International Journal of Social Robotics
https://doi.org/10.1007/s12369-017-0447-2

CrossMark

Faces of Emotion: Investigating Emotional Facial Expressions Towards a Robot

Isabelle M. Menne[1] · Frank Schwab[1]

Accepted: 15 November 2017

Abstract

Emotions have always been an intriguing topic in everyday life as well as in science. As robots are starting to move from industry halls to our private homes, emotions have become a vital theme for the field of human–robot interaction. Since Darwin, research suggests facial expressions are associated with emotions. Facial expressions could provide an ideal tool for a natural, social human–robot interaction. Despite a growing body of research on the implementation of emotions in robots (mostly based on facial expressions), systematic research on users' emotions and facial expressions towards robots remains largely neglected (cf. Arkin and Moshkina in Calvo R, D'Mello S, Gratch J, Kappas A (eds) The Oxford handbook of affective computing. Oxford University Press, New York, pp 483–493, 2015 on challenges in effective testing in affective human–robot interaction). We experimentally investigated the multilevel phenomenon of emotions by using a multi-method approach. Since self-reports of emotions are prone to biases such as social desirability, we supplemented it by an objective behavioral measurement. By using the Facial Action Coding System we analyzed the facial expressions of 62 participants who watched the entertainment robot dinosaur Pleo either in a friendly interaction or being tortured. Participants differed in the type and frequency of Action Units displayed as well as in their self-reported feelings depending on the type of treatment they had watched (friendly or torture). In line with a previous study by Rosenthal-von der Pütten et al. (Int J Soc Robot 5(1):17–34, 2013. https://doi.org/10.1007/s12369-012-0173-8), participants reported feeling more positive after the friendly video and more negative after the torture video. In the torture condition, participants furthermore showed a wide range of different Action Units primarily associated with negative emotions. For example, the Action Unit 4 ("Brow Lowerer") that is common in negative emotions such as anger and sadness was displayed more frequently in the torture condition than in the friendly condition. The Action Unit 12 ("Lip Corner Puller") however, an Action Unit commonly associated with joy, was present in both conditions and thus not necessarily predictive of positive emotions. The findings indicate the importance for a thorough investigation of the variables of emotional facial expressions. In investigating the Action Units participants display due to an emotional situation, we aim to provide information on spontaneous facial expressions towards a robot that could also serve as guidance for automatic approaches.

Keywords Facial expression · Facial Action Coding System · Emotion · Emotional response · Empathy · Human–robot interaction

1 Introduction

A new "form of life" is moving from industry halls to our private homes. It comes in the shape of lawn mower robots, vacuum cleaner robots or entertainment robots. The appearance of robots outside of their traditional industrial context leads to new challenges and questions, such as: How are we going to treat social robots? Are we able to accept them as part of our life? Is it morally acceptable to harm a social robot? Can we mistreat them, tear them apart and sell them without feeling bad? In other words, how deeply ingrained is our ability to treat artificial social entities as humans? To answer these questions, a thorough understanding of the determinants and mechanisms of social reactions towards robots is needed.

The aim of our present study was to further investigate on the profoundness and specific expression of emotional reactions towards robots by using a multimethod approach

✉ Isabelle M. Menne
isabelle.menne@uni-wuerzburg.de

[1] University of Würzburg, Würzburg, Germany

Published online: 30 November 2017

for additional information and cross-validation (e.g. [2]). Since emotions are a multi-level phenomenon, a multi-method approach is recommendable (e.g. [2,56]) by using self-reports as well as observational methods. A combination of these methods avoids the flaws of self-reports (social desirability etc.) by adding a different, more objective source of information and thus a higher level of validity can be obtained [2,8]. For emotional reactions we chose to use verbalized emotions as well as nonverbalized emotions. We assessed verbalized emotions using self-reports. Nonverbalized emotions were obtained by observing facial emotional expressions.

2 Theoretical Background

2.1 A Natural Human–Robot Interaction

With robots entering our domestic lives, unspecialized non-expert humans not trained in interacting with robots come into contact with robots. However, the average person is used to and good at interacting with other humans and even pets through verbal and nonverbal communication. Current research has identified the need to capitalize on this existing ability, to equip robots with the ability of assessing the affective displays of the human communication partner. That way, people will intuitively interact with robots in a natural social manner if the robot can perceive, interpret, and appropriately respond with familiar human social cues.

2.2 Significance of the Face

The face is considered to be the most important mode of non-verbal communication (e.g. [1,22]). It is involved in most interactions among living beings. As the face houses the majority of our sensory organs (eyes, mouth, nose, ears), they are also involved in communication. Even though nonverbal communication is a multimodal process, the visual sense is the most important channel, and thus, more research focuses on the face and facial expressions than on any other nonverbal channel [39]. Since faces present an extraordinarily potent emotional and social stimuli, humans have a functional sensitivity for faces and facial expressions (e.g. [35,46]) and the facial expression of humans is the main nonverbal channel for socially interacting with others.

2.3 Facial Expression and Emotion

A natural interaction includes the interaction partner's ability to infer the affective state of its communication partner based on external observable cues. Cues that are unobtrusively observable are facial expressions. Research indicates, facial expressions are associated with emotions (e.g. [28]). Although there is a large body of research on emotion and facial expressions of emotion, most of this research focuses only on human–human interaction (see Sect. 2.5). Regarding facial expressions, social robotic research is almost solely concerned with either creating facial expressions on robots (e.g. [4,7,12,37]) and/or recognizing facial expressions posed by robots (e.g. [15,29,45,57]).

2.4 Research on Facial Expressions in Human–Robot Interaction

Kappas et al. [39] identified effective testing and evaluation of the social responses the robot invokes in people as one of the major challenges of affective human–robot interaction (HRI). However, systematic research on user's facial expressions in HRI remains rather scarce. Research not especially focused on emotional (facial) expressions provides anecdotal evidence on facial displays of emotion towards robots (e.g. [11,36]). Breazeal [11] reports subjects responded empathetic towards a robot's saddened face. One participant would "look to the experimenter with an anguished expression on her face, claiming to feel 'terrible' or 'guilty'" [11].

A study with a focus on special facial muscles in reaction to emotional stimuli [51] found a deactivation of m. corrugator supercilii (involved in Action Unit (AU) 4/Brow Lowerer) and an activation m. zygomaticus major (associated with AU 12/lip corner puller) in response to happy robotic faces, whereas an increase in corrugator activity (AU4) and a decrease of zygomaticus activation (AU12) occurred while looking at sad robotic faces. Indications for emotion expressions corresponding to the robot's facial emotional expression are found for the robot Feelix [13]. Since imitating facial expressions is an unconscious reaction in the process of sympathizing with someone else [18], analyzing facial expressions presents an opportunity for studying emotional reactions towards robots.

2.5 Research on Facial Expressions in Human–Human Interaction

Research on emotions and facial expressions has a long research tradition in psychology and the life and social sciences. We will only give some examples of (emotional) facial expression research (for an overview see e.g. [28]). Dimberg et al. (e.g. [19,20]) studied facial responses towards static pictures of emotional faces using facial electromyography (EMG). They could show that mere photographs of angry facial expressions induced spontaneous muscle activity in the corrugator supercilii (involved in brow lowering) and photographs of happy faces induced spontaneous muscle activity in the zygomatic major (involved in lip corner pulling). Sato and Yoshikawa [55] studied whether these muscle activities produced visible appearance changes. By using FACS, they found that in dynamic presentations of angry expressions, brow lowering occurred, whereas in happy expressions pulling of the lip corners could be observed. Thus, they found that appearance changes congruent with the type of facial expression presented could be observed in the subject's face. Facial expressions corresponding to emotional experiences were also found in research on negative emotional experiences (e.g. for pain: [16]; for media effects: [58]).

2.6 The Facial Action Coding System

The Facial Action Coding System (FACS) [27] is an objective and standardized method for measuring visual appearance changes in the face. It is the most comprehensive and most widely used method for facial expression analysis (e.g. [28]). Because of its descriptive power, FACS has emerged as the criterion measure of facial behavior in multiple fields including computer vision (e.g. [43]) and social (e.g. [28]) studies of emotion, among many other disciplines.

FACS defines AUs as the smallest unit of muscular activity observable in the human face. Thus, every movement is describable by either a single AU or a combination of several AUs. The assessment of AUs and interpretation of facial activity is usually a process that runs in two steps. In a first step, external signs of facial behavior, AUs, such as the lifting of the corners of the lips (AU 12), are objectively gathered. In a second step, those AUs are interpreted as indicators of psychological processes if the empirical evidence is sufficient. Thus, observable facial behavior is interpreted as an indicator for unobservable psychological processes. FACS thus does not impose meaning categories but allows for an objective description of appearance changes in the face, which presents one of the major strengths of this system.

We chose the smallest level of emotion analysis, facial appearance changes, without prior assumptions a-bout prototypical emotion expressions. This provides several advantages: First, as there is no consistent emotion theory available yet, facial behavior is discussed to be multifunctional (e.g. [21,54,56]). Second, facial indicators of emotional processes are often very subtle and change very rapidly (e.g. "micro momentary expressions" [24]). Third, the analysis on a signal basis of facial indicators of emotions allows for an objective approach to emotional phenomena. Fourth, since research has shown that facial expressions of emotion are more often partial than complete [14], questions of how single action units and combinations can be interpreted become crucial [56]. FACS allows reliable coding of facial expressions on a microanalytic level independent of prior assumptions about emotion expressions. Working with the AU codes instead of working with the emotion codes could be useful for testing the validity of emotion-specific expression typologies. Furthermore, analyzing AUs might help observers distinguish between a true felt emotion and a fake unfelt emotion [23].

Spontaneous facial expressions are often low in intensity or produce only subtle changes in the face. They are thus difficult to categorize as one of the prototypical emotions. However, since FACS is not based on categorizing expressions into prototypical emotions, but remains on a descriptive level, it is also able to capture most spontaneous facial behavior. The sensitivity of FACS to subtle expression differences could be shown by (e.g. differences between genuine and fake smile: [33]; characteristics of painful expression: [47].

2.7 Automatic Approaches

Regarding the recognition of facial expressions, automatic approaches prevail. Studies focus however mainly on the recognition of a small range of basic emotions (e.g. [9,31]). Furthermore, most automatic recognition approaches concentrate on acted rather than spontaneous facial expressions. Additionally, most AUs are still not reliably detectable automatically [9]. Thus, we aim to use this study as a starting point to investigate if humans show empathy towards a robot and if the emotional reactions can be observed in the face. Furthermore, we want to investigate if observations based on facial expressions could match data obtained from self-report.

2.8 The Media Equation and Matters of Measurement

The Media Equation states that people mindlessly treat computers, TV and other media entities like real people, interacting with them in the same way as they do with people. Computers can be seen as social actors [49] eliciting automatic and unconscious reactions to media entities, which are not limited to special groups of people, but are fundamentally rooted in evolved human nature. Due to the unconsciousness of reactions, Reeves and Nass [49] emphasize, that simply asking questions in a survey is not sufficient to validate the media equation. Rather, objective observation methods seem to be an enhancement or even a better option. However, the use of observational methods has been rather scarce. Self-reports prevail and seem to be the method of choice (cf. [8]). Even though self-reports are economical in their implementation, they come with certain limitations in reliability and validity (e.g. [3,30,60]). Despite self-reports' susceptibility to biases such as social desirability, they are often simply not applicable (e.g. infants; people with intellectually handicaps; unconscious processes etc.). However, self-reports can serve as supplemental methods, e.g. for cross-validation with observational measurements. Bethel and Murphy [8] as well as Arkin and Moshkina [2] advocate a multimethod approach in evaluating affective HRI to obtain a more comprehensive understanding and convergent validity.

3 Study Outline and Hypotheses

3.1 Study Outline

Anecdotal evidence already suggests people feel pity with a robot they had to harm with electric shocks in a replication of Milgram's study or had to "kill" with a hammer [5,6]. Participants also reported empathic concerns for a robot that was put back in a closet during a game [38]. Systematic research on users' emotional reactions towards robots remains rather scarce with a few exceptions: To investigate whether emotional reactions towards robots are profound and deeply ingrained, Rosenthal-von der Pütten et al. [52], studied unconscious emotional responses towards a robot shown in different situations (being tortured or being treated friendly) by measuring physiological arousal as well as using self-reports. They found an increased level of physiological arousal during the torture video. Subjects also reported empathic concern. Even though heart rate did not differ, skin conductance influenced emotional reactions. Since participants usually cannot manipulate the responses of their autonomic nervous system, these findings already indicate a profoundness in emotional reactions. In another study [53], the authors found neural activation when comparing human–human torture with human–robot torture. Every stimuli (human, robot, box) evoked neural activation in limbic systems. Thus, even though findings already indicate a profound effect, further research is needed to investigate on the consistency of the effect. Furthermore, due to their inherent unspecificity, physiological measurements cannot answer questions regarding the valence of emotional reactions, or distinguish whether participants experience joy or anger, (e.g. [2]). Furthermore, physiological measurements commonly suffer from certain limitations concerning their obtrusiveness (e.g. interfering natural body movement, lack of comfort, etc.). Considering these issues, the aim of our present study was to further investigate on the profoundness and specific expression of emotional reactions towards robots by using a multimethod approach for further information and cross-validation (e.g. [2]).

The study by Rosenthal-von der Pütten et al. [52] inspired and guided our research. We made some changes concerning the design and the methods. Instead of using a within-subjects design for the dependent variable "type of video" we used a between-subjects design. By using this design, people were not able to directly compare the videos, since every participant was only in one condition. Thus, influences of demand characteristics or social desirability on the results could be eliminated. Furthermore, the concern for demand characteristics applies only to self-reports, not behavioral observations. Thus, we added the motor expression component by using FACS. Several reasons were important for this step: First, emotions are a complex multilevel phenomenon, thus their measurement could profit from a multi-method approach to obtain a more comprehensive understanding and increase the validity of results [2,8]. Research indicates that facial expressions are associated with the experience of emotions (e.g. [25,28]). Second, we believe facial expressions could serve as an important input channel for further investigations on emotional responses towards robots, especially concerning their value for a natural, unobtrusive human–robot interaction. Third, as emotional processes are often unconscious and therefore not accessible to self-report [21], FACS provides an objective observation method. Fourth, most computer scientists refer to FACS when emotion is included, especially concerning the recognition of human facial expressions by robots as well as the robot's simulation of emotion with facial expressions (e.g. [50]). Thus, we studied whether a human's emotional reaction towards a robot can be observed in the face. To the best of our knowledge, there are no studies that far that systematically show human's emotional reaction towards a robot is also visible in the face. After all participants could physiologically show feelings of emotional arousal and report feeling emotions, but show no such signs in the face at all or even laugh at a robot being tortured due to the artificialness of the situation. Thus, especially due to facial expressions' significant value as input channel for processing emotions, it is important to study whether people show signs of emotional reactions in their face at all.

In this paper we refer to Menne et al. [44]. We extended our findings by providing detailed description of all AUs observed in participants. We also reanalyzed the data to include more AUs than previously analyzed and thus hope to obtain a more thorough understanding of complex (emotional) facial reactions. Even though not the main objective, but as part of the multilevel approach, we address the question, which physiological, verbal or facial aspects do best "express" the emotional reaction and is thus best used as an input for the robot's sensory.

3.2 Hypotheses

By experimentally manipulating the treatment of the entertainment dinosaur robot Pleo (either friendly treatment or torture treatment) we assume the different conditions have implications on the (emotional) facial responses of users. The friendly condition should evoke facial expressions associated with joy, such as AU 12 (Lip Corner Puller) (e.g. [19,20,48,55]). We thus hypothesize that participants will show the AU 12 ("Lip Corner Puller") more frequently in the friendly condition than in the torture condition (Hypothesis 3, H3). The torture condition, on the other side, should evoke facial expressions associated with negative emotions such as AU 4 (Brow Lowerer) (e.g. [16,48,58]). We hypothesize participants will show the AU 4 ("Brow Lowerer") more frequently in the torture condition than in the friendly condition

(Hypothesis 2, H2). Additionally, we assume participants show more AUs associated with positive emotions (AUs 6 and 12 [26]) in the friendly condition (Hypothesis 4a, H4a) and more AUs associated with negative emotions (AUs 1, 2, 4, 5, 9, 10, 14, 15, 17, 24 [26]) in the torture condition (Hypothesis 4b, H4b). Assuming participants experience the torture condition as a negative experience, AUs commonly associated with negative emotions should be displayed more frequently in the torture condition and vice versa respectively. As there are commonly more AUs associated with negative emotions than for positive emotions, we hypothesize that participants will show more facial expressions in the torture condition than in the friendly condition (Hypothesis 1, H1). Expecting similarities to previous research [52] we hypothesize that participants will feel more positively after watching the friendly video (Hypothesis 5a, H5a) and more negatively after the torture video (Hypothesis 5b, H5b).

4 User Study

We chose a one-factorial, between-subjects-design to test the hypotheses. Type of video was the independent variable. The participants were randomly assigned to one of the two conditions: "friendly interaction" or "torture interaction". The dependent variable was the emotional reaction toward the robot, captured with self-reports (H5a, H5b) and the type and frequency of specific facial expressions (H1, H2, H3, H4a, H4b).

A total of 62 participants (12 male) between 18 and 29 years of age (M = 20.32; SD = 1.91) took part in the experiment.

4.1 Stimulus Material

The entertainment robot Pleo by Innvo Labs was used for creating the two video clips. Its appearance is inspired by a baby Camarasaurus. The robot is built to be "a life form" (Innvo Labs), thus Pleo is intended to be used as an entertaining companion in everyday life. Associations with real animals are intended since Pleo does react autonomically and realistically to its environment. Sensors placed all over its body allow it to react to touch, movement, temperature and light. Through a speaker, Pleo is able to make utterances. Pleo offers much potential for observing facial expressions and emotional reactions towards a robot that expresses emotions like joy and pain believably through noises and movement.

Films are often used for emotion-elicitation tasks in emotion research (e.g. [25,34]). Lazarus [42] argued that emotions induced by films are real emotions. Since we laid a major focus on reliable and replicable robotic behaviors, we chose to use videos instead of live interaction (for a discussion on the use of films in HRI see [61]).

We used Pleo to create two short video clips, containing either a "friendly interaction" or a "torture interaction". The content and design of the videos was inspired by Rosenthal-von der Pütten et al. [52]. Both videos lasted 1 min each. Every video consisted of five sequences lasting 10 s each and separated by a 2 s black screen. The friendly video contained sequences of Pleo being caressed, Pleo being fed with a leaf, Pleo being fed with a mushroom, Pleo awakening and being gently stroked. The torture video contained sequences of Pleo being hit on the head, Pleo being shaken by the tail, Pleo's head hit on the table, Pleo choking due to a plastic bag over his head and Pleo being punched.

4.2 Conduction of the Study

The experimental setup took place in a booth like setting of the university specifically designed for quiet and unobtrusive video as well as audio recording of the participants. Inside a 3 m × 4 m room, a 2 m × 2 m booth containing a table, a chair, a computer, a shelf, where a surveillance camera (night mode) was placed, loudspeakers, infrared illumination lamps and ventilation. Outside the box, the experimenter could control the participant's computer and monitor the experiment. Figure 1 shows a floor-plan representation of the experimental environment. Each participant was seated at a computer workstation facing the "north" wall. The participant's face

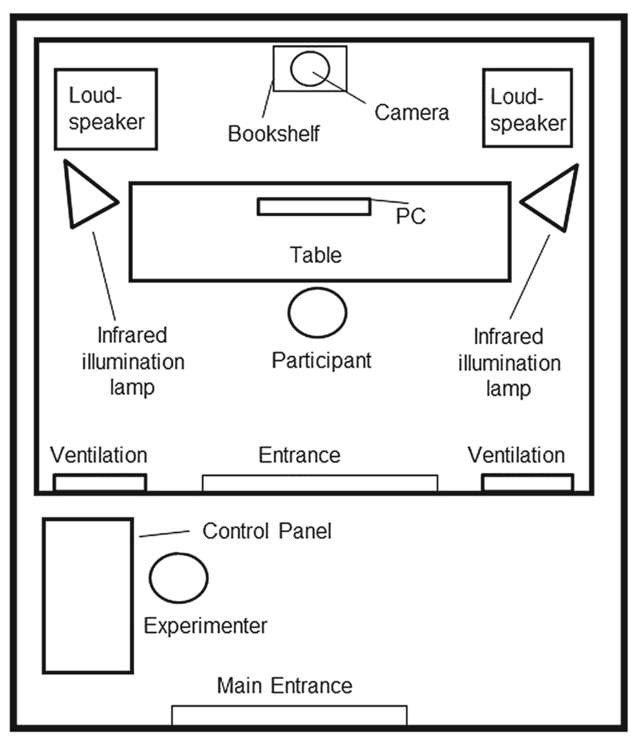

Fig. 1 Schematic drawing of the experimental setup

Table 1 Action units observed in participants

AU no.	Appearance changes	AU no.	Appearance changes
1	Inner Brow Raiser	12	Lip Corner Puller
2	Outer Brow Raiser	14	Dimpler
4	Brow Lowerer	15	Lip Corner Depressor
5	Upper Lid Raiser	16	Lower Lip Depressor
6	Cheek Raiser and Lid Compressor	17	Chin Raiser
7	Lid Tightener	20	Lip Stretcher
9	Nose Wrinkler	24	Lip Presser
10	Upper Lip Raiser	25	Lips Part

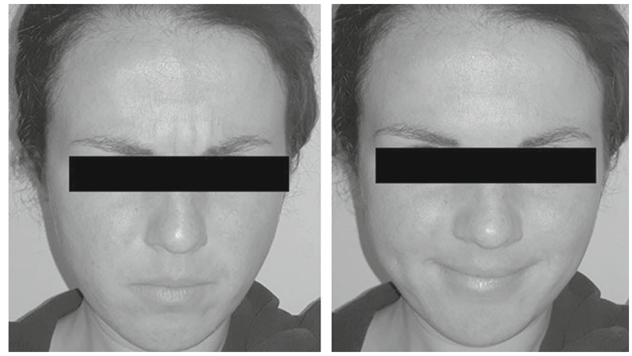

Fig. 2 Example pictures (anonymized) of AUs. Left picture: AU 4. Right picture: AU 12

was easily visible with the surveillance camera while the participant watched the videos on the computer.

To avoid participants being influenced by the presence of other participants and to ensure an undisturbed reception of the video, the study was conducted in single sessions. Participants were randomly assigned to one of the two conditions. Half of the participants watched the videos in reverse order to avoid sequence effects. After arrival, the participant received general information about the study procedure, data privacy, voluntariness and anonymity and then signed an informed consent of the video and audio recording. After completing the first web based (SosciSurvey) questionnaire concerning demographic data and their past experience with Pleo, participants watched either the friendly video or the torture video. While the participant watched one of the two videos, his/her face was recorded with an unobtrusive camera. After that, participants completed the Positive and Negative Affect Schedules (PANAS) questionnaire.

We used the German adaptation of the Positive and Negative Affect Schedules (original PANAS: [59]). The German PANAS was reliable (Cronbachs alphas $\geq$ .84). The factor-based scales for "Positive Affectivity" and "Negative Affectivity" contained ten items each. The items were rated on a five-point likert scale ranging from "nothing or very little" to "very strong" . Sums of both subscales were calculated.

4.3 Observations Using the Facial Action Coding System

Videos of the participant's facial activity were coded manually by a certified FACS coder (a coder trained in FACS and having achieved at least 80% interrater reliability in the FACS Final Test, cf. [26]). According to Ekman et al. [26], the videos were first watched in real-time, as soon as a change in the facial appearance was observed, the whole change was replayed in slow-motion to detect all Action Units that appeared. According to Friesen and Ekman [32], the apex of the following Action Units was coded (Table 1). Videos of the participants' facial activity were coded by a certified FACS coder according to Ekman et al. [27] and Friesen and Ekman [32].

To prove the reliability of the coding, 20% of participants were independently coded by a comparison coder certified in FACS. Reliability for coding was satisfying (Cohen's k = 0.76). Examples of facial appearance changes coded can be seen in Fig. 2.

5 Results

We hypothesized participants would differ in their number and type of AUs shown, depending on the video condition. Furthermore, we assumed a match between answers derived from self-reports and facial expressions. To test our hypotheses, we used a one-factorial between-subjects design with type of video as independent variable and the number and type of AUs as well as the PANAS questionnaire as the dependent variables. We analyzed the data statistically.

5.1 Overview of Action Units

How much facial activity happened? What type of facial activity could be observed? To answer these questions, we analyzed data from our video recordings as described in Sect. 4.3. The AU 12 was the AU most frequently observed in both conditions with an occurrence rate of 29.55% of all AUs that occurred more than 20 times in both conditions. AU 4 was the second most common AU that occurred (17.56%). The following AUs occurred subsequently: AU 10 (11.56%), AU 17 (10.49%), AU 25 (9.64%), AU 15 (8.57%), AU 1 (6.85%) and AU 2 (5.78%). The occurrence frequencies of AUs, ordered by treatment condition, are in Fig. 3 for the friendly condition and in Fig. 4 for the torture condition.

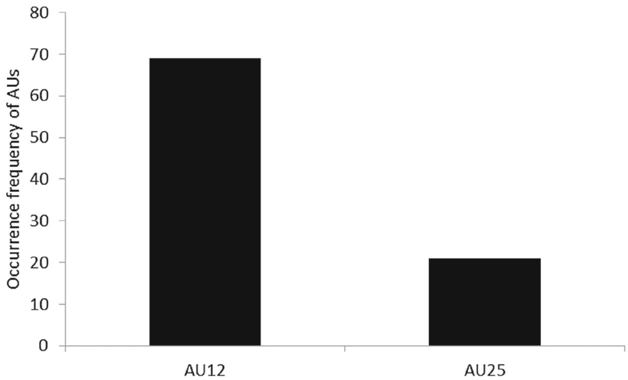

Fig. 3 Frequency of AUs in the friendly condition. *Note*: displayed here are only AUs that occurred at least 20 times

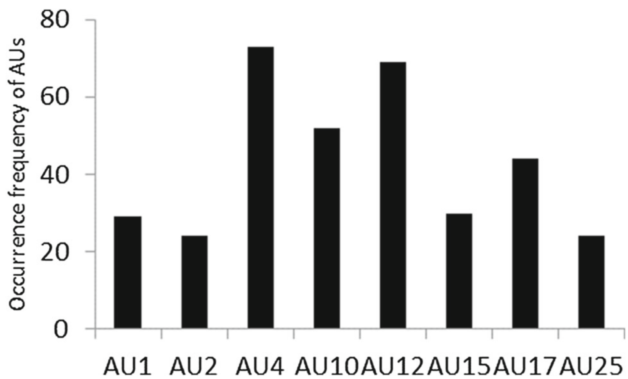

Fig. 4 Frequency of AUs in the torture condition. *Note*: displayed here are only AUs that occurred at least 20 times

Table 2 Occurrence frequency (in percentage) of action units expressed during the friendly condition scenes

Friendly condition (FC)	AU4	AU12	AU25
Caressing (body)[a]	11.11	15.94	19.05
Feeding (leaf)[a]	11.11	23.19	33.33
Stroking[a]	11.11	14.49	19.05
Feeding (mushroom)[a]	44.44	23.19	28.57
Caressing (head)[a]	22.22	23.19	0.00
Overall occurrence			
Frequency in FC	10.98	50.00	46.67

[a]Numbers are in percentage and refer to the occurrence frequency of each AU within the friendly condition. Displayed here are only AUs with an occurrence frequency of at least 10% for each scene of the friendly condition

For details regarding the occurrence frequency of AUs in each scene, please refer to Table 2 for scenes in the friendly condition and Table 3 for scenes in the torture condition.

5.2 Differences Between Facial Expressions Depending on Type of Video

For the analysis of facial expressions in relation to the different types of treatment, we used the Mann–Whitney test since our data was not normally distributed. For the self-report data we used the unpaired t-test for calculating differences between the groups. Next to the test statistics we also report the effect sizes (r).

5.2.1 H1: Differences in Facial Expressions

Participants displayed significantly more AUs in the torture condition ($Mdn = 10.50$) than in the friendly condition ($Mdn = 2.50$), $U = 763.00$, $z = 4.00$, $p < .000$, $r = .51$.

5.2.2 H2: Differences in Action Unit 4

Participants displayed the AU 4 significantly more frequently in the torture condition ($Mdn = 1.00$) than in the friendly condition ($Mdn = 0.00$), $U = 686.50$, $z = 3.32$, $p < .001$, $r = .42$.

5.2.3 H3: Differences in Action Unit 12

Participants did not display the AU 12 significantly more frequently in the friendly condition ($Mdn = 1.00$) than in the torture condition ($Mdn = 1.00$), $U = 512.00$, $z = 0.46$, $p > .05$, $r = .06$.

5.2.4 H4a: Differences in Action Units Associated with Positive Emotions

Participants did not display AUs associated with positive emotions significantly more frequently in the friendly condition ($Mdn = 1.00$) than in the torture condition ($Mdn = 1.00$), $U = 511.00$, $z = 0.45$, $p > .05$, $r = .06$.

5.2.5 H4b: Differences in Action Units Associated with Negative Emotions

Participants displayed significantly more AUs associated with negative emotions in the torture condition ($Mdn = 8.00$) than in the friendly condition ($Mdn = 0.00$), $U = 834.50$, $z = 5.09$, $p < .000$, $r = .65$.

5.3 Differences Between Self-Reported Feelings Depending on Type Of Video

5.3.1 H5a: Differences in Self-Reported Positive Feelings

Participants reported significantly more positive feelings ($M = 3.41$, $SD = .53$) after the reception of the friendly video,

Table 3 Occurrence frequency (in percentage) of AUs expressed during the torture condition scenes

TC[b]	Hitting[a]	By tail[a]	On the side[a]	On table[a]	Choking[a]	Sum.[c]
AU1	24.14	17.24	10.34	20.69	27.59	90.63
AU2	25.00	16.67	16.67	29.17	12.50	88.89
AU4	20.55	20.55	10.96	26.03	21.92	89.02
AU10	21.15	26.92	17.31	21.15	13.46	96.30
AU12	21.74	20.29	24.64	23.19	10.14	50.00
AU14	23.08	23.08	0.00	30.77	23.08	92.86
AU15	33.33	13.33	10.00	30.00	13.33	75.00
AU17	27.27	20.45	11.36	15.91	25.00	89.80
AU25	20.83	20.83	25.00	20.83	12.50	53.33

[a]Numbers are in percentage and refer to the occurrence frequency of each AU within the torture condition. Displayed here are only AUs with an occurrence frequency of at least 10% for each scene of the torture condition.
[b]Torture condition
[c]Overall occurrence frequency in torture condition

than after the reception of the torture video ($M = 3.08$, $SD = .60$), $t(60) = 2.29$, $p < .05$, 95% CI [.04, .62], $r = .28$.

5.3.2 H5b: Differences in Self-Reported Negative Feelings

Participants reported significantly more negative feelings ($M = 2.23$, $SD = .60$) after the reception of the torture video than after the reception of the friendly video ($M = 1.45$, $SD = .49$), $t(60) = 5.58$, $p < .001$, 95% CI [−1.06, −.50], $r = .58$.

6 Discussion

The aim of this paper was to gain a more comprehensive understanding of emotional processes in relation to robots. We studied facial expressions of participants using FACS and tested its applicability for analyzing emotional facial expressions in HRI. Unlike most studies in the context of HRI (cf. [8]) we used a multimethod approach for analyzing the multilevel phenomenon of emotional processes.

We assumed that witnessing a robot being tortured and being treated friendly would evoke different emotional reactions and hypothesized that these reactions would be visible in the face. Analyzing the type of AU that occurred in both conditions, AU 12 was among those AUs that occurred most frequently and with the same occurrence rate in both conditions. AU 4 followed behind, but with an occurrence rate of 89% in the torture condition. Since there seem to be more AUs for negative emotions than for positive emotions (e.g. [26]), and participants reported feeling more negative after the torture condition, we found there were indeed more facial expressions in the torture condition than in the friendly condition. Analyzing whether these differences were due to the type of AUs displayed, the results indicate that participants showed more AUs associated with negative emotions in the torture condition. However, we found no difference for AUs associated with positive emotions for the friendly condition. Focusing on the AU 12, results indicated no difference between the conditions. However, for AU 4, there was a significant difference between the conditions. The Brow Lowerer AU occurred more frequently in the torture condition than in the friendly condition. Results concerning self-reports are in line with [52,53]. Participants reported feeling more positive after the positive video and more negative after the negative video.

The torture condition evoked more negative feelings than the friendly condition and participants showed the AU 4 more frequently in the torture condition. This is in line with previous research (e.g. [19,20,51,55]). As AU 4 is among those AUs that are difficult to fake [48], the occurrence of AU 4 suggests an automatic and profound reaction independent of demand characteristics or social desirability. There is a consensus that the corrugator activity (AU 4) is commonly associated with emotions of negative valence (e.g. [19,20,48,51,55,56]. Thus, there seems to be a match for self-reports of negative emotions and reactions visible in the face.

However, AU 12 did not occur more frequently in the friendly condition than in the torture condition. The occurrence rate of AU 12 was comparable during both conditions. Even though AU 12 is typically associated with positive emotions (e.g. [19,20,26,48,51,55]) some evidence suggests the zygomaticus (involved in AU 12) is also active during negative states (e.g. [40,41]). Several things could play a role for the occurrence of AU 12 in both conditions. First, situations inducing negative emotions might trigger intrapsychic coping strategies so as to handle possible overpowering affects (e.g. [56]). It is not unusual for smiles to be used for regulating negative emotional states (e.g. [23]). Interestingly, Bartneck and Hu [5] also reported participants giggling and laughing during mistreating a Crawling Microbug Robot. During coding we observed differences in the duration of AU 12. In the friendly condition, AU 12 was displayed longer whereas in the torture condition, AU 12 occurred mostly for short durations and was frequently observed to occur in combinations of AUs associated with negative emotions. We believe it could be worthwhile analyzing this anecdotal evidence for a deeper understanding of the function of AU 12 in distinguishing different emotional facial expressions. Overall, the results support our hypotheses that torturing a robot compared to treating a robot in a friendly way elicits negative emotions and these negative emotions are also visible in the face. Using a between-subjects design and a multimethod approach we are confident we could eliminate potential influences such as demand characteristics or social desirability and could show that differences in facial behavior are profound and (mostly) unconscious processes that play an important part in emotion processing. The findings offer confirmation that we react socially and emotionally towards robots and integrate well into existing literature (e.g. [5,52,53]).

Since facial expressions are objective, unconscious and widely understood, they seem to be an ideal input for the robot's sensory of emotions. An argument for the use of videos instead of live interaction with the robot could be made since the medium of video provides a certain degree of disconnection between participants and the events shown within the videos. Had the participant been the cause of the actions against the robot, differences in participants' ability to empathize could have possibly been observed. We do not believe these differences concern the type of AUs but the intensity of AUs shown. Compared to a live interaction, the video mediated interaction might not trigger a high level of emotional reactions, due to a possible disconnection. However, facial behavior could be observed, suggesting the video condition is sufficient enough to produce emotional reactions and a certain level of involvement. For more intense emotional reactions, a live interaction is recommendable (for a discussion see [61]).

Another limitation concerns the effort for manually coding behavior. High time workload is one of the major weakness of this approach together with the absence of a comprehensive emotion theory. However, automatic approaches focus mainly on the six basic emotions (e.g. for an overview: [31]) and AUs are not yet reliably detectable automatically [9]. Research on the effects of inaccurate responses of robots towards users' affective states shows negative effects such as decreases in trust and influences on the human–robot relationship [17]. Whereas robots capable of correctly recognizing another's affect are more successful at establishing and maintaining a positive relationship with users [10]. Thus, the implementation of automatic emotion recognition systems in robotic platforms demands for a certain level of caution. Considering the impact inaccurate recognition of emotional states has on users, we think that identifying the smallest unit of facial behavior seems a promising starting point for an objective analysis of subsequent linking of facial expressions to emotional processes in human–robot interaction.

7 Conclusion and Future Work

In the present study, we could find further indications for the applicability of objective measures of nonverbal emotions by analyzing facial expressions based on FACS. By using a multimethod approach and a between-subjects design, we are also confident to have eliminated potential influences of biases such as social desirability and demand characteristics. Thus, the results indicate human emotional responses towards robots are profound and observable. In future studies, we plan to extend our study to a broader set of different robot types for examining possible differences or consistencies in facial expressions. Furthermore, we also plan for a more thorough investigation of the types of AUs occurring in different conditions. In a time, when social interactions between humans and robots become normal, determinants and mechanisms influencing affective processes should be well known for a successful human–robot interaction.

Acknowledgements We gratefully acknowledge the assistance of Christin Schnellbacher in the conduction of the study.

Compliance with Ethical Standards

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

References

1. Ambady N, Rosenthal R (1992) Thin slices of expressive behavior as predictors of interpersonal consequences: a meta-analysis. Psychol Bull 111(2):256–274. https://doi.org/10.1037/0033-2909.111.2.256
2. Arkin RC, Moshkina L (2015) Affect in human–robot interaction. In: Calvo R, D'Mello S, Gratch J, Kappas A (eds) The Oxford handbook of affective computing. Oxford University Press, New York, pp 483–493
3. Austin EJ, Deary IJ, Gibson GJ, McGregor MJ, Dent J (1998) Individual response spread in self-report scales: personality correlations and consequences. Pers Indiv Differ 24(3):421–438. https://doi.org/10.1016/s0191-8869(97)00175-x
4. Bartneck C (2002) eMuu: an embodied emotional character for the ambient intelligent home. Ph.D. thesis, University Eindhoven. http://www.bartneck.de/publications/2002/eMuu/bartneckPHDThesis2002.pdf
5. Bartneck C, Hu J (2008) Exploring the abuse of robots. Interact Stud 9(3):415–433. https://doi.org/10.1075/is.9.3.04bar
6. Bartneck C, Rosalia C, Menges R, Deckers I (2005) Robot abuse—a limitation of the media equation. In: Proceedings of the interact 2005 workshop on agent abuse, Rome, Italy
7. Becker-Asano C, Ishiguro H (2011) Evaluating facial displays of emotion for the android robot Geminoid F. In: 2011 IEEE workshop on affective computational intelligence (WACI), Paris, France. https://doi.org/10.1109/waci.2011.5953147

8. Bethel CL, Murphy RR (2010) Review of human studies methods in HRI and recommendations. Int J Soc Robot 2(4):347–359. https://doi.org/10.1007/s12369-010-0064-9
9. Brais M, Valstar MF (2016) Advances, challenges, and opportunities in automatic facial expression recognition. In: Kawulok M, Celebi ME, Smolka B (eds) Advances in face detection and facial image analysis. Springer, Cham, pp 63–100
10. Brave S, Nass C, Hutchinson K (2005) Computers that care: investigating the effects of orientation of emotion exhibited by an embodied computer agent. Int J Hum Comput Stud 62(2):161–178
11. Breazeal C (2002) Regulation and entrainment in human–robot interaction. Int J Robot Res 21(10–11):883–902. https://doi.org/10.1177/0278364902021010096
12. Breazeal C (2003) Emotion and sociable humanoid robots. Int J Hum Comput St 59(1–2):119–155. https://doi.org/10.1016/s1071-5819(03)00018-1
13. Cañamero LD (2002) Playing the emotion game with Feelix, chap. Socially intelligent agents. Springer, Boston, pp 69–76
14. Carroll JM, Russell JA (1997) Facial expressions in Hollywood's protrayal of emotion. J Pers Soc Psychol 72(1):164–176. https://doi.org/10.1037/0022-3514.72.1.164
15. Costa S, Soares F, Santos C (2013) Facial expressions and gestures to convey emotions with a humanoid robot. In: Herrmann G, Pearson MJ, Lenz A, Bremner P, Spiers A, Leonards U (eds) Social robotics. Springer, Cham, pp 542–551
16. Craig KD, Hyde SA, Patrick CJ (1991) Genuine, suppressed and faked facial behavior during exacerbation of chronic low back pain. Pain 46(2):161–171. https://doi.org/10.1016/0304-3959(91)90071-5
17. Cramer H, Goddijn J, Wielinga B, Evers V (2010) Effects of (in) accurate empathy and situational valence on attitudes towards robots. In: 5th ACM/IEEE international conference on human–robot interaction (HRI). IEEE, Istanbul, pp 141–142
18. Davis MH (1983) Measuring individual differences in empathy: evidence for a multidimensional approach. J Pers Soc Psychol 44(1):113–126
19. Dimberg U (1982) Facial reactions to facial expressions. Psychophysiology 19(6):643–647. https://doi.org/10.1111/j.1469-8986.1982.tb02516.x
20. Dimberg U, Thunberg M (1998) Rapid facial reactions to emotional facial expressions. Scand J Psychol 39(1):39–45. https://doi.org/10.1111/1467-9450.00054
21. Ekman P (1993) Facial expression and emotion. Am Psychol 48(4):384–392. https://doi.org/10.1037/0003-066x.48.4.384
22. Ekman P, Friesen WV (1969) The repertoire of nonverbal behavior: categories, origins, usage, and coding. Semiotica. https://doi.org/10.1515/semi.1969.1.1.49
23. Ekman P, Friesen WV (1982) Felt, false, and miserable smiles. J Nonverbal Behav 6(4):238–252. https://doi.org/10.1007/bf00987191
24. Ekman P, Friesen WV (2003) Unmasking the face: a guide to recognizing emotions from facial clues. Malor Books, Los Altos
25. Ekman P, Friesen WV, Ancoli S (1980) Facial signs of emotional experience. J Pers Soc Psychol 39(6):1125–1134. https://doi.org/10.1037/h0077722
26. Ekman P, Friesen WV, Hager JC (2002) FACS investigators guide. Research Nexus eBook, Salt Lake City
27. Ekman P, Friesen WV, Hager JC (2002) The facial action coding system. Research Nexus eBook, Salt Lake City
28. Ekman P, Rosenberg EL (2005) What the face reveals. Oxford University Press, New York
29. Endo N, Momoki S, Zecca M, Saito M, Mizoguchi Y, Itoh K, Takanishi A (2008) Development of whole-body emotion expression humanoid robot. In: 2008 IEEE international conference on robotics and automation. IEEE, Phoenix. https://doi.org/10.1109/robot.2008.4543523
30. Fan X, Miller BC, Park KE, Winward BW, Christensen M, Grotevant HD, Tai RH (2006) An exploratory study about inaccuracy and invalidity in adolescent self-report surveys. Field Methods 18(3):223–244. https://doi.org/10.1177/152822x06289161
31. Fasel B, Luettin J (2003) Automatic facial expression analysis: a survey. Pattern Recogn 36(1):259–275. https://doi.org/10.1016/s0031-3203(02)00052-3
32. Friesen WV, Ekman P (1983) Emfacs-7: emotional facial action coding system. Unpublished manual
33. Giudice MD, Colle L (2007) Differences between children and adults in the recognition of enjoyment smiles. Dev Psychol 43(3):796–803. https://doi.org/10.1037/0012-1649.43.3.796
34. Gross JJ, Levenson RW (1995) Emotion elicitation using films. Cogn Emot 9(1):87–108. https://doi.org/10.1080/02699939508408966
35. Haxby JV, Hoffman EA, Gobbini M (2000) The distributed human neural system for face perception. Trends Cogn Sci 4(6):223–233. https://doi.org/10.1016/s1364-6613(00)01482-0
36. Heerink M, Kröse B, Evers V, Wielinga B (2009) Relating conversational expressiveness to social presence and acceptance of an assistive social robot. Virtual Real 14(1):77–84. https://doi.org/10.1007/s10055-009-0142-1
37. Hegel F, Eyssel FA, Wrede B (2010) The social robot Flobi: key concepts of industrial design. In: Proceedings of the 19th IEEE international symposium in robot and human interactive communication (RO-MAN 2010), Viareggio, Italy, pp 120–125
38. Kahn PH, Kanda T, Ishiguro H, Freier NG, Severson RL, Gill BT, Ruckert JH, Shen S (2012) Robovie, you will have to go into the closet now: childrens social and moral relationships with a humanoid robot. Dev Psychol 48(2):303–314. https://doi.org/10.1037/a0027033
39. Kappas A, Krumhuber E, Küster D (2013) Facial behavior. In: Knapp ML (ed) Hall JA. Handbook of communication science Nonverbal communication. deGruyter, Berlin, pp 131–166
40. Kappas A, Pecchinenda A (1998) Zygomaticus major activity is not a selective indicator of positive affective state in ongoing interactive tasks. Psychophysiology 35:S44
41. Larsen JT, Norris CJ, Cacioppo JT (2003) Effects of positive and negative affect on electromyographic activity over zygomaticus major and corrugator supercilii. Psychophysiology 40(5):776–785. https://doi.org/10.1111/1469-8986.00078
42. Lazarus RS (1991) Progress on a cognitive-motivational-relational theory of emotion. Am Psychol 46(8):819–834. https://doi.org/10.1037/0003-066x.46.8.819
43. Lien JJJ, Kanade T, Cohn JF, Li CC (2000) Detection, tracking, and classification of action units in facial expression. Robot Auton Syst 31(3):131–146. https://doi.org/10.1016/s0921-8890(99)00103-7
44. Menne IM, Schnellbacher C, Schwab F (2016) Facing emotional reactions towards a robot. an experimental study using FACS. In: Agah A, Cabibihan J, Howard A, M MS, He H (eds) Social robotics. Springer, Cham, pp 372–381
45. Mirnig N, Strasser E, Weiss A, Kühnlenz B, Wollherr D, Tscheligi M (2014) Can you read my face? Int J Soc Robot 7(1):63–76. https://doi.org/10.1007/s12369-014-0261-z
46. Öhman A (2002) Automaticity and the amygdala: nonconscious responses to emotional faces. Curr Dir Psychol Sci 11(2):62–66
47. Prkachin KM (1992) The consistency of facial expressions of pain: a comparison across modalities. Pain 51(3):297–306. https://doi.org/10.1016/0304-3959(92)90213-u
48. Reed LI, Zeglen KN, Schmidt KL (2012) Facial expressions as honest signals of cooperative intent in a one-shot anonymous prisoners dilemma game. Evol Hum Behav 33(3):200–209. https://doi.org/10.1016/j.evolhumbehav.2011.09.003
49. Reeves B, Nass C (1996) The media equation. How people treat computers, television, and new media like real people and places. CSLI Publications, Stanford

50. Ribeiro T, Paiva A (2012) The illusion of robotic life. In: Proceedings of the seventh annual ACM/IEEE international conference on human–robot interaction—HRI 12. ACM Press, Boston. https://doi.org/10.1145/2157689.2157814
51. Riether N (2013) On the profoundness and preconditions of social responses towards social robots. Experimental investigations using indirect measurement techniques. Ph.D. thesis, Universität Bielefeld
52. Rosenthal-von der Pütten AM, Krämer NC, Hoffmann L, Sobieraj S, Eimler SC (2013) An experimental study on emotional reactions towards a robot. Int J Soc Robot 5(1):17–34. https://doi.org/10.1007/s12369-012-0173-8
53. der Pütten Rosenthal-von AM, Schulte FP, Schulte SC, Sobieraj S, Hoffmann L, Maderwald S, Brand M, Krämer NC (2014) Investigations on empathy towards humans and robots using fMRI. Comput Hum Behav 33:201–212. https://doi.org/10.1016/j.chb.2014.01.004
54. Russell JA, Fernández-Dols JM (1997) The psychology of facial expression. Cambridge University Press, New York
55. Sato W, Yoshikawa S (2007) Spontaneous facial mimicry in response to dynamic facial expressions. Cognition 104(1):1–18. https://doi.org/10.1016/j.cognition.2006.05.001
56. Scherer KR, Schorr A, Johnstone T (2001) Appraisal processes in emotion: theory, methods, research. Oxford University Press, New York
57. Takahashi Y, Hatakeyama M (2008) Fabrication of simple robot face regarding experimental results of human facial expressions. In: 2008 international conference on control, automation and systems. IEEE, Seoul, Korea. https://doi.org/10.1109/iccas.2008.4694495
58. Unz D, Schwab F, Winterhoff-Spurk P (2008) TV news—the daily horror? Media Psychol 20(4):141–155. https://doi.org/10.1027/1864-1105.20.4.141
59. Watson D, Clark LA, Tellegen A (1988) Development and validation of brief measures of positive and negative affect. The PANAS scales. J Pers Soc Psychol 54(6):1063–1070. https://doi.org/10.1037/0022-3514.54.6.1063
60. Wilcox RR (2011) Introduction to robust estimation and hypothesis testing. Academic Press, Waltham
61. Woods S, Walters M, Koay KL, Dautenhahn K (2006) Comparing human robot interaction scenarios using live and video based methods: towards a novel methodological approach. In: 9th IEEE international workshop on advanced motion control. IEEE, Istanbul. https://doi.org/10.1109/amc.2006.1631754

Isabelle M. Menne is a research assistant at the University of Würzburg, Department of Media Psychology. She graduated in Psychology (Master's degree) and is currently doing her Ph.D. on social robotics. She is interested in emotions, facial expressions and social, "emotional" robotics.

Frank Schwab is Professor for Media Psychology at the University of Würzburg. His research focuses on evolutionary and emotional aspects of entertainment media as well as the analysis of facial expressions.