Emotion and Mood Blending in Embodied Artificial Agents: Expressing Affective States in the Mini Social Robot

Enrique Fernández-Rodicio[1] · Marcos Maroto-Gómez[1] · Álvaro Castro-González[1] · María Malfaz[1] ·
Miguel Á. Salichs[1]

Abstract

Robots that are devised for assisting and interacting with humans are becoming fundamental in many applications, including in healthcare, education, and entertainment. For these robots, the capacity to exhibit affective states plays a crucial role in creating emotional bonding with the user. In this work, we present an affective architecture that grounds biological foundations to shape the affective state of the Mini social robot in terms of mood and emotion blending. The affective state depends upon the perception of stimuli in the environment, which influence how the robot behaves and affectively communicates with other peers. According to research in neuroscience, mood typically rules our affective state in the long run, while emotions do it in the short term, although both processes can overlap. Consequently, the model that is presented in this manuscript deals with emotion and mood blending towards expressing the robot's internal state to the users. Thus, the primary novelty of our affective model is the expression of: (i) mood, (ii) punctual emotional reactions to stimuli, and (iii) the decay that mood and emotion undergo with time. The system evaluation explored whether users can correctly perceive the mood and emotions that the robot is expressing. In an online survey, users evaluated the robot's expressions showing different moods and emotions. The results reveal that users could correctly perceive the robot's mood and emotion. However, emotions were more easily recognized, probably because they are more intense affective states and mainly arise as a stimuli reaction. To conclude the manuscript, a case study shows how our model modulates Mini's expressiveness depending on its affective state during a human-robot interaction scenario.

Keywords Social robotics · Affective state · Emotion · Mood · Emotional expression · Human-robot interaction

1 Introduction

In the last few decades, researchers in artificial life have discussed whether it is necessary, or even moral, to endow embodied artificial agents with affective capabilities [4,5,46]. Affect is a term that typically involves two different events: mood and emotion. While these terms are usually exchangeable, many authors agree that emotion and mood, although related, represent different concepts [1]. In particular, the results of the study conducted by Beedie et al. [1] showed that their main differences rely on the duration, cause, or intentionality of the event. Numerous authors, such as Frijda, have studied affective generation in humans [22].

Other classical studies—such as Russel's circumplex model of affect [51], Ekman's studies about the nature of emotions [13–15], or Plutchik's psycho evolutionary theory of emotions [48]—provide significant insights for modelling affective states. These authors agree that affect is a critical component of human behaviour and it plays an important role in fundamental human processes, such as social interaction [41] or ethical decision-making [55]. Similarly, relevant research reviews the role of emotions in robotics and

✉ Enrique Fernández-Rodicio
enrifern@ing.uc3m.es

Marcos Maroto-Gómez
marmarot@ing.uc3m.es

Álvaro Castro-González
acgonzal@ing.uc3m.es

María Malfaz
mmalfaz@ing.uc3m.es

Miguel Á. Salichs
salichs@ing.uc3m.es

1   Systems Engineering and Automation Department, University Carlos III of Madrid, Av de la Universidad, 30, Leganés, 28911 Madrid, Spain

how emotions should be included in artificial systems [11]. However, despite the previous progress in human behaviour modelling, how emotional processes are evoked in human brains remains unclear. Consequently, endowing robots with these capabilities is still an arduous task [20]. As the authors conclude in [35], expressing emotion and mood is imperative for an interaction to be natural among humans and robots. These cases show the importance of endowing embodied artificial agents, such as social robots, with the ability to generate and express affective states.

In this work, we present the design and development of an affective architecture for the Mini social robot to enable it to show affective expressiveness while interacting with people. These expressions use different modulation profiles, which define the effect that a particular affective state has over the robot's interfaces. The proposed system considers mood and emotion blending, including their temporal dimension. Emotions and moods are differentiated from each other in intensity and duration. On the one hand, emotions are intense, short-lived impulsive affective episodes that appear as a response to stimuli. Their effects rapidly disappear when the stimuli are no longer perceived. On the other hand, mood acts as a baseline affective state with a longer-lasting duration and influence on the robot [36]. Our model generates the robot's affective state considering the elicitation of emotions and mood as a response to environmental stimuli. It also shapes how their intensity decays with time once stimuli are no longer perceived, and how emotion and mood blend[26].

Our architecture allows mood and emotion blending. This enables the activation of one mood and several emotions at the same time. Nonetheless, because Mini can only show one expression at a time, we define Mini's affective state as the combination of the emotion with the highest intensity level (dominant emotion) and its mood. The expressive generation of emotion and mood is predefined and stored in profiles. Because we consider emotions to be short-lived, intense experiences, our model prioritises the expression of the dominant emotion above mood, using its intensity to define the weight that the emotion should have over the robot's expressiveness. Finally, our architecture is also able to react to the stimulus that elicited a particular emotion through the execution of emotional expressions that have been handcrafted for each possible stimulus. For example, if the performance of the user during a game triggers the elicitation of the *joy* emotion, then the robot will perform an expression that expresses happiness while acknowledging the user's excellent work. The proposed architecture was validated using an online survey where the participants had to identify individual mood and emotion expressions in different videos. Once our expressions were validated, we designed an interaction scenario where Mini expressed its affective state addressing mood and emotion blending, and their evolution with time.

This manuscript continues in Section 2 with a review of the recent literature about affective generation in humans and expressiveness in social robots. In Sect. 3, we describe how the model generates the robot's affective state considering important studies in neuroscience, and we also describe how we deal with emotion and mood blending. Section 4 describes how the robot expressiveness translates affective states into actuation commands that transmit the robot's affective state to the user. Section 5 presents the experimental setup and the evaluation process, which were carried out to assess whether or not the participants correctly perceived the affective cues that the robot is expressing. Section 6 shows a case of use of the operation of the system during human-robot interaction, where the robot and a user interact by playing a quiz game. The robot reacts to the user's behaviour and the game's dynamics during the interaction by modulating its expressiveness. In Sect. 7, we present the results that were obtained during the evaluation of the experiment described in Sect. 5. Next, in Sect. 8, we discuss the most notable outcomes of the evaluation. Finally, Sect. 9 summarizes the novelties derived from this study.

2 Related Work

Endowing robots with the capacity to generate and express affect has attracted much interest in human-robot interaction. Furthermore, the study of how affective states should be modelled and expressed by robots has gained significant attention over the years. In 2014, Paiva et al. [45] presented a review of the advances made in emotion modelling. At the core of this review is the concept of the *affective loop*, where the expression of emotions by the user leads to the elicitation of emotions in the agent, which will in turn affect the user's affective state. This review also surveys the different types of emotional architectures that can be found, regarding to the inspiration they take, which affective states they model, how they integrate emotion and cognition, and what expressiveness capabilities they integrate. In the area of expressiveness, the authors highlight how inspiration has been taken traditionally from the world of arts and (more specifically) animation. This work finishes with an analysis of how to achieve robots that have empathic capabilities and a review of the challenges that should be tackled in the future. In 2017, M.F. Jung [33] reviewed the perspectives on emotion in HRI. According to this work, emotion and emotion regulation have to be understood in the context of interactions between participants, and not just as something that happens within and on-top of individuals. To achieve this, it is important that the participants in the interaction achieve affective grounding, which is a common framework for interpreting and responding to behaviour from an emotion point of view. This means that robots have to be endowed with the

ability to take part in this affect coordination. In 2021, Yan et al. [65] conducted a survey of the available literature in the areas of emotion classification, emotional robots, and emotion space modelling for Human-Robot Interaction. Based on their review, the authors presented a few insights about the direction that research in these areas should take. They proposed that the use of techniques extracted from areas such as data mining or machine learning could help with the task of discriminating the different types of emotions. Other findings from this review are: (i) the task of emotion recognition needs improvement; (ii) physiological and non-physiological data may be combined with the use of advanced control theory and simulations of the human cerebral cortex and neural system; (iii) effective HRI ingrained with robust emotion space models and standards should be considered; and (iv) the study of the relationship between emotion and cognition should be intensified.

In this section, we will present an evaluation of some of the approaches for generating and expressing affective states in social robots. This review will be divided into two parts: the first part includes authors that focused mainly on emotions, while the second part includes authors that also used mood (either as a state that can be expressed or as a variable that affects emotion generation).

2.1 Generation and Expression of Emotion in Robots

Among the many different affective states that a robot can express, emotion has attracted more attention in research. Some authors have focused more on the process of emotion generation, while others paid more attention to how robots can express emotions. Following the first approach, in 2018 Correia et al. [6] presented a model for group-based emotions (i.e., emotions that result from an event related to the social group where the individual is integrated) in social robotic partners. The system maintains the context of the interaction and the social groups that are present, elicits emotions based on events happening in the current social context, and decides how to express these emotions. The novelty of this work is emotion generation, which can be elicited based on the individual robot actions or the whole social group. The authors evaluated this approach with an experiment that included 48 participants, where two persons would team with two robots to play a collaborative card game. One group of the participants teamed with a robot that generated emotions based on group actions, while the other group teamed with a robot that generated emotions based on its actions. The results show that participants playing with the robot generating group-based emotions showed a more robust group identification and perceived their partner as more likeable. The participants reported a higher degree of group trust towards the robot that has group-based emotions. In the same year, Javed and Park [32] presented the design of a custom animated

character, in the shape of a penguin, with emotional expressiveness capabilities, including an algorithm for regulating a user's emotions through empathic interaction. This character was designed to interact with children with Autism Spectrum Disorder. The animated penguin was integrated in an iPod-based mobile platform and was able to use facial expressions, motions, colour, and sounds. The interaction framework at the core of this robotic platform was consensus-based, and includes three agents: the human user, the robotic platform, and an emotion goal state that the agent tries to steer the user towards. Emotions are modelled in a 2D circumplex model. Under this model, the penguin's emotional state is computed as a function of the user's state and the emotion goal. The authors validated their approach through a user study where the participants played an emotion game. The results of this study showed that the proposed framework was able to regulate the emotions of the participants in the study appropriately, and that longer interactions with the robot led to higher levels of engagement.

As mentioned earlier, other authors have focused on how to express emotions using robots. In 2013, Yilmazyildiz et al. [66] presented an emotional interface for the Probo social robot. In their model, emotions are represented by a vector inside a 2D space (valence-arousal) that is mapped to the robot's degrees of freedom. Emotions are expressed using facial expressions and affective gibberish speech (i.e., vocalisation of meaningless strings of speech). In total, 35 children evaluated the system by reviewing videos of the facial expressions, audio clips of speech, and both modalities combined in the context of the robot telling a story with pictures. The participants then matched each clip with the perceived emotion (i.e., anger, disgust, fear, joy, sadness, and surprise). The results show that children recognised the emotions more easily using the audio clips than using the videos. However, the combination of video and audio yielded the best outcomes.

In 2015, Cameron et al. [3] studied how life-like facial expressions in a humanoid robot affected children's behaviour and attitude towards the robot. The participants played the Simon Says game with the robot. The results showed significant effects when dividing the participant's answers according to gender. In particular, female participants gave lower ratings for the extent to which the robot liked them. Male participants in the expressive condition reported more enjoyment than those in the non-expressive condition, while the opposite happened with the female participants. In the same year, Bretan et al. [2] presented an emotional system for robots that lack facial expression and complex humanoid design. This system uses a series of parameters and mathematical functions that were derived from Darwin's research to continuously generate movements that fit different emotions and intensities. They tested how this system could express emotions using predefined postures and gestures, and generated emotional motions with random parameters. The

participants reviewed either static postures or gestures, in person or through a video. They then labelled each gesture or pose with one of the six emotions, a valence, and rated how well it represents the emotion. The results show that dynamic gestures are better at transmitting emotion and that there are no significant differences between face to face evaluations and using videos for the postures, although there were differences in the ratings of the gestures. The participants showed more problems identifying the postures of fear, happiness, and surprise. The results for the evaluation of generated gestures suggest that specific characteristics of pose and motion are tied to particular emotions.

Gácsi et al. [27] studied if a robot could express emotions using basic behaviours inspired by animal behaviours. In particular, they focused on expressive dog behaviours. In total, 78 participants watched five videos of the robot and five videos of a dog performing emotional behaviours showing joy, sadness, anger, fear, and a neutral state. They then described these clips on an open-ended and multiple-choice questionnaire. The robot used movements and predefined sounds to execute the behaviours. Their results show that the participants tended to attribute emotions to both the robot and the dog in open-ended questions, and they selected the correct emotion in multiple-choice questions. Feldmaier et al. [19] evaluated the possibility of using colours and dynamic light patterns through an RGB-LED display to convey four emotions: happiness, sadness, anger, and fear. The authors defined a specific pattern for each emotion combining colours: a fire pattern for anger, a rain pattern for sadness, a rainbow pattern for happiness, and a purple pattern for fear. The authors used a two-wheeled robot that displayed LED patterns. The robot moved using emotion-specific motion patterns. In the evaluation, the participants observed videos of the robot for each emotion, and they then evaluated the level of excitement and pleasure. Their results show that there was a significant difference in the excitement perceived by the participants for the emotions with high arousal (happiness and anger) and low arousal (sadness and fear). On the pleasure axis, only anger shows a significant difference.

In line with Feldmaier et al.'s research, Song et al. [56] studied the use of colour, sound, and vibration to express emotions. They developed a series of expressions that used one or several modalities to convey happiness, anger, sadness, and relaxation. Their approach was evaluated with a user study, where the participants observed the robot performing different expressions and then selected which emotion they thought the robot was expressing in each case. Their results show that sadness and anger were the emotions that the participants perceived more easily, while none of the expressions designed to express happiness was correctly perceived. Their results also show that colour is the essential modality for communicating affective states, while sound and vibration showed bias in particular emotions. Löffler et al. [39]

also focused on displaying artificial emotions using colour, motion, and sound. Their goal was to quantify the information content of each modality and find how they could be combined more effectively. Several light patterns, beeping sounds, and motions were designed for each emotion (i.e., joy, sadness, fear, and anger). The expressions were first evaluated through an online survey, and the highest-rated emotion and modality were used in a second study. The participants watched the robot performing 28 of these expressions and they then selected the most appropriate emotion, and rated their confidence on that answer. The results show that anger and joy are better expressed through colour, sadness through sound, and motion is preferred to display fear. The results also show that multimodal expressions (two or more modalities) present higher classification accuracy and confidence. Motion is the modality that performs the best, while the combination of motion and colour is the best multimodal alternative. Finally, the effectiveness of each modality depends on the emotion that is expressed.

In 2018, Tuyen et al. [61] proposed an incremental learning model to select the user's representative emotional expressions based on the user's cultural traits. First, the robot clusters human emotional expression samples, affected by their cultural background, based on the similarity of the movements. The system then selects an expression from the most significant cluster and maps it into the robot's motion space. The authors evaluated the behaviour selection model through long-term interactions. In this experiment, the robot interacted with users, recorded their bodily expressions, and estimated their emotions from facial information. An online survey was conducted with 30 participants, who watched the expressions and then matched them with the corresponding emotion, and assigned appropriate arousal and valence values. Their results show that the participants could correctly identify the expressions of happiness, while they had more problems recognising sadness. All of the participants correctly assigned the happy gestures high levels of valence and arousal, while most of them correctly assigned the "sad" gestures to low levels of arousal and valence.

In 2020, Suguitan et al. [57] proposed a method for modulating affective expressions using a neural network. Their system uses a variational autoencoder with an emotion classification to adapt the movements of the robot. The autoencoder compresses the original movement into a latent embedding space. The arousal and valence of the movement are then modified in this latent space. Finally, the autoencoder decodes the latent representation of the new movement. An online survey was conducted to evaluate the subjective effectiveness of this work. Five samples of movement were extracted and then modulated into the other two emotions for each of the three emotions considered (i.e., anger, happiness, and sadness). The participants watched 30 movements that were randomly extracted from the resulting dataset, rated

how well they exhibited each emotion and then selected the emotion that better described each movement. Their results show that although there were no significant differences in the participants' ratings for how well the modified and original gestures could transmit the target emotion, differences did appear for the recognition accuracy with some modifications.

Finally, some studies have combined both the generative and the expressiveness aspects of affect in the robot. Tielman et al. [59] proposed a model for adaptive emotional expression. Their model uses information from the environment and the user's emotional state to infer the model's internal parameters value. These parameters affect the colour of the robot's eyes, the volume of the voice, and the type and size of the gestures executed by the robot. Additionally, static poses were developed for each emotion. In total, 19 children carried out an experiment playing a quiz game with two NAO robots, one group with the model and the other group without. The researchers evaluated the children's expressions during the interaction and their opinion after completing the session. Their results show that children express more positive expressions when interacting with the affective robot, although there were no significant differences between the likeability ratings given by the children.

Hong et al. [29] presented a multimodal emotional interaction architecture for social robots. This architecture is composed of three subsystems: i) the multimodal affect recognition subsystem determines the user's affective state from body language and vocal signals; ii) the robot emotion model determines the robot's deliberative emotion according to the user's state, the robot's desires, drives and previously displayed emotion, and its reactive emotion from the information retrieved from the touch sensors and the 2D camera; and iii) the interaction activity subsystem determines the most appropriate behaviour according to the task at hand. The robot selects between the deliberative and the reactive emotions depending on its priority. The architecture then generates a dynamic combination of different interaction modalities based on the emotion selected. An experiment was performed to investigate the impact on the interaction and the user's experience with the robot. The participants interacted with either a robot that used the proposed system to display emotions or one that did not. The results show that, on average, users rated the robot's valence when displaying emotions more highly. Participants in that condition also found the interaction more pleasant, and they found the emotions to be real and understandable. Regarding the usefulness of the robot, both the expressive and neutral robots displayed similar results. Finally, the participants ranked vocal intonation as the dominant modality, followed by body language and eye colours.

2.2 Generation and Expression of Moods in Robots

Emotions are the most common affective state that is implemented in robotic platforms, but they are not the only one. Some authors have also focused on endowing robots with mood expressiveness. In 2013, Han et al. [28] presented a method based on mood transition for autonomous emotional human-robot interaction. Their model recognises the user's emotional state and adapts to it. First, the system extracts the user's state by analysing their face, updates the robot's mood according to the user's state and the robot's personality, and finally generates an appropriate facial expression. The authors evaluated the model using the user's emotional state recognition and the response of an artificial face when using different personalities and moods. They then assessed the system in three different robots using a questionnaire. Robots A and B had opposite personalities, whereas robot C did not exhibit any personality and its emotional response depended on the user's perceived state. The participants watched videos of a person interacting with the three robots and then evaluated the interaction through a questionnaire. The participants rated interactions with the robots with a personality as more natural. The results also show that their mood transition model enables the robot to behave in a human-like manner.

Similarly, Xu et al. [64] presented a method for expressing mood through the modulation of behaviours that were designed with a specific task in mind. Under this model, the behaviours are selected based on the task, while the mood modulates some behaviour parameters related to the robot's pose and motion. Combining both elements creates expressions that can achieve the communicative goal that is requested by the task, while simultaneously expressing affect. The authors conducted an evaluation where the participants designed behaviours for the robot to express specific moods. The participants' behaviours were compared with those created by the authors. This comparison showed that the parameter settings for expressing each mood were consistent with those chosen by the participants.

Instead of exclusively focusing on endowing a robot with the capacity to express moods, several authors have opted to combine mood and emotion when designing their architectures. Some authors have developed systems that only used the mood as an internal variable that influences the generation of emotions, which is then expressed. An example of this solution is the work of Itoh et al. [30], which was published in 2009. In this work, the robot's mood, identity, and conversation content are used to generate the robot's emotional state appropriately, conveyed through its facial expression. The mood is updated based on the accumulation of past emotions. The evaluation of the system was carried out by numerous participants who observed the interactions between the robot and the user, and then answering a questionnaire. Their

results showed that personality could be expressed by changing the robot's mood transitions.

Other authors have endowed their robots with the ability to express both mood and emotion. In 2008, Leite et al. [37] presented a study where they evaluated the effects that adding emotion and mood expression to an embodied agent has when this agent acts as a game companion. In particular, when focusing on the role of affective states, the authors studied if the addition of emotional behaviours can help users understand the game that is being played. In this experiment, a robot plays chess with an user, and its affective state is determined by the current state of the game. This task is performed on three levels: (i) the game module is in charge of appraising the state of the game and controlling the plays made by the robot; (ii) the emotion module receives the state of the game from the game module, and then updates the robot's emotions and mood; and (iii) the animation module receives the game actions and the affective states that have to be conveyed, and then generates the appropriate motions and facial expressions. During the experiment, the participants played a game of chess with the robot and were asked if they believed the robot to be winning or loosing the game. Their responses were then compared with an evaluation made by the chess engine's evaluation feature. The results of the evaluation shows that conveying an emotional behaviour helped users to better perceive the game. The participants also had to evaluate the state of the game according to the robot's expression. In 2011, Moshkina et al. [44] presented the TAME framework, which is an affective software architecture that has been applied to humanoid robots. This framework encompasses various affective phenomena (i.e., affect, personality traits, affective attitudes, moods, and emotions), which can change through time. The affective module generates a series of variables that modify the robot's behaviours or force the robot to execute specific expressive affective behaviours. This framework was implemented in a NAO robot and evaluated through an online survey. In total, 26 participants watched videos of the robot displaying different emotions (i.e., anger, joy, sadness, interest, fear, or disgust), moods (i.e., positive or negative), and traits (i.e., extroverted or introverted) through affective expressions. Their results show that participants correctly perceived the emotions and traits displayed but had problems perceiving the negative mood.

In 2015, P. Gebhard [24] proposed ALMA, which is a model to represent affective states that takes into account three characteristics: emotions (short-term), moods (medium-term), and personality (long-term). This work combines the OCC model of emotion with the five factor model of personality, with the addition of moods represented through a Pleasure-Arousal-Dominance space that is divided in eight octants. Emotions in this model are elicited in response to events happening in the world, and they decay over time. The set of active emotions then influences mood transitions,

where the mood becomes more intense if multiple events that support this mood happen. The default mood of a character is computed based on its personality. In ALMA, the global parameters used for affect computation and the personality profile for the character are defined in an XML-based modelling language. This personality profile contains the set of rules that will be used to define how different events are appraised in relation to emotion elicitation. The character affective state then affects the following aspects of its expressiveness: (i) wording and phrasing of utterances; (ii) dialogue strategies used; (iii) idle gestures; (iv) features of conversational gestures; and (v) facial expressions. In 2017, Woo et al. [63] presented the integration of their modular cognitive model for a smart-device-based robot partner. An emotional model among the modules in this system allows the robot to express emotions, feelings, and moods during interactions. Emotions are triggered based on input information captured by the robot. Feelings are computed as a summation of emotions. Finally, the mood is updated based on changes in feelings. The robot can then express these affective states through three modalities: speech, motions, and facial expressions. The mood is used during speech generation to change the robot's utterances in a particular situation. Regarding motions, the robot's feelings will affect the parameters being used for movement generation. Finally, the robot's facial expression is selected from among a set of candidates using a fuzzy approach, where the robot's feeling is used to determine the type of expression that has to be used, while the robot's mood is used to determine the intensity of the expression that has to be selected. The authors presented three study cases to demonstrate how their architecture works.

2.3 Comparing Approaches

In this work, we focus on how social robots generate and express different affective states. Based on this goal, the works related to affect expression will be compared using the following characteristics:

- **Modalities:** The expressive modalities that the robot uses to transmit affective states.
- **Affective states:** Components of affect that the robot can express (e.g., mood, emotions, or both), and how these states evolve with time.
- **Expression design:** How the expressions used to express affect are designed or modified.

Table 1 shows a summary of the related works evaluated. The works reviewed in this section provide a great diversity of modes for expressing affect because this characteristic is tied to the morphology of each robot and its communica-

Table 1 Comparative of the related works reviewed in this section, according to the three characteristics considered: modalities used, affective states implemented, and expression design

Ref.	Modalities	Affective states	Expression design
Correia et al. [6]	Body posture, speech (content), animations	Individual or group-based emotion generation	Handcrafted
Javed and Park [32]	Facial expressions, motions, colour, sounds	Emotion generation	Handcrafted
Yilmazyildiz et al. [66]	Facial expressions, speech	Emotion expression	Handcrafted (basic emotions), joint poses interpolated (intermediate emotions)
Cameron et al. [3]	Facial expressions	Emotion expression	Handcrafted
Bretan et al. [2]	Body posture, motions	Emotion expression	Parameter-based gesture modulation
Gácsi et al. [27]	Motions, sound	Emotion expression	Remotely controlled. Based on dog behaviours
Feldmaier et al. [19]	Colours, motions, lights	Emotion expression	Handcrafted
Song and Yamada [56]	Colours, sound, vibrations	Emotion expression	Handcrafted
Löffler et al. [39]	Colours, sound, motions	Emotion expression	Handcrafted
Tuyen et al. [61]	Body motions	Emotion expression	Culture-based incremental learning from samples
Suguitan et al. [57]	Body motions	Emotion expression	Gestures modulated using a variational autoencoder
Tielman et al. [59]	Eye colour, speech (prosody), body gestures	Emotion expression and generation	Parameter-based gesture modulation.
Hong et al. [29]	Eye colour, speech (prosody), kinesics	Emotion expression and generation	Handcrafted (deliberative and reactive emotions)
Han et al. [28]	Facial expression	Mood expression (affected by personality)	Weighted fusion of expressions
Xu et al.[64]	Kinesics	Mood expression	Parameter-based gesture modulation
Itoh et al. [30]	Facial expressions	Emotion expression (affected by mood)	Handcrafted
Leite et al.[37]	Facial expressions, motions	Mood and emotion expression	Handcrafted (emotions), generated (mood)
Moshkina et al.[44]	Body motions	Mood, emotion, and personality expression	Handcrafted
Gebhard [24]	Speech (prosody, content) gestures, facial expressions	Mood, emotion, and personality expression	Modulated dynamically
Woo et al. [63]	Speech (prosody, content), motions, facial expressions	Mood, feeling, and emotion expression	Modulated dynamically

tive channels. Among the different interaction modalities that may be used, the one that appears in most of these works is kinesics (i.e., body motion and body postures). Examples of this approach appear in the works of Moshkina et al. [44], Xu et al. [64], Gácsi et al. [27], Löffler et al. [39], or Suguitan et al. [57]. The work developed by Hong et al. [29] and Tuyen et al. [61] provide great insights regarding the modulation of body motions. Meanwhile, Correia et al. [6] is the only study where body postures are used to express affect without controlling body motions. Affect state expression can be done by modulating the parameters that define the robot's posture and motions, as seen in the works of Suguitan et al. [57], Xu et al. [64] or Bretan et al. [2]; or by using expressions created specifically to express a particular emotion or mood, as shown in the works of Gácsi et al. [27], Tuyen et al. [61], or Woo et al. [63].

Along with kinesics, the other most popular communicative channels are speech and facial expressions. Examples of this can be found in the works of Leite et al.[37], Itoh et al. [30], Cameron et al. [3], or Yilmazyildiz et al. [66]. The works presented by Woo et al. [63] and P. Gebhard [24] also use speech and facial expressions to express affect, although they separate themselves from the other works by having the robot's affective state change the utterances that are being generated and not only altering features of the robot's voice. The work of P. Gebhard also alters the dialogue strategies that the robot follows. Other interfaces that are less common and can express affect states are LED patterns or non-speech sounds. Authors such as Correia et al. [6] or Tielman et al. [59] used these types of communication interfaces. The work of Javed and Park [32] also used colour as part of the modalities for conveying emotion.

Regarding the type of affect states that the robot can express, it can be observed that the majority of authors opted to only implement emotions. Researchers such as Cameron et al. [3], Song et al. [56], or Löffler et al. [39] implemented discrete emotions, represented by a label (e.g. happy, sad, angry etc.). Meanwhile, like Moshkina et al. [44], other authors have also considered an intensity for each possible discrete emotion. Another popular solution is to represent emotions in a continuous 2D space (either pleasure-arousal or valence-arousal). Finally, in some approaches the affect state of the user can influence the robot's state. Examples of this can be seen in the works presented by Tielman et al. [59], or Hong et al. [29]. It is interesting to mention here the work of Javed and Park [32] because they not only use the user's emotional state to influence the emotions of the robot but they also seek to guide the user to a preset emotion goal state, which will also play a role in the elicitation of emotions in the robot. Several different approaches can be observed among those works where the robot's mood is also used. The work presented by Itoh et al. [30] uses the robot's mood as a variable for computing its emotional state, which is then

expressed. The works presented by Moshkina et al. [44] and Leite et al. [37] can express both mood and emotions. Meanwhile, Han et al. [28] and Xu et al. [64] only consider the robot's mood and not its emotional state. Finally, authors such as Tielman et al. [59], Moshkina et al. [44], or P. Gebhard [24] also included personality traits among the internal states that the robot can express, while Woo et al. [63] considered the robot's feeling in addition to the emotions and moods.

The last feature that is considered in the analysis of the works presented in this section is generating the robot's expressions. Two main strategies can be highlighted: either the robot uses expressions designed to express specific affective states, or the robot's expressiveness architecture can modify the robot's actions to adapt them to the robot's affective state. Among the former, we can find the works of Cameron et al. [3], Feldmaier et al. [19], or Löffler et al. [39]; while among the latter we can find the works of Tielman et al.[59] or Bretan et al.[2]. Usually, entirely handcrafted expressions tend to be used in systems that model the robot's affect state as a discrete value, while works, where the affect state is modelled as a 2D space (or as a set of discrete states with a continuous intensity level) tend to rely on the modulation/parametrisation of predefined templates. Finally, it is interesting to highlight the works of Tuyen et al. [61] because their system can learn from the user's actions to generate affective expressions and Woo et al. [63] because their system considers that not all affective states modulate all of the robot's interfaces.

2.4 Our Approach

Following Moshkina et al. [44], our approach implements mood and emotional expression in a social robot. From a generation standpoint, both the emotions and moods are modelled as 2D variables in a valence-arousal space. When expressing the robot's affect state, emotions are considered as continuous variables with an associated level of intensity, while moods are discrete. Emotional intensities increase with the perception of environmental stimuli. When the associated stimulus disappears, its intensity starts decaying with time. In this aspect, our research can be compared to the works of Moshkina et al.[44] because it also presents a robot that is able to express mood and emotion simultaneously, with a temporary decay of emotional intensities.

While the system selects predefined expressions based on the affective state in this work, we combine handcrafted affective expressions with the modulation of non-affective expressions using modulation profiles. The work presented by Woo et al. [63] could also be compared with our system because emotions play a similar role in our architecture to the role being played by feeling in their system. The dif-

ference is that in our case all of the robot's interfaces will be modulated based on a combination of both affect states, as defined by the modulation profiles, while in the work of Woo et al. speech is modified based on mood, motions are modified based on feeling, and the facial expression is a product of both. Itoh et al. [30] also presented a system where the robot experiences emotions and moods, but the mood is only used to generate new emotions and not express them.

The approach that our system uses to modulate the generic expressions based on the robot's affective state is similar to the one presented by Bretan et al. [2] or Xu et al. [64]. In the work of Bretan et al. [2], a series of control parameters can be used to alter the robot's motions. The robot's state is used to modify the control parameters, which are then used to alter the motion primitive selected. Three main differences can be observed with our approach. The first is that we compute the combined effect of the emotion and the mood for each possible control parameter considered (e.g., the pitch of the voice, the speed of the motions, or the blinking frequency). The second difference is that we combine the modulation of the robot's expressions with the performance of emotional expressions that have been handcrafted to respond to a specific stimulus that triggers a particular emotion. The last difference is that we combine the modulation of expressions to convey affective state with the correct development of the robot's tasks. This makes our approach similar to the work presented by Xu et al. [64], although we combine moods and emotions.

Tielman et al.'s [59] work could be compared with our approach. In their work, the system maintains three internal parameters (i.e., valence, arousal, and extroversion) that are updated based on emotional occurrences. When an emotion is triggered, the system can use poses implicating the entire body to express the emotion and use the internal parameters to modify the robot's expressions. In our approach, instead of using internal parameters to perform the modulation of the robot's expressions, we define a series of modulation profiles, where developers can define the specific effect that each possible affect state has over the robot's interfaces. This effect is then amplified or reduced depending on the intensity of the emotion. Our approach also allows us to combine the effect of mood and emotions instead of focusing on emotions only. Finally, the emotional expressions in our approach are connected to the specific stimulus that triggered the emotion instead of being a generic expression of that particular emotion. Although the use of handcrafted profiles to control the affective state of the robot is similar to the approach followed by P. Gebhard [24], we define the effect that each affective state has over the robot's interfaces, while Gebhard uses profiles to define the global parameters that are used to compute the robot's affective state and its personality.

3 Affect Generation

This section describes how the affective architecture that is presented in this manuscript generates the robot's affective state based on mood and emotion. Our model yields emotion and mood using the stimuli perceived by the robot. Stimuli (which are denoted as affective elicitors following Velásquez's [62] definition) are responsible for activating emotions and defining mood. Two valence-arousal spaces are used to calculate emotion and mood. The valence-arousal axes in both spaces range from 0 to 100 units.

Stimuli can have different effects on mood and emotion. The valence axis represents the robot's pleasantness (if the stimulus is positive or harmful to the robot) and the arousal axis represents the robot's excitation. Therefore, we model the effect of stimuli with a valence and arousal effect. We define the affective state (AE) as a time-dependent state where emotion and mood blend. In our model, the affective state is formed by the dominant emotion (de) and the current mood (cm), as Equation 1 shows. If there is no dominant emotion, then the dominant emotion is defined as none. This fact occurs when no emotion has an intensity over the activation threshold of 20 units. However, there is always an active mood that, by default, is the neutral one.

$$AE = \{de, cm\} \tag{1}$$

Next, we describe the affective elicitors recognised by Mini, their effects on emotion and mood, and how mood and emotion are generated and blended, which define Mini's affective state.

3.1 Affective Elicitors

Our affective state depends on how we perceive the stimuli that emerge from the internal and external situations that we experience. For example, fear usually arises due to external factors, such as when facing a dangerous situation. In contrast, some affective states depend on our interpretation of the stimuli around us, such as feeling happy when meeting a good friend. The way in which we consider the effect of stimuli on affect follows Velásquez's [62] classification and interpretation of stimuli. Velásquez's approach was notably influenced from Izard's multi-system and emotion activation [31], and Tomkins [60] beliefs about cognitive and non-cognitive emotional elicitors. Unlike the previous works, our model applies to robots rather than humans. Therefore, we have based the effect and influence of the affective elicitors on the expressiveness that we pretend our social robot Mini to exhibit. This expressiveness aims to transmit certain affective responses to improve human-robot interaction, which allows the robot's users to perceive how the robot affectively feels. In the model,

Table 2 Features of the affective elicitors considered in this work by the perception system of the robot. Each elicitor has a different effect on the valence and arousal dimensions

Activation system	Elicitor	Appraisal	Effect on Valence	Arousal
Sensorimotor	Touch stimulus	Hit	−45	30
Sensorimotor	Touch stimulus	Stroke	0	80
Cognitive	User answer	Correct	70	15
Cognitive	User answer	Incorrect	−60	−5

affective elicitors simultaneously impact the robot's emotions and mood differently.

Table 2 shows the affective elicitors that the robot can perceive, and their effects on the robot's pleasantness and arousal. The robot evaluates touch stimuli (physical contact) as strokes (positive) and hits (negative). When the robot plays a quiz game, it evaluates whether the answers provided by the user are correct or incorrect. The effects of the stimuli have been empirically determined to promote the activation of specific emotions. Therefore, in our model hits favour anger activation, strokes lead to surprise, correct answers elicit joy, and incorrect answers elicit sadness. Because the robot can perceive more than one stimulus simultaneously, more than one emotion can be active simultaneously. Nevertheless, the mood considers the aggregate effect of all stimuli, so only one mood can be active simultaneously.

Each stimulus has an optimal set point that is situated in the valence-arousal space, which defines how much pleasure (valence) and excitement (arousal) the robot feels when it perceives the stimulus. If the stimulus is not perceived, then it has no effect. When perceiving the stimulus, its effects on the valence and arousal axes move towards their setpoints, as indicated in Table 2.

3.2 Emotion

Emotions are biological processes that define the human's affective state in the short term [38]. They are typically elicited after perceiving specific stimuli and they trigger specific expressions [12]. In this work, these expressions depend on the robot's interpretation and the model's definition. The assessment of stimuli is a personal experience that varies across individuals, which lead to different affective responses depending on the person [12]. Taking a strong influence from the Cathexis architecture developed by Velásquez [62], in this study we define emotions as short-lived experiences whose intensity value e_i depends on the effect of stimuli. If more than one stimuli are perceived, then their effects are computed separately so each stimulus can trigger a different emotion.

Considering Ekman's study about basic emotions across cultures [14], we opted to include the emotions of joy, sadness, anger, and surprise. Mathematically, the intensity of each emotion is represented by a continuous signal ranging [0, 100] that evolves with time since valuable studies
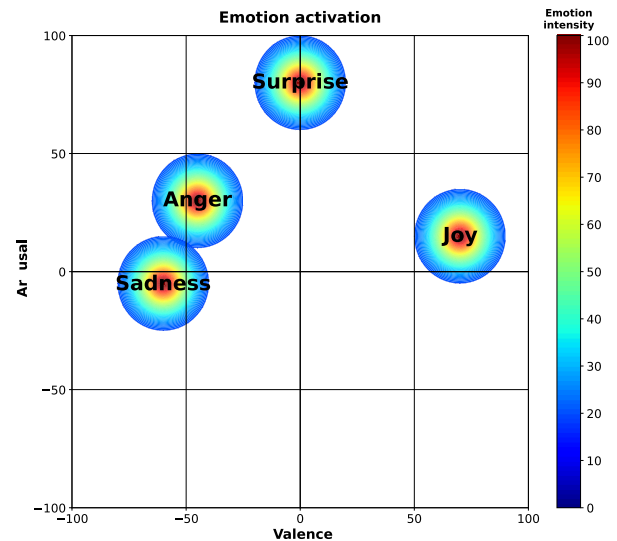

Fig. 1 Valence-arousal space to calculate emotion. Each emotion (joy, sadness, anger, and surprise) has an activation region inside the bidimensional space obtained from Russel's study [51]

support that emotions are triggered with different intensities inside a valence-arousal space [48–50]. As explained in the previous section, we tied each affective elicitor to one emotion (hits-anger; strokes-surprise; correct answers-joy; wrong answers-sadness). Figure 1 shows how emotions are distributed in the valence-arousal space following the study performed by Russell [51]. The heat map shown in Fig. 1 represents how the intensity of the emotions increases when the valence-arousal effect of a stimulus gets closer to the optimal value of the emotion (located in the centre of the heat map). Therefore, when Mini perceives a stimulus, its valence-arousal effect specified in Table 2 moves inside the valence-arousal space, which activates their corresponding emotion. Because each stimulus has its own effect and is computed separately, different emotions can be triggered simultaneously.

The perception of an stimulus makes its associated emotion increase its intensity following Equation 2 [16] (where $sf = 10$ is the increase speed factor). Equation 2 corresponds to the heat map that represents how the intensity of each emotion increases when its associated stimulus is perceived. If the robot keeps perceiving the stimulus and the
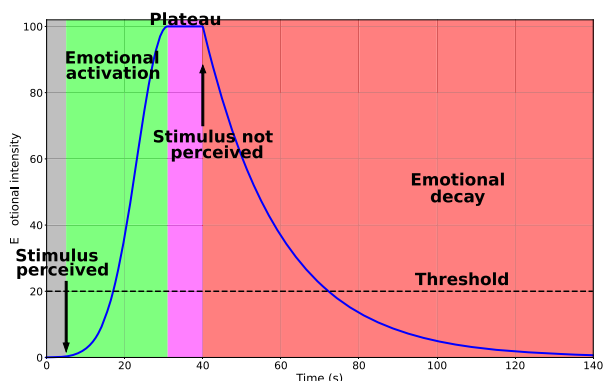
Fig. 2 Emotional activation example. When a stimulus is perceived, its related emotion gets active increasing its intensity following Equation 2. When the maximum intensity (100) is attained, the emotional intensity is maintained by entering a plateau zone. Once the stimulus is not perceived, the emotional intensity starts decaying according to Equation 3

emotion reaches its maximum intensity of 100 units, then the intensity level maintains in 100 units until the stimulus is no longer perceived. The intensity of the emotion then starts decaying exponentially according to Equation 3 [16,18] by a decay rate $dre = -0.05$. We empirically set the decay rate considering that the emotional intensity takes approximately 100 seconds to decay from 100 units to 0 units.

Figure 2 shows the evolution of the intensity associated with a particular emotion. The figure depicts the moment when the robot perceives the stimulus, causing the increase of the emotional intensity, the plateau once the emotion reaches the maximum intensity, and the exponential decay when the robot stops perceiving the stimulus.

$$\text{Emotional increase: } e_i(t) = 100 * e^{\frac{(-t+1)^2}{sf}} \tag{2}$$

$$\text{Emotional decay: } e_i(t) = 100 e^{dre*t} \tag{3}$$

If the robot perceives more than one affective elicitor simultaneously, then several emotions can be triggered. However, because the Mini social robot can only express one emotion, we opted to define a winner-take-all strategy drawing on similar works in affective generation [23]. Thus, the emotion with the highest level of intensity is the dominant emotion (de), as Equation 4 shows. To avoid shallow emotions becoming dominant, we empirically set a threshold value of 20 units for an emotion to become dominant (see Fig. 2). If no emotion has an intensity greater than 20 units, then there is no dominant emotion and the robot's mood determines the robot's affective state.

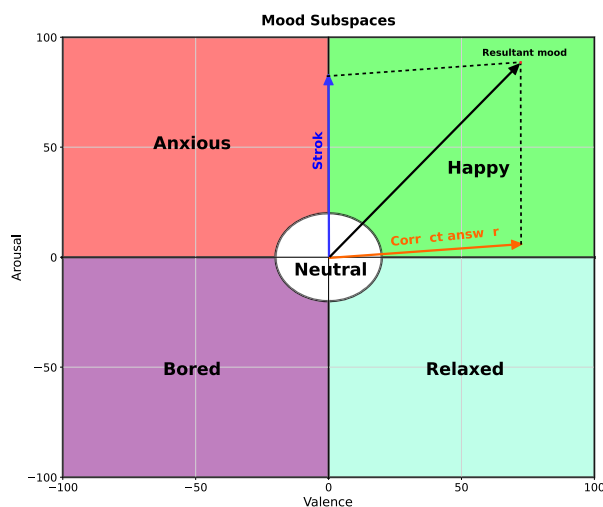$$de = \max(e_i : i = 0, ..., n) \tag{4}$$

Fig. 3 Valence-arousal affective space used to define the robot's mood. In this space, the possible moods are distributed drawing on Zhang et al. model [67]. The perception of a stroke results in the blue vector and the perception of a correct answer in the orange vector. The active region after the summation of both vectors defines the agent's mood (happy). (Color figure online)

3.3 Mood

In the last few years, theoretical studies have explored how our mood is generated using dimensional approaches (Russell [51], Plutchik [48], or Mehrabian [42]). These authors agree on including at least two axes where moods situate as subregions of the entire space.

As mentioned in the previous section, in this model we opted to define a bidimensional valence-arousal space, where moods situate as Fig. 3 shows. Drawing on how Zhang et al. [67] shape moods using the Mehrabian [42] theory, we opted to define five mood states: neutral, happy, anxious, bored, and relaxed. We model moods as the long-lasting side of the robot's affective state that act as the baseline in the absence of impulsive emotions. Note that unlike Zhang et al., we have included a neutral mood placed on the stability interval of the valence-arousal axes (situated in the middle of each interval) to set a default affective state in the absence of a dominant emotion and a more specific mood.

Mood generation is affected by stimuli in a different way than how they affect the generation of emotions. The bidimensional space used to define the robot's mood is the same (valence-arousal) as the one used for emotions, but we use independent data points for calculating each component of the robot's affective state (emotion and mood). Since only one mood can be active simultaneously, we consider the aggregate effect of stimuli to compute the robot's mood. This aggregated effect is represented as a vector considering the effects of each stimulus set in Table 2, as Fig. 3 shows.

The robot's mood evolves with time, moving in the direction of the aggregating vector following Equations 5 and 6. In both equations, the rate $drm = 0.001$ represents how fast the valence-arousal values move in the direction of the vector that aggregates the effects of all stimuli that Mini perceives. The valence-arousal values are updated every 0.5 second. The rate dr was empirically determined to let each valence-arousal values move from 0 to 100, and vice versa, in approximately two hours. The resulting vector is recalculated if the robot perceives new stimuli or stops perceiving a stimulus. The resulting vector is null if the robot does not perceive any stimulus, so the valence-arousal slowly returns to the neutral mood. Note that the decay rate of the mood calculation drm is much lower than the decay rate set for the emotions dre. This solution states the long-term effect of mood in the robot's affective state, reinforcing the dominance of the mood in the long run.

$$v(t) = v(t-1) + e^{\pm drm*t} \tag{5}$$
$$a(t) = a(t-1) + e^{\pm drm*t} \tag{6}$$

An illustrative example of how the aggregated effect of stimuli works in our model is shown in Fig. 3. If the robot perceives a stroke (blue vector) and a correct answer (orange vector), then the mood of the agent is determined by the active subregion in the valence-arousal result of the summation of both vectors. In this example, the mood of the robot is happy.

4 Affect Expression

In this work, the robot's affective state will be used to modulate its expressiveness. First, we introduce the robot Mini, which is the platform where the affective state generation architecture presented in Sect. 3 has been integrated. In particular, we focus on describing its software architecture and its expressiveness capabilities. The core of this section presents the proposed approach for expressing affective states. We then describe the software architecture that modulates the robot's expressiveness to transmit its affective states and the design stage that we have followed in this contribution.

4.1 The Mini Social Robot

Mini [52], shown in Fig. 4, is a social robot that aims to assist older adults in their daily lives. Mini can provide entertainment to the user by playing games, playing videos or songs, or reading the latest news. Given that Mini was designed for older adults that present mild cases of cognitive impairment, it can also provide cognitive stimulation therapies that are based on the recommendation of the user's therapist.

Fig. 4 The Mini social robot, which is used as the platform where our affective model has been implemented to generate and express affective states

Mini exhibits a fully autonomous intelligent behaviour that is controlled by a decision-making system [40] that is in charge of sequencing which behaviour is the most suitable for each situation in which the robot is involved. Its software architecture includes a Human-Robot Interaction (HRI) manager [21], which is used for human interaction. These functionalities are possible as a consequence of Mini's sensorimotor capabilities. Mini's perception system contains a $3D$-stereo camera to perceive the user's presence, tactile sensors to sense physical contact on its foamy case, a microphone to recognise ambient noise and understand human speech, and a tablet device through which cognitive stimulation and similar exercises are performed.

The actuation capabilities of the robot are fundamental for expressing different affective states (emotions and moods in this work). Mini has five motorised degrees of freedom that are placed in its head, neck, arms, and hip, which allow it to perform different gestures and movements. Inside its head, two animated screens simulate the appearance of natural eyes, which can blink at an indicated frequency and can be configured with different expressions. A blue LED array placed on its mouth simulates the voice modulation, and an RGB led placed at the front of its chest simulates the heartbeat. In these LEDs, we can control the blinking frequency, colour, and intensity. These actuation capabilities endow Mini with an affective expressiveness that seeks to improve human interaction by expressing emotional cues. The strategy that is developed for this will be presented in the following section.

4.2 Affect Expression in the Mini Robot

Overall, affective state expression in Mini works as follows. The affective generation system presented in Sect. 3 updates the robot's state at a specific rate. Every time that this update

is received in the expressiveness architecture, it checks if a new emotion was elicited, the robot's mood has changed, or the intensity of the dominant emotion changed. If any of these things are true, then the expressiveness architecture alters the parameters used to modulate the robot's expressions so they match the new affective state. Every time that Mini has to perform an expression, the expression is modulated to display the robot's current affect state. Based on this, the expression of affective states in Mini can be divided into three tasks. Mini will constantly experience a mood (we consider that the neutral state is one of the possible moods).

Thus, the first task that must be performed is to modulate every expression performed by Mini to display the appropriate mood. The second task modulates the robot's expressions to convey emotions with different intensity levels. During interactions, stimuli coming from the environment or the user can cause the elicitation of emotion. Nevertheless, at the same time that this emotion is triggered, Mini also experiences a particular mood. This means that the expressions need to convey the combined effect of the emotion triggered and the mood. The expressiveness architecture then needs to find the effect that the emotion has over the robot's expressions, attenuate this effect to represent the intensity of the expression, and then combine the effects of the emotion and mood. Finally, there will be situations when the stimulus that elicited the activation of emotion requires a specific response by the robot. For example, if the robot is playing a game with the user and it wins, then this might cause the robot to be joyful but at the same time it would be necessary to have an expression that references the fact that the robot won the game. To do this, the last task related to the expression of affective states is the display of emotional expressions to react to stimuli that trigger a specific emotion. The modulation to express both moods and emotions will be performed based on a series of modulation profiles that indicate a particular affective state's impact on Mini's expressiveness. In contrast, the expressions for reacting to the stimuli have been handcrafted beforehand. These three tasks will be presented in more depth in the following sections.

4.2.1 Modulating Mini's Expressiveness

As is the case for humans, at any given time, Mini will experience a particular mood or emotion (assuming that a neutral mood exists). This means that any expression that the robot performs while performing a task has to convey the robot's affective state and achieve the particular communicative goal that expression has been designed for. Our solution to this problem is to allow expression designers to define only those features that are essential to achieve the desired communicative goal while using the remaining features to express the affective state. In line with the works of Song et al. [56], or Löffler et al. [39], which highlighted the importance of mul-

Table 3 Parameters of each interface that can be controlled for expressing affective states

Interface	Parameters
Speech	Articulation rate, pitch, volume
Motions	Position, speed
Gaze	Expression, blink frequency, pupil position, pupil size
Heart	Colour, intensity, fade speed

timodality when expressing affective states, we have also opted to use multiple communication channels to express Mini's emotions and moods. The parameters that Mini can use for expressing affective states with each interface are presented in Table 3. It is worth noting that these parameters can be controlled when expressing affective states but not all of them are needed to express all states. For example, happiness can be transmitted without changing the fade speed of the coloured LEDs or the size of the pupils in the robot's eyes. The relationship between these features and each particular affective state that Mini can express will be presented in Sect. 4.2.2.

To endow Mini with the ability to express affective states, we have developed a modulation approach where roboticists can define in a series of modulation profiles how Mini can express each particular affective state with the features shown in Table 3. In particular, our system includes one profile to store the modulations for all possible moods and another for all possible emotions. We have imposed limits for the parameters in Table 3 to avoid unnatural expressions (e.g., using an excessive pitch). Every time that a new mood or emotion is elicited, the expressiveness architecture retrieves the configuration related to the new state from the profiles. It then uses the configuration to modulate the robot's expressions from that point onwards until the next update of the robot's affective state.

The process that the robot's expressiveness architecture follows when modulating expressions changes depending on whether the robot is experiencing a change in mood or an emotion has been triggered. While moods are discrete, emotions can have an associated intensity level that must be considered. Whenever Mini experiences a change in its affective state, its expressiveness architecture retrieves the configuration that is related to the new state from the corresponding modulation profile. If the mood has changed, then the architecture uses the information extracted from the modulation profile to find the appropriate values for the features in Table 3 that are connected to the new mood. Every time that an expression has to be performed, the parameters that do not have a value defined by the developer are filled with the values related to Mini's mood.

As stated earlier, Mini always has a mood (considering that *neutral* is one of the possible moods), while emotions

are not always active but rely on the perception of stimuli. Emotions include an intensity level that vary from 0 to 100. This event means that Mini's expressiveness will reflect its mood at any given time. Whenever an emotion is triggered, the expressiveness stops conveying the mood and instead displays only the emotion with the highest level of intensity of all emotions. As the intensity of the emotion decays, its effect on Mini's expressiveness also starts to fade away, while the effect that the mood has over the expressiveness starts to be perceived, which blends the expression of both affective states. The combined effect that Mini's mood and dominant emotion have over each of the features in Table 3 is calculated as shown in Equation 7,

$$effect(t) = e_m + (e_e - e_m) * i(t) \tag{7}$$

where e_m is the impact that the robot's mood has over a particular feature; e_e is the dominant emotion's effect on the feature; and i is the emotion's intensity, which is normalised to a value between 0 and 1. The effect computed using Equation 7 is used to find the value that each of the features of Mini's interfaces has to take whenever a new expression has to be performed. This effect will be updated every time that the intensity changes until the emotion disappears completely and Mini's expressiveness goes back to expressing only its mood.

Besides altering its expressiveness to express whatever affective state it is experiencing, Mini can also react to the stimulus that triggered a specific emotion by executing an appropriate emotional expression. However, this cannot be done by simply modulating the robot's expressiveness because each stimulus will require a detailed response. For example, if the user hits the robot, then it is not enough for Mini to seem to be angry. Its reaction should also acknowledge the fact that the user hit it and complained accordingly. These emotional responses can only be caused by stimuli that trigger highly intense emotional occurrences to avoid the robot constantly reacting to every stimulus.

The different stimuli that can elicit an emotion are connected to the expression used as a reaction for each emotion. Whenever an event triggers a change in Mini's affective state that leads to the elicitation of a new emotion, our system evaluates which emotion has been triggered, the stimulus that elicited it, and its intensity level. If the intensity is deemed to be high enough (the threshold has been empirically defined as 80), then our approach checks if the stimulus that triggered the emotion requires an emotional response and which expression should be used in response. In the case of a new stimulus triggering a different emotion while Mini is still responding to a previous stimulus, it will be ignored until the expression being performed has been completed.

4.2.2 The Effect of Moods and Emotion on Mini's Expressiveness

To define the effect that each of the moods and emotions described in Sect. 3 should have over Mini's expressiveness, we decided that the best course of action was to take inspiration from how humans express these particular affective states to achieve a natural expressiveness. There has been a wide range of research focused on identifying how humans express mood and emotions. The pioneer in this area of research is thought to be Charles Darwin, with his book "The expression of emotions in man and animals" [10]. Many researchers have since focused on finding the effect of the affective state that a person is experiencing over their communicative actions.

For each communication interface that Mini can use (i.e., voice, motions and body posture, gaze, and coloured LED), we have analysed several works that reviewed how to express the affective states present in Mini with that interface. Based on these findings, we designed the effect that the emotions and moods should have over Mini's expressiveness. The result of this process can be observed in Tables 4 and 5. In these tables, in addition to the effect that the emotions and moods have over each parameter of Mini's interfaces, we have included a list of all the works from which we have drawn inspiration when designing our modulation profiles. It is important to mention here that the modulation described for the robot's moods will always be smaller than those described for emotions, because emotions always have a higher intensity.

There are a few particularities regarding Mini's eyes for displaying affect. On the one hand, authors in the works reviewed during the design of the modulation profiles focused on how to convey affect through facial expressions. However, the only feature that we can alter in Mini's face is the eyes. Thus, we needed to design gazes that could convey the same information that humans transmit with their entire faces. Therefore, a specific gaze has been designed for each of the robot's affective states. To reflect the different levels of intensity that the emotions can show, three different versions for each expression have been developed: the first expresses the corresponding emotion with high intensity, the second is used for a medium level of intensity, and the third expresses the emotion with low intensity. An example can be seen in Fig. 5. This step was done by interpolating the eyelids' position for the neutral gaze and then for the gaze connected to each emotion.

5 Evaluation

This section presents the experiment setup we conducted to evaluate the robot's expressions.

Table 4 The effect that the different emotions that the robot can experience have on the robot's interfaces

Channel	Refs.	Parameter	Joy	Sadness	Anger	Surprise
Voice	[17,25,34,43,53,54]	Pitch	Increase	Decrease	Increase	Increase
		Rate of articulation	Small increase	Decrease	Increase	Small decrease
		Volume	Increase	Decrease	Increase	None
Motions & Poses	[7,9,10]	Speed	Small increase	Decrease	Increase	Increase
		Head pose	Raised	Lowered	Normal	Normal
		Motion amount	High	Low	High	Low
Heart	[8,58]	Colour	Yellow	Blue	Red	White
		Heart rate	High	Low	High	Low
		Intensity	High	Low	High	Normal
Eyes	[10,47,60]	Expression	Happy	Sad	Angry	Surprised
		Blinking frequency	Normal	Low	High	Low
		Pupil position	Centred	Down	Centred	Centred

Table 5 The effect that the different moods that the robot can experience have on the robot's interfaces

Channel	Refs.	Parameter	Happy	Bored	Relaxed	Anxious
Voice	[17,25,34,43,53,54]	Pitch	Increase	Small decrease	Small increase	Small increase
		Rate of articulation	Small increase	Small decrease	Small decrease	Small increase
		Volume	Increase	None	Small decrease	None
Motions & Poses	[7,9,10]	Speed	Small increase	Small decrease	Small decrease	Normal
		Head pose	Raised	Normal	Normal	Normal
		Motion amount	High	Small decrease	Small decrease	Normal
Heart	[8,58]	Colour	Yellow	Green	Green	Purple
		Heart rate	High	Low	Low	High
		Intensity	High	Low	Normal	High
Eyes	[10,47,60]	Expression	Happy	Neutral	Neutral	Afraid
		Blinking frequency	Normal	Low	Low	High
		Pupil position	Centred	Left/Right	Centred	Continuous motion

5.1 Experimental Setup

The affective system that is presented in this work allows the robot to express five moods and four emotions. The generation and expression of affective states in a social robot requires us to ensure that the users who interact with Mini can correctly perceive and recognise these affective states. To validate our system, we designed a case of use for how the proposed architecture can be used in a real interaction. We recorded a video showing the dynamics of affective modulation in the Mini social robot during a real human-robot interaction. This case of study, which is presented in Sect. 6, shows how the robot's affective state varies with time while the robot plays a quiz game with a user. In the video, the robot reacts by showing a happy expression where the robot congratulates the user for giving a correct answer to a question, a sad expression where the robot shows pity because the user gave a wrong answer, an anger expression for reacting to being hit by the user and a surprised expression where the robot reacts to a sudden stroke by the user. Once the effect of punctual emotional reactions disappears, it is possible to perceive how the emotional intensity decays with time, giving way to the expression of the robot's mood.

To prove that the case of use described is feasible, we also designed an experiment to evaluate if real users would be able to properly perceive the affective states that appear in this interaction, as follows: all four emotions, all four emotional expressions for reacting to stimuli, and two of the five moods considered (neutral and happy). We limited the amount of moods displayed to two to maintain a realistic scenario. Moods in our approach are considered to be long-term states, and thus having a video where the robot displays all possible moods would imply either recording an excessively long interaction or forcing unnatural mood transitions. We conducted a video-based evaluation where participants watched a series of videos of the robot expressing each of the affec-
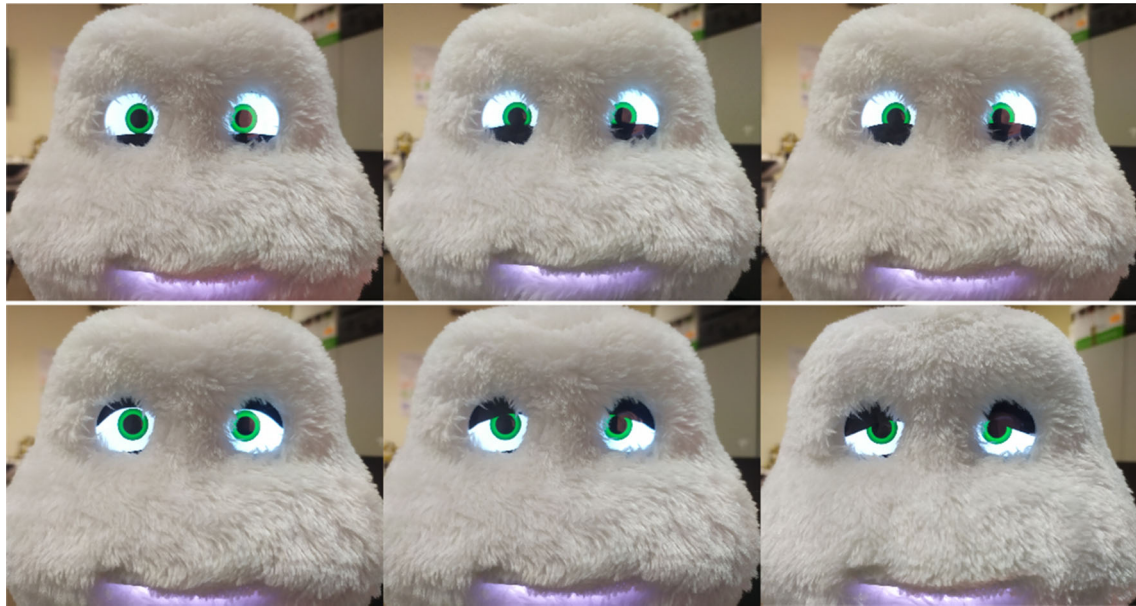
Fig. 5 Effect of the emotion intensity on Mini's gaze. The top image shows Mini displaying a happy gaze, while the bottom image shows Mini displaying a sad gaze. In both cases, the order of the images (from left to right) is low intensity, medium intensity, and high intensity

tive states shown in the case of use [1]. Each video includes an emotional reaction to unexpected stimuli, an emotional decay, or a robot's mood duration between 20 and 30 seconds. It is important to mention here that the videos show the robot performing a single expression without any other context (i.e., no other actions performed by the robot, or the user actions that might have caused the emotional response in the first play). In total, 55 participants watched and evaluated the videos showing the robot's mood and the punctual emotional reactions to stimuli. Regarding emotional decay, 36 participants watched and evaluated the videos showing emotional decay with time. The participants that validated the videos were all Spaniards.

5.2 Validation

The system was validated by sharing an online questionnaire with the participants of the experiment. In the first section of the questionnaire, the participants watched a video for each of the two moods considered in the case of use. After watching the video, the participants were first allowed to describe the mood that they perceived in the robot with a free text answer. They were then asked to select one of the options from the Mehrabian moods (i.e., happy, bored, relaxed, anxious, and neutral) [42]. The options were randomly presented to users. These moods define the bidimensional pleasantness-arousal space by which our model was inspired.

The participants repeated the same process in the second section of the questionnaire but the videos showed the robot expressing different emotions instead of moods. In the first place, the participants watched the videos concerning emotional reactions to the four stimuli considered in this work (i.e., strokes, hits, correct and incorrect answers). They then had to describe the emotion that they perceived in the robot with a free text answer. Finally, they were asked to select one of the options from a predefined list containing angry, joy, sad, surprised, disgust, and fear—the six basic emotions of Ekman [14]—, which were sorted randomly.

The third section of the questionnaire repeated the same process to evaluate the robot's emotions. However, in this case the videos showed the robot expressing the four emotions considered in this study but without reacting to any stimuli. The participants then had to provide a free text description of the emotion that they thought the robot was conveying. Finally, they were asked to select an emotion from a predefined set of alternatives randomly presented: angry, joy, sad, surprised, disgust and fear, which are Ekman's six basic emotions [14]. Finally, the last section of the questionnaire allowed the participants to leave comments and suggestions regarding the experiment that they had just participated in. It is worth mentioning essential aspects of this study here. Because all of the participants were Spanish, we formulated the questions in Spanish. In the videos showing an emotion not tied to any specific stimuli, the participants could indicate that they did not perceive any emotion in the robot (i.e., none label).
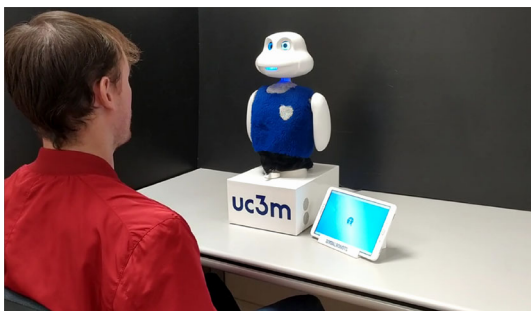
---

[1] The videos used in the experiment can be seen in the following YouTube playlist: https://youtube.com/playlist?list=PLxGCA0SJbjmliOZ_URX-1Ck2FCKYz1yI2

Fig. 6 Case study where the Mini social robot interacts by playing a quiz game. The robot's affective state is modulated depending on the dynamics of the interactions and expressed to inform the user how it feels

6 Case of Use

In this case study, the Mini social robot and a participant interacted while playing a quiz game. This scenario is shown in Fig. 6. The game consists of the robot presenting three different questions about the preferred category. During the game, the user has to select the correct answer from four different options displayed on the touch screen. After each response, the robot provides meaningful information about the question, explaining why the participant was correct or wrong. The dynamics of the experiment, and the evolution of the robot's mood and emotions are shown in Fig. 7.

The interaction begins when the user sits in front of the robot and presses the start button on the touch screen. Mini, who was previously sleeping and with a neutral mood, awakes, salutes the participant and starts the game's introduction. The introduction consists of Mini explaining that they will play a quiz game, and the user has to answer questions by using the touch screen device on the table. First, the user is requested to select the game category from geography, history, science, sport, art and literature, or entertainment. Suppose that the participant decides to select history, driving the game to start. The first question that the robot asks is about when did the attack on Pearl Harbour took place. Four different options appear on the touch screen: 1942, 1939, 1940, and 1941. Suppose that the participant provided a correct answer (1941) and the robot interprets this as a positive stimulus, then this is translated into an emotional reaction of joy. Next, Mini briefly explains the date of the attack, its origin and its consequences. During the explanation, the emotion of joy remains active but starts decaying with time. The experiment continues with the robot asking the participant whether they like history questions. Suppose that the participant answers yes, then the robot continues by introducing the next question. It is worth noting here that before starting the second question of the game, the effect of the joy emotional response con-

sequence to the participant's correct answer has disappeared completely.

The robot formulates the second question in a happy mood after the previous correct answer. This time, the robot asks where the most ancient city in south American is located, and presents the participant with four alternatives: La Paz, Valparaiso, Caral, or Arequipa. Suppose that the participant answers La Paz, which is wrong, then this leads the robot to react with an emotional expression of sadness. During the question's explanation, the emotion of sadness stays active but its effect decays with time. By the end of the explanation, the emotion has completed vanished, returning the robot to express a happy mood. The robot then asks the participant if they have been to South America, and suppose that the user replies that they have not been there. The robot then encourages them to visit this city and continues with the next question.

The third question that the robot formulates is *"Which emperor was Cleopatra married to?"*. Suppose that the participant, who manifests confidence in their answer, selects Julius Caesar instead of Ptolemy XIV, Mark Antony or the option of "none of the above". The wrong answer provided by the participant leads the robot to express sadness again but suddenly the participant hits the robot on the belly, arguing that Mini is cheating. The robot then responds with an emotional expression of anger, scolding the user for hitting it. This situation represents how affective architecture deals with emotional blending. Sadness was active while the hitting was produced but anger became dominant because its intensity is the highest and rules the affective state and expressiveness of the robot in the short term. Then, the robot continues playing and is still angry with the participant but its effect rapidly disappears after explaining the results of the question to the participant and asks them if he likes Egyptian mythology.

The last question the robot asks the participant is what age the Renaissance epoch belonged to, showing four alternatives on the touch screen: Middle ages, Contemporary age, Ancient history, or Modern age. Suppose that the participant feels confident and provides the correct answer by selecting the Modern age. Mini responds by expressing a joyful emotional expression due to the correct answer. Suddenly, the participant strokes the robot on the belly. The robot responds to the stroke by showing an expression of surprise. It is then possible to perceive how the surprise intensity decays in the robot. The game finishes with Mini explaining the answer to the question while the effect of the surprise emotion expression decays, returning the robot to express a happy mood. Finally, the robot communicates the result of the game and says goodbye by hoping to see the participant again.

Fig. 7 Evolution of the robot's mood and emotion during the case study. The user's actions influence the robot's affective state, provoking the activation of emotions in particular moments. Mood evolves during the experiment as a long-term variable that defines the robot's affective state in the long run. The expressiveness of the robot is modulated by the dominant emotion and by the robot's mood if there is no dominant emotion
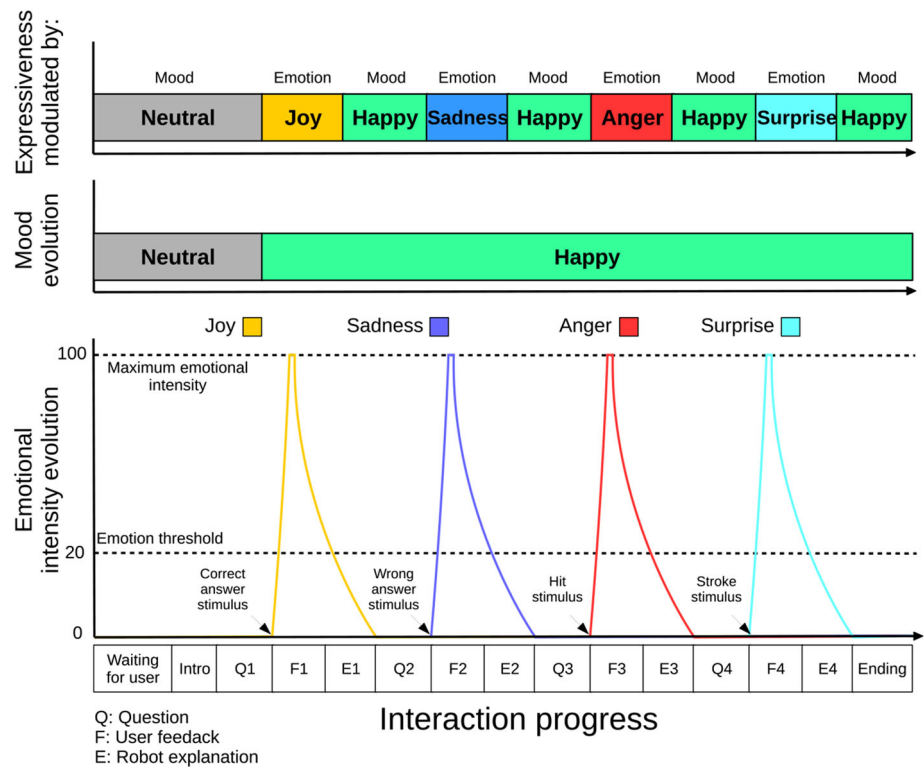
7 Results

In this section, we present the results of the evaluation that we conducted to test if the affective states that the robot displayed during the case study could be perceived correctly by the users. Figure 8 shows that nearly half of the participants could successfully identify the correct mood term in both videos when asked to use terms from a closed set. More specifically, in the case of a neutral mood, the preferred choice was neutral (44%), followed by bored (40%), relaxed (11%), and anxious (5%). Looking at the result of the robot expressing a happy mood, 49% of participants recognised the robot's mood successfully, presenting both the relaxed and neutral options a rating of 22% and anxious 7%.

The recognition rates are lower when the participants were given the option of provide an open text answer, as shown in Fig. 8. One of the possible reasons behind this is the way in which we analysed the participants' responses. To be as strict as possible, we grouped only those answers that used terms that we considered to be synonyms. For example, we considered that answers such as "joyful" and "happy" referred to the same state, while terms such as "Serious" and "Neutral" were considered to be different answers. Following this approach, we observed that 36% of participants correctly perceived the happy mood, while 20% did not perceive any mood in the robot expression and 7% provided a term out of context. In addition, a broad number of participants labelled the mood using positive terms closely related to happiness
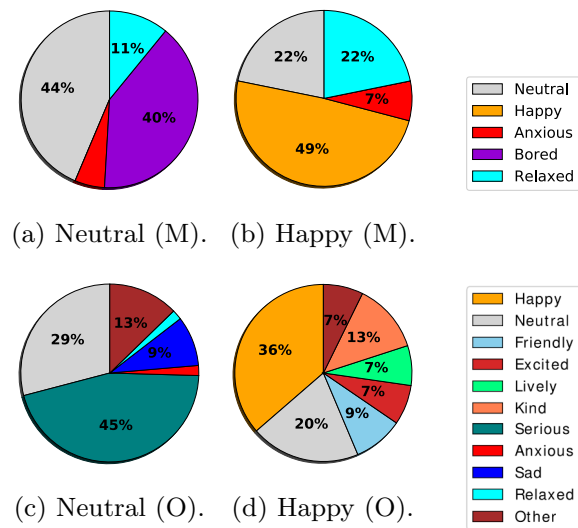

(a) Neutral (M).   (b) Happy (M).

(c) Neutral (O).   (d) Happy (O).

Fig. 8 Values obtained in the evaluation of the robot's mood expressiveness (i.e., neutral and happiness) in multi-choice (M) and open questions (O)

like kind (13%), friendly (9%), or lively (7%). Regarding the neutral mood, only 29% of participants perceived the mood correctly, while 45% of them perceived the robot as serious (as explained earlier, neutral is hard to define and serious could be an acceptable definition).

The results for the validation of the effect that emotions have on the robot's expressiveness are shown in Fig. 9. These
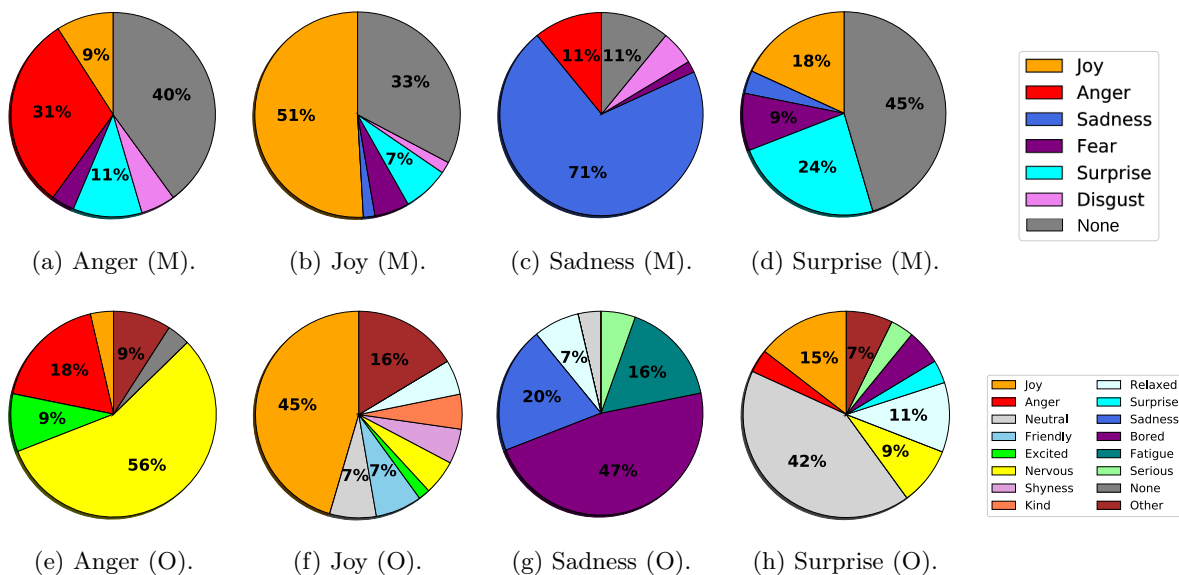
Fig. 9 Values obtained in the evaluation of the effect that emotions (i.e., anger, joy, sadness, and surprise) have over the robot's expressiveness, using multi-choice (M) and open question (O) approaches

results show that recognising emotional states was a complicated task for the study's participants. Thus, when the robot was expressing anger, the participants correctly perceived the emotion in 31% of the occasions. However, the participants selected surprise with the 11%, joy 9%, disgust 6%, and fear 3%. Meanwhile, 40% of participants did not perceive any emotional state in the robot. The expression of the emotional state joy yielded valuable results because 51% of participants successfully recognised the emotion. Nevertheless, 33% of participants did not identify any emotional state, 7% identified surprise, 5% fear, and 2% sadness and disgust. Sadness was correctly perceived by 71% of participants, which is a very positive success rate. In total, 11% perceived anger, 4% disgust, and 3% fear. However, 11% of users did not perceive any specific emotional state in the robot during the visualisation of the video. Finally, the study's participants found that recognising the surprise emotional decay was quite difficult because just 24% correctly recognised the emotional state that the robot was expressing. Meanwhile, 45% of people did not perceive any emotional state in the robot, 18% perceived joy, 9% fear, and 4% sadness.

When considering the responses given through open text answers, the results are again worse than those obtained through a closed set of possible answers, as could be expected. These results are shown in Fig. 9. Just 18% of participants correctly perceived the emotional state of anger. Most participants (56%) perceived the robot as nervous, followed by 9% feeling the robot as excited and 4% as joy. The evaluation of the joy emotional state yielded more positive outcomes: 45% of participants correctly identified the emotion that the robot was expressing, which is the most repeated

answer. However, the results for sadness were not as good as expected: 47% of participants misidentified the emotional state as bored, while 20% of the participants successfully perceived the emotional state as sadness. The next more repeated response was fatigued (16%). Finally, the assessment of the emotional state of surprise produced weak results because just 7% of participants provided a correct response. Instead, the most given options were neutral (42%), joy (15%), and relaxed (11%).

Figure 10 shows the results obtained during the evaluation of the emotions as a response to stimuli perceived by Mini. As Fig. 10a shows, anger was successfully recognised by 92% of participants, while the rest of the options present residual percentages below 3%. Joy was correctly perceived by 94% of participants (see Fig. 10b), which presents the same rate of success as surprise (see Fig. 10d). Interestingly, when the robot expressed surprise, disgust was the alternative, despite its selection percentage being below 6%. In this case, sadness was the emotion that the participants perceived with more difficulty. However, it was correctly perceived in 69% of the cases, which is still a high rating. In the evaluation, 17% of participants selected anger as the alternative emotion, followed by disgust with a 14%. For open text answers, anger was correctly perceived by 72% of participants (see Fig. 10e), joy by 78% (see Fig. 10f), sadness by 47% of the participants (the second most repeated answer was disappointment, chosen by 22% of participants, as shown in Fig. 10g), and surprise by 50% (the next closest answer was joy, selected by 17% of the participants, as shown in Fig. 10h).
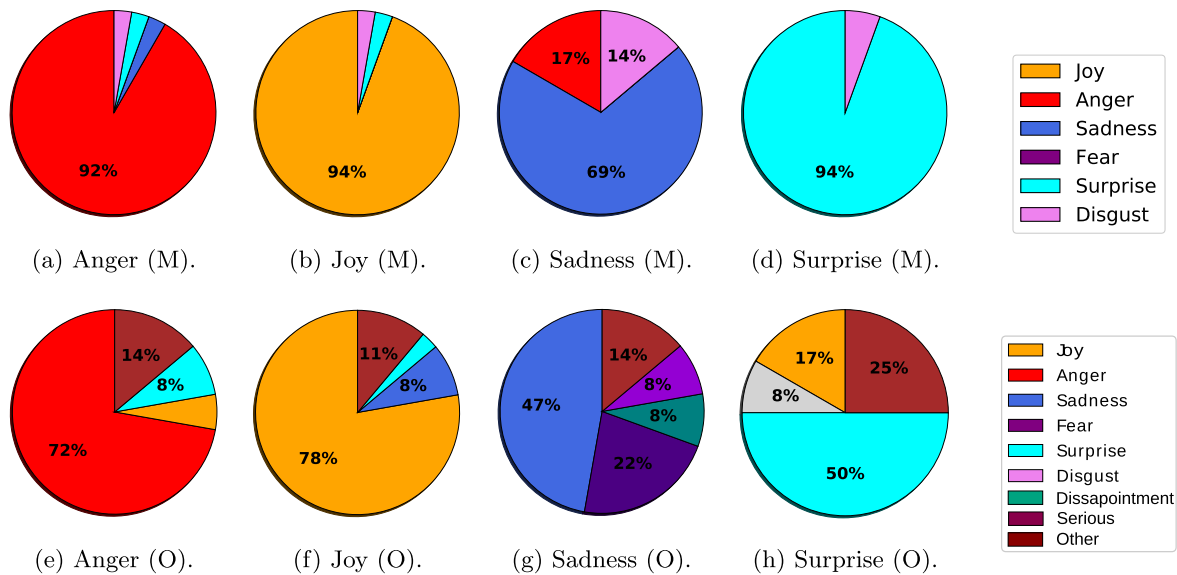
Fig. 10 Values obtained in the evaluation of the emotional expressions of the robot (i.e., anger, joy, sadness, and surprise) using multi-choice (M) and open question (O) approaches

8 Discussion

The results obtained for recognising moods and emotions in the multi-choice setting showed that the participants could recognise the affective state displayed by the robot above the chance level when given a closed set of possible answers. In particular, participants were given five options to choose from when evaluating Mini's mood, setting the chance level at 20%. In total, 44% of participants recognised the happy mood correctly, while 49% of participants successfully identified the neutral mood. Regarding the recognition of emotions, the participants were given seven options to choose from, thus having a 14% chance of selecting the correct state with a random answer. For the case where participants had to identify the robot's emotional states decaying with time (without reacting to a specific stimulus), 31% of participants successfully recognised anger, 51% joy, 71% sadness, and 24% surprise. In the case of the emotional reactions to stimuli, the difference between the recognition success rate and the chance level is more significant for emotions than in both previous cases (mood and emotional modulation) because 94% of participants identified the joy and surprise correctly, 92% identified the angry emotion, and 69% recognised the sad emotion successfully. These results show that participants had more success when asked to identify certain emotions than Mini's moods. We expected this issue from the beginning because we consider emotions to be more intense than moods, and thus they lead to more aggressive changes in the robot's expressiveness. This makes it easier for them to be recognised. When comparing the results obtained with the open and closed questions, we perceive that the recog-

nition accuracy is lower for all affective states when the participants were presented with an open question (as could be expected). The participants still correctly identified the emotional expressions used when the robot perceives stimuli that requires an emotional response, as well as the happy mood. The results obtained for the neutral mood are also acceptable because a large majority of participants identified the robot as being either serious (the majority) or neutral, which could be considered similar states, at least in Spanish. Finally, regarding the modulation used to convey emotion, the results for open text answers were less promising because only happiness was correctly perceived by the majority of the participants. Anger was confused with nervousness (two states that involve high locomotor activation, prosody rate and voice pitch), and sadness was confused with boredom (both states involve low locomotor activation, prosody rate, and voice pitch). The use of video-based evaluations could be one of the reasons behind these results because the robot's eye expression (one of the features that could allow to differentiate these states) could be harder to appreciate in videos. Finally, the recognition rates for surprise were the lowest (similar to the results observed when using multiple choice questions).

The results for emotion recognition presented in Sect. 7 show that the participants found sadness the most difficult to recognise. When comparing our results with the works reviewed in Sect. 2 (in those that presented results for recognition accuracy), we observed that our findings are in line with those reported by Gácsi et al. [27] and Tuyen et al. [61]. However, there is no consensus regarding which emotion is more difficult to recognise, as Suguitan et al. [57]

reported. The participants in their experiment had more problems recognising anger than happiness or sadness, while Bretan et al. [2] reported that participants had more problems recognising surprise than any of the other emotions (these are the only authors who considered surprise as an emotion among the works reviewed in Sect. 2). A possible cause for these variations is that several factors play a role in how easy it is for a person to recognise the robot's affective state, including the design of the expressions of the robot, the communicative modes that these expressions used, or the robot used in the experiment (among others). However, further tests would be required to evaluate this.

Overall, the results concerning punctual emotional reactions to stimuli presented in Sect. 7 are among the top recognition accuracies observed in similar works reviewed in Sect. 2 for happiness, anger, and surprise (all three above 90%). Nevertheless, the recognition rate for sadness was low when compared with the results reported by other authors. This fact points to the need to improve the modulation that is currently being used to convey sadness. Regarding mood, none of the works reviewed that provided results for mood recognition accuracy used a discrete set of moods like ours and instead focused on transmitting only a positive or negative mood. Like mood, the emotional decay with time has not been assessed in human-robot interactions previously because all evaluations focus on emotional reactions. These findings suggest that it we should continue evaluating more moods because social robots have a gap in mood generation and expression.

8.1 System Limitations

There are several limitations in our work that must be mentioned. First, we retrieved the results exclusively using a video-based evaluation, where the participants could only see videos of the robot performing the expressions. On the one hand, this approach makes it easier to add new participants to an experiment and ensures that all participants observe the same interaction. On the other hand, video-based evaluations can fail to capture some interactions that might be relevant for the evaluation. From the works reviewed in Sect. 2, Bretan et al. [2] had participants who evaluated the robot's expressions both face-to-face and using videos. Both approaches reported differences in how accurate the participants were when recognising the robot's emotions, although these differences were minor. A second limitation of the proposed approach is that it depends on how roboticists create the robot's expressions. We consider that all of the features of an expression defined by the developers are essential for the communicative task that the expression will perform, and thus cannot be used to express affective states because the modulation could change the message that the expression tries to convey. Thus, if the developers define values for most of an expression's features, then it will be hard to convey affective states with the remaining features. Finally, although the proposed affective architecture may have multiple emotions active at the same time, with different intensity levels, this will not affect the robot's expressiveness because the proposed approach focuses exclusively on the robot's mood and dominant emotion (the emotion with the highest intensity), ignoring all other emotions. Thus, Mini's expressiveness will be the same when only anger is elicited as when both anger and sadness are elicited if anger is the dominant emotion.

9 Conclusion

This work presents a software architecture that can be used to generate and express affective states in a social robot. The proposed system can generate moods and emotions and combine affective states' effects over the robot's expressiveness through the use of a series of modulation profiles. Emotions are generated as a response to a set of stimuli that are modelled in a bidimensional valence-arousal space. When the robot perceives a particular stimulus, the intensity of the associated emotion starts growing until the maximum intensity is reached. At this point, a plateau is maintained while the stimulus is still present. Once the stimulus disappears, the intensity starts exponentially decaying. In contrast, moods are discrete and do not have an associated intensity level. As occurs with emotions, the stimuli that the robot can perceive change the robot's valence and arousal levels. These changes can lead to the variation of the robot's mood. The values for valence and arousal will decay with time until the robot goes back to a neutral state.

The main contribution of this paper is the proposal of a method for endowing a social robot with the ability to express affective states. This method is based on the use of modulation profiles. In these profiles, developers can specify the effect of a particular emotion or mood over the robot's expressiveness. For each robot's communicative interface, the profiles define a configuration for the interface's parameters (e.g., the speed of the motions, the pitch of the voice, or the heart rate). The setup consists of a percentage inside the range between the parameter's lower and higher setting (e.g., a 0 for the motion speed would mean that the robot would move at the minimum speed, while a 100 would mean that these motions would be performed at max speed). When a change in the robot's affective state occurs (in mood or emotional intensities), the expressiveness architecture retrieves the effects related to the new state from the modulation profiles. It then computes the new configuration for the robot's output interfaces based on these effects and the intensity level of the dominant emotion, and uses the new configuration to modulate all of the expressions that the robot performs. This

expressiveness approach also allows us to combine the effects of the dominant emotion with those connected to the robot's mood.

Finally, certain stimuli require that the robot reacts by performing a particular expression. To allow this, an emotional response module has been designed (in the proposed system, moods do not have any associated expressions). The affective state of the robot is sent to this module at a specific rate. If a new emotion is triggered with an intensity level above a particular threshold, then the emotional response module finds the expression associated with the stimulus that triggered the emotion and sends it to the robot's expressiveness architecture to be performed.

The proposed architecture has been integrated in Mini, which is a social robot developed to assist older adults with mild cognitive impairment. An evaluation was conducted to test the proposed affective expressiveness approach. Individual videos were recorded for the two moods, the four emotional reactions to stimuli, and the four emotional modulations that the robot Mini can express. We also recorded an additional video showing the affective modulation of Mini during a real human-robot interaction.

The evaluation of individual affective expressions consisted of 10 videos (four emotional reactions, four emotions and two moods). Participants in the evaluation watched all of the videos twice. The participants first provided an open text answer for each video and they were then asked to choose from a predefined list of alternatives. The results of the experiment show that participants were able to correctly perceive all the affective states integrated into the robot, with recognition rates ranging from 94% when presented with a happy or surprised robot to 44% when presented with a robot in a neutral state (although the recognition rate is below 50%, it was still the most selected option for that particular video). This result shows that the proposed architecture can correctly handle the expression of affective states, conveying them in a recognisable way for users who are interacting with the robot.

Data Availability The datasets generated during and/or analysed during the current study are not publicly available due to the confidentiality agreements signed with the participants in the study, where it was stated that the data would not be shared with third parties, but are available from the corresponding author on reasonable request.

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

Informed Consent Informed consent was obtained from all individual participants included in the study.

References

1. Beedie C, Terry P, Lane A (2005) Distinctions between emotion and mood. Cognition & Emotion 19(6):847–878

2. Bretan M, Hoffman G, Weinberg G (2015) Emotionally expressive dynamic physical behaviors in robots. International Journal of Human-Computer Studies 78. https://doi.org/10.1016/j.ijhcs.2015.01.006

3. Cameron D, Fernando S, Collins E, et al (2015) Presence of life-like robot expressions influences children's enjoyment of human-robot interactions in the field. In: Proceedings of the AISB Convention 2015

4. Coeckelbergh M (2010) Moral appearances: emotions, robots, and human morality. Ethics Inf Technol 12(3):235–241

5. Coeckelbergh M (2020) Moral appearances: emotions, robots, and human morality. In: Machine Ethics and Robot Ethics. Routledge, p 117–123

6. Correia F, Mascarenhas S, Prada R, et al (2018) Group-based emotions in teams of humans and robots. In: Proceedings of the 2018 ACM/IEEE international conference on human-robot interaction, pp 261–269, https://doi.org/10.1145/3171221.3171252

7. Coulson M (2004) Attributing emotion to static body postures: Recognition accuracy, confusions, and viewpoint dependence. J Nonverbal Behav 28:117–139. https://doi.org/10.1023/B:JONB.0000023655.25550.be

8. Da Pos O, Green-Armytage P (2007) Facial expressions, colours and basic emotions. Colour: design & creativity 1(1):2

9. Dael N, Goudbeek M, Scherer K (2013) Perceived gesture dynamics in nonverbal expression of emotion. Perception 42:642–57. https://doi.org/10.1068/p7364

10. Darwin C (1872) The expression of the emotions in man and animals. John Murray

11. Dimitrievska V, Ackovska N (2020) Behavior models of emotion-featured robots: A survey. Journal of Intelligent & Robotic Systems 100(3):1031–1053

12. Dolcos F, Katsumi Y, Moore M et al (2020) Neural correlates of emotion-attention interactions: From perception, learning, and

memory to social cognition, individual differences, and training interventions. Neuroscience & Biobehavioral Reviews 108:559–601

13. Ekman P (1992) Are there basic emotions? Psychological Review

14. Ekman P (1999) Basic emotions. Handbook of cognition and emotion 98(45–60):16

15. Ekman PE, Davidson RJ (1994) The nature of emotion: Fundamental questions. Oxford University Press, Oxford

16. Elster J (2009) Urgency. Inquiry 52(4):399–411

17. Erdemir A, Walden TA, Jefferson CM et al (2018) The effect of emotion on articulation rate in persistence and recovery of childhood stuttering. J Fluency Disord 56:1–17

18. Fang B, Zhang Q, Wang H et al (2018) Personality driven task allocation for emotional robot team. Int J Mach Learn Cybern 9(12):1955–1962

19. Feldmaier J, Marmat T, Kuhn J, et al (2016) Evaluation of a rgb-led-based emotion display for affective agents. arXiv

20. Fellous JM (2004) From human emotions to robot emotions. Cross-Disciplinary Foundations, American Association for Artificial Intelligence, Architectures for Modeling Emotion, pp 39–46

21. Fernández-Rodicio E, Castro-González Á, Alonso-Martín F et al (2020) Modelling multimodal dialogues for social robots using communicative acts. Sensors 20(12):3440

22. Frijda NH (2017) The laws of emotion. Psychology Press

23. Gadanho SC, Hallam J (2001) Robot learning driven by emotions. Adapt Behav 9(1):42–64

24. Gebhard P (2005) Alma: A layered model of affect. In: Proceedings of the Fourth International Joint Conference on Autonomous Agents and Multiagent Systems. Association for Computing Machinery, New York, NY, USA, AAMAS '05, p 29-36, https://doi.org/10.1145/1082473.1082478

25. Goberman AM, Hughes S, Haydock T (2011) Acoustic characteristics of public speaking: Anxiety and practice effects. Speech Commun 53(6):867–876

26. Goldenberg A, Garcia D, Halperin E et al (2020) Collective emotions. Curr Dir Psychol Sci 29(2):154–160

27. Gácsi M, Kis A, Faragó T et al (2016) Humans attribute emotions to a robot that shows simple behavioural patterns borrowed from dog behaviour. Comput Hum Behav 59:411–419. https://doi.org/10.1016/j.chb.2016.02.043

28. Han MJ, Lin CH, Song KT (2013) Robotic emotional expression generation based on mood transition and personality model. Cybernetics, IEEE Transactions on 43:1290–1303. https://doi.org/10.1109/TSMCB.2012.2228851

29. Hong A, Lunscher N, Hu T, et al (2020) A multimodal emotional human-robot interaction architecture for social robots engaged in bidirectional communication. IEEE Transactions on Cybernetics PP:1–15. https://doi.org/10.1109/TCYB.2020.2974688

30. Itoh C, Kato S, Itoh H (2009) Mood-transition-based emotion generation model for the robot's personality. In: 2009 IEEE International Conference on Systems, Man and Cybernetics, pp 2878–2883, https://doi.org/10.1109/ICSMC.2009.5346563

31. Izard CE (1993) Four systems for emotion activation: Cognitive and noncognitive processes. Psychol Rev 100(1):68

32. Javed H, Park CH (2019) Interactions with an empathetic agent: Regulating emotions and improving engagement in autism. IEEE robotics & automation magazine 26(2):40–48

33. Jung MF (2017) Affective grounding in human-robot interaction. In: 2017 12th ACM/IEEE International Conference on Human-Robot Interaction (HRI, IEEE, pp 263–273

34. Kappas A, Hess U, Scherer K (1991) Voice and emotion, Editions de la Maison des Sciences de l'Homme, pp 200–234

35. Kirby R, Forlizzi J, Simmons R (2010) Affective social robots. Robot Auton Syst 58(3):322–332

36. Kontaris I, East BS, Wilson DA (2020) Behavioral and neurobiological convergence of odor, mood and emotion: A review. Front Behav Neurosci 14:35

37. Leite I, Pereira A, Martinho C, et al (2008) Are emotional robots more fun to play with? In: RO-MAN 2008 - The 17th IEEE International Symposium on Robot and Human Interactive Communication, pp 77–82

38. Levenson RW (1999) The intrapersonal functions of emotion. Cognition & Emotion 13(5):481–504

39. Löffler D, Schmidt N, Tscharn R (2018) Multimodal expression of artificial emotion in social robots using color, motion and sound. In: Proceedings of the 2018 ACM/IEEE International Conference on Human-Robot Interaction, pp 334–343, https://doi.org/10.1145/3171221.3171261

40. Maroto-Gómez M, Castro-González Á, Castillo JC et al (2018) A bio-inspired motivational decision making system for social robots based on the perception of the user. Sensors 18(8):2691

41. Massey DS (2002) A brief history of human society: The origin and role of emotion in social life. Am Sociol Rev 67(1):1–29

42. Mehrabian A (1996) Pleasure-arousal-dominance: A general framework for describing and measuring individual differences in temperament. Curr Psychol 14(4):261–292

43. Montero JM, Gutiérrez-Arriola J, Colás J, et al (1999) Analysis and modelling of emotional speech in spanish. In: Proc. of ICPhS, pp 957–960

44. Moshkina L, Park S, Arkin RC et al (2011) Tame: Time-varying affective response for humanoid robots. Int J Soc Robot 3(3):207–221. https://doi.org/10.1007/978-3-642-03986-7_1

45. Paiva A, Leite I, Ribeiro T (2014) Emotion modeling for social robots. The Oxford handbook of affective computing p 296

46. Paiva A, Mascarenhas S, Petisca S, et al (2018) Towards more humane machines: Creating emotional social robots. Towards more humane machines: creating emotional social robots pp 125–139

47. Perkins AM, Inchley-Mort SL, Pickering AD et al (2012) A facial expression for anxiety. J Pers Soc Psychol 102(5):910

48. Plutchik R (1980) A general psychoevolutionary theory of emotion. In: Theories of emotion. Elsevier, p 3–33

49. Reisenzein R (1994) Pleasure-arousal theory and the intensity of emotions. J Pers Soc Psychol 67(3):525

50. Ross AA (2013) Mixed emotions. In: Mixed Emotions. University of Chicago Press

51. Russell JA (1980) A circumplex model of affect. J Pers Soc Psychol 39(6):1161

52. Salichs MA, Castro-González Á, Salichs E, et al (2020) Mini: A new social robot for the elderly. International Journal of Social Robotics pp 1–19

53. Scherer K (1995) Expression of emotion in voice and music. Journal of voice?: official journal of the Voice Foundation 9:235–48. https://doi.org/10.1016/S0892-1997(05)80231-0

54. Shen P, Changjun Z, Chen X (2011) Automatic speech emotion recognition using support vector machine. In: Proceedings of 2011 International Conference on Electronic & Mechanical Engineering and Information Technology, IEEE, pp 621–625

55. So J, Achar C, Han D et al (2015) The psychology of appraisal: Specific emotions and decision-making. J Consum Psychol 25(3):359–371

56. Song S, Yamada S (2017) Expressing emotions through color, sound, and vibration with an appearance-constrained social robot. In: 2017 12th ACM/IEEE International Conference on Human-Robot Interaction, pp 2–11

57. Suguitan M, Gomez R, Hoffman G (2020) Moveae: Modifying aafective robot movements using classifying variational autoencoders. In: Proceedings of the 2020 ACM/IEEE International Conference on Human-Robot Interaction, pp 481–489, https://doi.org/10.1145/3319502.3374807

58. Terada K, Yamauchi A, Ito A (2012) 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, pp 314–321, https://doi.org/10.1109/ROMAN.2012.6343772

59. Tielman M, Neerincx M, Meyer Jj, et al (2014) Adaptive emotional expression in robot-child interaction. In: 2014 9th ACM/IEEE International Conference on Human-Robot Interaction (HRI), pp 407–414, https://doi.org/10.1145/2559636.2559663

60. Tomkins SS (1962) Affect imagery consciousness: Volume I: The positive affects, vol 1. Springer publishing company, Berlin

61. Tuyen N, Jeong S, Chong N (2018) Emotional bodily expressions for culturally competent robots through long term human-robot interaction. In: 2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), pp 2008–2013, https://doi.org/10.1109/IROS.2018.8593974

62. Velásquez JD (1996) Cathexis–a computational model for the generation of emotions and their influence in the behavior of autonomous agents. PhD thesis

63. Woo J, Botzheim J, Kubota N (2017) System integration for cognitive model of a robot partner. Intelligent Automation & Soft Computing pp 1–14

64. Xu J, Broekens J, Hindriks K, et al (2013) Mood expression through parameterized functional behavior of robots. IEEE

65. Yan F, Iliyasu AM, Hirota K (2021) Emotion space modelling for social robots. Eng Appl Artif Intell 100(104):178

66. Yilmazyildiz S, Henderickx D, Vanderborght B, et al (2013) Multimodal emotion expression for affective human-robot interaction. In: Proceedings of the Workshop on Affective Social Speech Signals (WASSS 2013), Grenoble, France

67. Zhang J, Zheng J, Magnenat-Thalmann N (2016) Modeling personality, mood, and emotions. In: Context aware human-robot and human-agent interaction. Springer, p 211–236