensive `rsc_description` package providing kinematic models, simulation environments, and control interfaces.

2. A modular `rsc_movements_emotions` package encapsulating coordinated color and motion for each of the six emotions.

The final videos, voice assets, and facial expression modules are hosted on my GitHub thesis repository [107].

We then evaluated the system in a video–based cross-cultural user study (N=47). Participants distinguished positive versus negative valence with over 95% accuracy and achieved an overall emotion-recognition rate of 58.16%, demonstrating moderate but meaningful efficacy of the multimodal expressions. User Experience Questionnaire (UEQ) scores were consistently positive, especially for Perspicuity and Attractiveness, while qualitative interviews highlighted both the strengths of our design and areas for refinement.

Cross-group analyses revealed systematic differences by gender, academic level, and nationality. These insights point to the necessity of adaptive, personalized expressivity profiles and culturally aware calibrations.

By releasing both the digital twin and emotion library as open-source ROS 2 packages, this work lays a robust foundation for reproducible research and rapid prototyping in affective human–robot interaction.

Together, these contributions advance both the theory and practice of emotionally intelligent learning companions, moving us closer to scalable, personalized educational robots capable of enhancing student motivation, well-being, and learning outcomes.

Acknowledgements

First and foremost, I would like to express my deepest gratitude to Estonia and to Tartu University for welcoming me so warmly and providing an environment in which I could learn, grow, and feel truly at home. This beautiful country and its people have captured my heart, and I can honestly say that I have fallen in love with every aspect of life here.

I am profoundly grateful to my supervisors, Karl and Farnaz, whose unwavering support, insightful feedback, and patient guidance have been invaluable throughout this thesis journey. I also wish to thank Matevž for his thoughtful advice and encouragement.

I would like to thank all the wonderful people I've met here in Estonia:
Valeria, for being supportive from both Estonia and Italy, but also brutally honest when I deserved it. Our beautiful "cake&gossip" nights helped me survive my first Estonian winter, unforgettable moments.
Lella, for the craziest stories and adventures imaginable, nobody ever made me laugh quite like you. Tartu will never forget you, impossible.
Tova, for the instant and perfect connection when I needed support the most. I will always say that you and I are the same person, despite you being Swedish, so taller and blonde.
The Fun Table Caterina, Fabian, Riccardo, Matteo, and Nate, thank you for the most valuable and entertaining times, I miss you all dearly. I can't wait to see you all again.
Mason, the best flatmate you could ever imagine. Your energy has the power of a hurricane, you will achieve anything without a doubt, and I will be here waiting for you in Estonia. I know I was not your favourite flatmate but I can accept a second place.
Ilaria and Bianca, I only wish I had met you earlier, you are like a sunny, blue-sky day in the midst of an Estonian winter, always making me feel warm and understood. You're part of the reasons that keep me wanting to stay in Tartu.
My favorite Latvian, for being you and for being next to me.

Ai miei amici in Italia, grazie per avermi ricordato che la mia vita non è solo a Tartu, e che le vere amicizie non si fermano di fronte a 2 500 km di distanza. Un ringraziamento speciale a Gianluca e Vise.

Infine, alla mia famiglia: grazie per la vostra incrollabile fiducia in me e per l'incoraggiamento da lontano, anche quando le mie scelte non vi erano del tutto chiare. Scusate se talvolta non sono riuscita a starvi vicino come avrei voluto e come avreste meritato: sto imparando a essere una figlia più amorevole, una sorella più attenta e un'amica migliore. Prometto che non vi deluderò.

This thesis would not have been possible without the love, support, and laughter of each and every one of you. Thank you.

Bibliography

[1] Knut Inge Fostervold, Sten Ludvigsen, and Helge I. Strømsø. "Students' time management and procrastination in the wake of the pandemic". In: (2022). DOI: 10.1080/01443410.2022.2102582.

[2] Andriana Ioannidi. "Students' time management skills during the first year of studies and the support services of higher education institutions". In: (2025). DOI: 10.26220/aca.5250.

[3] Shanta Varma and Dawn McKell. "Impact of Feedback on Isolation in Distance Education Students". In: (2023).

[4] Adriana Benevides Soares et al. "Time management in the routine of university students: results of an intervention". In: (2023). DOI: 10.22235/cp.v17i2.2845.

[5] Saira Bano, Kanwal Atif, and Syed Atif Mehdi. "Systematic review: Potential effectiveness of educational robotics for 21st century skills development in young learners". In: 29.9 (2024), pp. 11135–11153. DOI: 10.1007/s10639-023-12233-2.

[6] Hae Won Park et al. "A model-free affective reinforcement learning approach to personalization of an autonomous social robot companion for early literacy education". In: 2019. DOI: 10.1609/aaai.v33i01.3301687.

[7] Yu-Ching Lin, Vivien Lin, and Nian-Shing Chen. "Examining Learner Engagement in Robot-Assisted Language Learning among Third Graders". In: 2023.

[8] Solveig Tilden et al. "Humanoid Robots as Learning Assistants? Useability Perspectives of Grade 6 Students". In: (2024). DOI: 10.1007/s10758-024-09731-8.

[9] Farnaz Baksh, Matevž Borjan Zorec, and Karl Kruusamäe. "Open-Source Robotic Study Companion with Multimodal Human–Robot Interaction to Improve the Learning Experience of University Students". In: (2024). DOI: 10.3390/app14135644.

[10] Karolina Zawieska. "HRI: From Interaction to (Lived) Experience". In: (2023). DOI: 10.3233/FAIA220635.

[11] Oxford English Dictionary. *Robot*. Accessed: 2025-03-15. URL: https://www.oed.com.

[12] iRobot. *Roomba 205 Dust Compactor Combo Robot*. Accessed: 2025-04-07. 2025. URL: https://www.irobot.com/en_US/roomba-205-dustcompactor-combo-robot/L124020.html?_gl=1*jt4dod*_up*MQ..*_ga*MTE2MjEyODMuMTc0NDAzMDcxMA..*_ga_CX1FERKJKP*MTc0NDAzMDcxMC4xLjEuMTc0

[13] J. Chestnutt et al. "Footstep Planning for the Honda ASIMO Humanoid". In: *Proceedings of the 2005 IEEE International Conference on Robotics and Automation*. 2005.

[14] Boston Dynamics. *Spot*. Accessed: 2025-04-07. 2025. URL: https://bostondynamics.com/products/spot/.

[15] B.R. Duffy et al. "What is a Social Robot?" In: (1999). URL: https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=42f34541ae8dad880c22fe6

[16] M.M.A. De Graaf, S. Ben Allouch, and J.A.G.M. Van Dijk. "What makes robots social?: A user's perspective on characteristics for social human-robot interaction". In: (2015). DOI: 10.1007/978-3-319-25554-5_19.

[17] Kirsikka Kaipainen, Aino Ahtinen, and Aleksi Hiltunen. ""Nice surprise, more present than a machine" Experiences evoked by a social robot for guidance and edutainment at a city service point". In: 2018. DOI: 10.1145/3275116.3275137.

[18] Manja Lohse, Frank Hegel, and Britta Wrede. "Domestic applications for social robots - An online survey on the influence of appearance and capabilities". In: (2008). DOI: 10.14198/JoPha.2008.2.2.04.

[19] Yao Song and Yan Luximon. "Trust in AI agent: A systematic review of facial anthropomorphic trustworthiness for social robot design". In: (2020). DOI: 10.3390/s20185087.

[20] Thomas B Sheridan. "A review of recent research in social robotics". In: (2020). DOI: 10.1016/j.copsyc.2020.01.003.

[21] Andrea F. Abate et al. "Contextual trust model with a humanoid robot defense for attacks to smart eco-systems". In: (2020). DOI: 10.1109/ACCESS.2020.3037701.

[22] Dmitry Dereshev and David Kirk. "Form, function and etiquette–potential users' perspectives on social domestic robots". In: (2017). DOI: 10.3390/mti1020012.

[23] Deepti Mishra et al. "An Exploration of the Pepper Robot's Capabilities: Unveiling Its Potential". In: (2024). DOI: 10.3390/app14010110.

[24] Fabien Badeig et al. "A distributed architecture for interacting with NAO". In: 2015. DOI: 10.1145/2818346.2823303.

[25] A. Balaji et al. "Design And Analysis Of Improving Human – Robot Interaction By NAO Robot, Awareness Of Human Facial Emotions". In: 2023. DOI: 10.1063/5.0178082.

[26] Manoj Ramanathan, Nidhi Mishra, and Nadia Magnenat Thalmann. "Nadine Humanoid Social Robotics Platform". In: (2019). DOI: 10.1007/978-3-030-22514-8_49.

[27] Iolanda Leite et al. "The influence of empathy in human-robot relations". In: (2013). DOI: 10.1016/j.ijhcs.2012.09.005.

[28] Paolo Dario et al. "Robot companions for citizens". In: 2011. DOI: 10.1016/j.procs.2011.12.017.

[29] In Ho Han et al. "Human-Robot Interaction and Social Robot: The Emerging Field of Healthcare Robotics and Current and Future Perspectives for Spinal Care". In: (2024). DOI: 10.14245/ns.2448432.216.

[30] Jiaqi Fang et al. "The Role of Social Robots in Alleviating Anxiety and Enhancing Mental Well-Being". In: (2025). DOI: 10.1007/s11469-025-01465-7.

[31] Antonio Vitale and Umberto Dello Iacono. "Using social robots as inclusive educational technology for mathematics learning through storytelling; [Usando robots sociales como tecnología educativa inclusiva para el aprendizaje de matemáticas a través de la narración]". In: (2024). DOI: 10.31637/epsir-2024-672.

[32] Juho Hamari Eshtiak Ahmed O˘guz 'Oz' Buruk. "Human–Robot Companionship: Current Trends and Future Agenda". In: (2024). DOI: 10.1007/s12369-024-01160-y.

[33] Naoto Yoshida et al. "Production of Character Animation in a Home Robot: A Case Study of LOVOT". In: (2022). DOI: 10.1007/s12369-021-00746-0.

[34] Muhammad Winal Zikril Zulkifli, Syamimi Shamsuddin, and Lim Thiam Hwee. "Survey on animal robot PARO in Malaysia: Perception and acceptance". In: (2018). DOI: 10.1007/978-981-10-8788-2_19.

[35] Kaoru Inoue, Kazuyoshi Wada, and Yuko Ito. "Effective application of paro: Seal type robots for disabled people in according to ideas of occupational therapists". In: (2008). DOI: 10.1007/978-3-540-70540-6_197.

[36] Cambridge Dictionary. *Emotion*. Accessed: 2025-03-16. URL: https://dictionary.cambridge.org.

[37] Klaus R. Scherer. *What are emotions? And how can they be measured?* Klaus R. Scherer, pp. 695–729. DOI: 10.1177/05390184050582.

[38] Ali Mahdi et al. "EmoGo: A Smart Wearable IoT System for Human Emotion Detection". In: 2022. DOI: 10.1109/URC58160.2022.10054216.

[39] Eva Hudlicka and Marjorie McShane. *Emotions in affective human-computer interaction*. 2023. DOI: 10.1515/9783110795486-041.

[40] Chandrasekar Vuppalapati et al. "Emotional health: A data driven approach to understand our emotions and improve our health". In: 2019. DOI: 10.1109/CSE/EUC.2019.00071.

[41] Seonjeong Ally Lee et al. "Does Consumers' Feeling Affect Their Quality of Life? Roles of Consumption Emotion and Its Consequences". In: (2015). DOI: 10.1002/jtr.1988.

[42] Maital Neta and Ingrid J. Haas. "Movere: Characterizing the role of emotion and motivation in shaping human behavior". In: (2019). DOI: 10.1007/978-3-030-27473-3_1.

[43] Ahmad Hammoudeh and Sara Tedmori. "Emotions and the Structure of the Language". In: 2019. DOI: 10.1109/JEEIT.2019.8717526.

[44] Dongho Koo et al. "Immersive Emotions: Translating Emotions into Visualization". In: 2022. DOI: 10.1145/3528575.3551430.

[45] Markus Haring, Nikolaus Bee, and Elisabeth Andre. "Creation and Evaluation of emotion expression with body movement, sound and eye color for humanoid robots". In: 2011. DOI: 10.1109/ROMAN.2011.6005263.

[46] Satoshi Yagi et al. "Can an android's posture and movement discriminate against the ambiguous emotion perceived from its facial expressions?" In: (2021). DOI: 10.1371/journal.pone.0254905.

[47] Nina Lvovna Sungurova Natalya Borisovna Karabuschenko* Aleksandr Vasilevich Ivashchenko and Ekaterina Mihailovna Hvorova. "Emotion Recognition in Different Cultures". In: (2016). DOI: 10.17485/ijst/2016/v9i48/109085.

[48] Xiaohui Liu Shakirat Abimbola Adesola Yongmin Li. "Effect of Emotions on Students Learning Strategies". In: (2019). DOI: 10.1145/3318396.3318408.

[49] Sidney D'Mello and Art Graesser. "Malleability of students' perceptions of an affect-sensitive tutor and its influence on learning". In: 2012.

[50] Yizhang Jiang y Jing Tan Jie Mao and Ming Gao. "The Influence of Academic Emotions on Learning Effects: A Systematic Review". In: (2021). DOI: 10.3390/ijerph18189678.

[51] Evert A. Van Doorn, Gerben A. Van Kleef, and Joop Van Der Pligt. "How instructors' emotional expressions shape students' learning performance: The roles of anger, happiness, and regulatory focus". In: (2014). DOI: 10.1037/a0035226.

[52] Facundo Froment and Manuel de-Besa Gutiérrez. "The prediction of teacher credibility on student motivation: Academic engagement and satisfaction as mediating variables; [La predicción de la credibilidad docente sobre la motivación de los estudiantes: el compromiso y la satisfacción académica como variables mediadoras]". In: (2022). DOI: 10.1016/j.psicod.2022.04.003.

[53] Junjie He et al. "A study on the effect of joyful learning application upon undergraduate English vocabulary learning". In: 2017. DOI: 10.1109/EITT.2017.76.

[54] Saba Hasanzadeh, Shaghayegh Shayesteh, and Reza Pishghadam. "Investigating the role of teacher concern in EFL students' motivation, anxiety, and language achievement through the lens of self-determination theory". In: (2024). DOI: 10.1016/j.lmot.2024.101992.

[55] Dinesh Chandra Agrawal et al. "Factors affecting student-teacher relationship in a private university of technology in Taiwan". In: (2019).

[56] Mark Loon and Robin Bell. "The moderating effects of emotions on cognitive skills". In: (2018). DOI: 10.1080/0309877X.2017.1311992.

[57] Mara Marini et al. ""I feel good with my teachers". The effects of positive teacher-student relationship on students' self-esteem and perceptions about their future". In: (2023). DOI: 10.13133/2724-2943/17983.

[58] Dovan Rai et al. "Repairing deactivating negative emotions with student progress pages". In: (2013). DOI: 10.1007/978-3-642-39112-5_115.

[59] Tomohiro Nagashima et al. "Designing Playful Intelligent Tutoring Software to Support Engaging and Effective Algebra Learning". In: (2022). DOI: 10.1007/978-3-031-16290-9_19.

[60] Theo Wubbels et al. *Teacher–Student Relationships and Student Achievement*. 2016.

[61] Hamideh Iraj et al. "Understanding Students' engagement with personalised feedback messages". In: 2020. DOI: 10.1145/3375462.3375527.

[62] Gerben A. van Kleef. "The interpersonal dynamics of emotion: Toward an integrative theory of emotions as social information". In: (2016). DOI: 10.1017/CBO9781107261396.

[63] Taehyun Kim, Vishesh Kumar, and Mike Tissenbaum. "Productive Anger? Changing Systems Understanding due to Negative Emotions". In: 2021.

[64] Ali Derakhshan and Lawrence Jun Zhang. "Introduction to the Special Issue: New Insights into the Study of Classroom Emotions: Emerging Research Methods for Exploring the Implications of Positive and Negative Emotions in Language Education Environments". In: (2024). DOI: 10.30466/ijltr.2024.121574.

[65] Joseph P. Mazer. "Students' Discrete Emotional Responses in the Classroom: Unraveling Relationships With Interest and Engagement". In: (2017). DOI: 10.1080/08824096.2017.1365233.

[66] Mehdi Ghayoumi and Arvind K. Bansal. “Emotion in robots using convolutional neural networks”. In: (2016). DOI: 10.1007/978-3-319-47437-3_28.

[67] Zhen-Tao Liu et al. “A multimodal emotional communication based humans-robots interaction system”. In: 2016. DOI: 10.1109/ChiCC.2016.7554357.

[68] Pedro Cárdenas et al. “Evaluation of Robot Emotion Expressions for Human–Robot Interaction”. In: (2024). DOI: 10.1007/s12369-024-01167-5.

[69] Shujie Zhou and Leimin Tian. “Would you help a sad robot? Influence of robots’ emotional expressions on human-multi-robot collaboration”. In: 2020. DOI: 10.1109/RO-MAN47096.2020.9223524.

[70] Temirlan Dzhoroev et al. “Human Perception on Social Robot’s Face and Color Expression Using Computational Emotion Model”. In: 2023. DOI: 10.1109/RO-MAN57019.2023.10309452.

[71] Tokitomo Ariyoshi, Kazuhiro Nakadai, and Hiroshi Tsujino. “Effect of facial colors on humanoids in emotion recognition using speech”. In: 2004.

[72] Jared Suttles and Nancy Ide. “Distant Supervision for Emotion Classification with Discrete Binary Values”. In: *Computational Linguistics and Intelligent Text Processing*. Springer Berlin Heidelberg, 2013, pp. 121–136. DOI: 10.1007/978-3-642-37256-8_11.

[73] Melissa Donaldson. “Plutchik’s wheel of emotions-2017 update”. In: *Six-Seconds. Retrieved from https://www. 6seconds. org/2017/04/27/plutchiksmodel-of-emotions* (2017).

[74] Kazunori Terada, Atsushi Yamauchi, and Akira Ito. “Artificial emotion expression for a robot by dynamic color change”. In: *2012 IEEE RO-MAN: The 21st IEEE International Symposium on Robot and Human Interactive Communication*. 2012, pp. 314–321. DOI: 10.1109/ROMAN.2012.6343772.

[75] Faeze Heydari, Majid Khalili-Ardali, and Ali Yoonessi. “Shades of Feeling: How Facial Color Variations Influence Emotional and Health Perception”. In: (2024). DOI: 10.1002/col.22968.

[76] Steven G. Young, Christopher A. Thorstenson, and Adam D. Pazda. “Facial redness, expression, and masculinity influence perceptions of anger and health”. In: (2018). DOI: 10.1080/02699931.2016.1273201.

[77] Hee Yeon Im et al. “Sex-related differences in behavioral and amygdalar responses to compound facial threat cues”. In: (2018). DOI: 10.1002/hbm.24035.

[78] Tetsuto Minami, Kae Nakajima, and Shigeki Nakauchi. “Effects of face and background color on facial expression perception”. In: (2018). DOI: 10.3389/fpsyg.2018.01012.

[79] Sandrine Gil and Ludovic Le Bigot. “Development of the Red-Negative Association: Motivation-based behaviors”. In: (2017). DOI: 10.1016/j.cogdev.2017.09.004.

[80] J.H. Xin et al. “Cross-regional comparison of colour emotions Part II: Qualitative analysis”. In: (2004). DOI: 10.1002/col.20063.

[81] Tomomi Onuki et al. “Designing robot eyes and head and their motions for gaze communication”. In: (2014). DOI: 10.1007/978-3-319-09333-8_66.

[82] Hyunsoo Song et al. “Design of a robot head for emotional expression: EEEX”. In: 2008. DOI: 10.1109/ROMAN.2008.4600667.

[83] J. P. Rodriguez Gomez et al. “An emoticon is well worth a few empathetic words”. In: *Proceedings of the 33rd Pacific Asia Conference on Language, Information and Computation, PACLIC 2019*. 2019, pp. 212–218.

[84] I. Leite, A. Paiva, and T. Ribeiro. “Emotion Modelling for Social Robots”. In: *The Oxford Handbook of Affective Computing*. Ed. by R. A. Calvo et al. 2014, pp. 296–308.

[85] Ehud Sharlin James E. Young Min Xin. “Robot expressionism through cartooning”. In: (2007). DOI: 10.1145/1228716.1228758.

[86] Gabriele Trovato et al. “Cross-Cultural Perspectives on Emotion Expressive Humanoid Robotic Head: Recognition of Facial Expressions and Symbols”. In: (2013). DOI: 10.1007/s12369-013-0213-z.

[87] Hui Jiang et al. “Virtual Characters Meet the Uncanny Valley: A Literature Review Based on the Web of Science Core Collection (2007-2022)”. In: 2022. DOI: 10.1109/CoST57098.2022.00088.

[88] Melanie L. Glocker et al. “Baby schema in infant faces induces cuteness perception and motivation for caretaking in adults”. In: (2009). DOI: 10.1111/j.1439-0310.2008.01603.x.

[89] Chien-Hsiung Chen and Xiaoyu Jia. “Research on the influence of the baby schema effect on the cuteness and trustworthiness of social robot faces”. In: (2023). DOI: 10.1177/17298806231168486.

[90] Melanie L. Glocker et al. “Baby schema modulates the brain reward system in nulliparous women”. In: (2009). DOI: 10.1073/pnas.0811620106.

[91] Lizhu Luo et al. “Neural systems and hormones mediating attraction to infant and child faces”. In: (2015). DOI: 10.3389/fpsyg.2015.00970.

[92] Fabian Löwenbrück and Ursula Hess. “Not all “caregivers” are created equal: Liking, caring and facial expression responses to the baby schema as a function of parenthood and testosterone”. In: (2021). DOI: 10.1016/j.biopsycho.2021.108120.

[93] Genta Yoshioka and Yugo Takeuchi. “Inferring affective states by involving simple robot movements”. In: 2015. DOI: 10.1145/2814940.2814959.

[94] Kazuhiko Takahashi, Mana Hosokawa, and Masafumi Hashimoto. “Remarks on designing of emotional movement for simple communication robot”. In: 2010. DOI: 10.1109/ICIT.2010.5472735.

[95] Jekaterina Novikova and Leon Watts. “A design model of emotional body expressions in non-humanoid robots”. In: 2014. DOI: 10.1145/2658861.2658892.

[96] Helen Andreae et al. “A study of Auti: A socially assistive robotic toy”. In: 2014. DOI: 10.1145/2593968.2610463.

[97] Martin Cooney and Anita Sant’Anna. “Avoiding Playfulness Gone Wrong: Exploring Multi-objective Reaching Motion Generation in a Social Robot”. In: (2017). DOI: 10.1007/s12369-017-0411-1.

[98] Roshni Kaushik and Reid Simmons. “Perception of emotion in torso and arm movements on humanoid robot quori”. In: 2021. DOI: 10.1145/3434074.3447129.

[99] Thi Le Quyen Dang et al. “Encoding cultures in robot emotion representation”. In: 2017. DOI: 10.1109/ROMAN.2017.8172356.

[100] Joe Crumpton and Cindy L. Bethel. "Validation of vocal prosody modifications to communicate emotion in robot speech". In: 2015. DOI: 10.1109/CTS.2015.7210396.

[101] Andreea Niculescu et al. "Making Social Robots More Attractive: The Effects of Voice Pitch, Humor and Empathy". In: (2013). DOI: 10.1007/s12369-012-0171-x.

[102] S. Hennig and R. Chellali. "Expressive synthetic voices: Considerations for human robot interaction". In: 2012. DOI: 10.1109/ROMAN.2012.6343815.

[103] J. James, C. I. Watson, and B. MacDonald. "Artificial Empathy in Social Robots: An analysis of Emotions in Speech". In: *2018 27th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*. 2018, pp. 632–637. DOI: 10.1109/ROMAN.2018.8525652.

[104] Akihiro Tanaka et al. "I feel your voice: Cultural differences in the multisensory perception of emotion". In: (2010). DOI: 10.1177/0956797610380698.

[105] Ilaria Torre and Sebastien Le Maguer. "Should robots have accents?" In: 2020. DOI: 10.1109/RO-MAN47096.2020.9223599.

[106] Inc. Sketchbook. *Sketchbook - For everyone who loves to draw*. 2021. URL: https://www.sketchbook.com/.

[107] Miriam Calafà. *miriam_thesis: Robot Study Companion – Designing Multimodal Emotional Expressions*. https://github.com/RobotStudyCompanion/miriam_thesis/tree/main. 2025.

[108] URL: https://veo2.ai/.

[109] EmmaMultilingual. *Voice Generator*. Accessed: 2025-04-08. URL: https://voicegenerator.io/.

[110] Raymond and Betty. *Narakeet*. Accessed: 2025-04-08. URL: https://www.narakeet.com/app/text-to-audio/?projectId=8bd93478-80ac-439d-b142-21ba63c91e9b.

[111] *ROS 2 Humble*. Accessed: 2025-04-08. 2025. URL: https://docs.ros.org/en/humble/index.html.

[112] *RViz 2 User Guide*. Accessed: 2025-04-08. 2025. URL: https://docs.ros.org/en/rolling/Tutorials/Intermediate/RViz/RViz-User-Guide/RViz-User-Guide.html.

[113] Farnaz Baksh. *MechDesign: Mechanical Design Files for the Robot Study Companion (RSC)*. https://github.com/RobotStudyCompanion/MechDesign.

[114] Miriam Calafà. *RSC_ROS2: Robot Study Companion – ROS 2 Simulation and Emotional Expression System*. https://github.com/RobotStudyCompanion/RSC_ROS2/tree/main. 2025.

[115] *Canva - Design, Share, and Print Business Cards, Presentations, Social Media Graphics, and More*. Accessed: 2025-04-08. URL: https://www.canva.com/.

[116] *User Experience Questionnaire (UEQ)*. Accessed: 2025-04-08. URL: https://www.ueq-online.org/.

Non-exclusive licence to reproduce thesis and make thesis public

I, Miriam Calafa'

1. herewith grant the University of Tartu a free permit (non-exclusive licence) to reproduce, for the purpose of preservation, including for adding to the DSpace digital archives until the expiry of the term of copyright,

   "Designing Multimodal Emotional Expression for a Robotic Study Companion"

   supervised by Karl Kruusamäe, Farnaz Baksh

2. I grant the University of Tartu a permit to make the work specified in p. 1 available to the public via the web environment of the University of Tartu, including via the DSpace digital archives, under the Creative Commons licence CC BY NC ND 3.0, which allows, by giving appropriate credit to the author, to reproduce, distribute the work and communicate it to the public, and prohibits the creation of derivative works and any commercial use of the work until the expiry of the term of copyright.

3. I am aware of the fact that the author retains the rights specified in p. 1 and 2.

4. I certify that granting the non-exclusive licence does not infringe other persons' intellectual property rights or rights arising from the personal data protection legislation.

Miriam Calafa'
20.05.2025