Human Perception of Social Robot's Emotional States via Facial and Thermal Expressions

DENIS PEÑA and FUMIHIDE TANAKA, University of Tsukuba, Japan

Facial and thermal expressions can be used by humans to interpret emotions. While facial expressions can be a voluntary reaction, the change of temperature in the body is often not. Thus, a facial expression may not always be consistent with the emotional state that is expressed by the body temperature. This article aims to study the human perception of the emotional expression and the emotional state of a robot that simultaneously uses its face and body temperature. To this end, a robot, named TherMoody, which has the capability to change its body temperature from 10–55°C, was used. Then, 25 combinations (5x5: one facial expression and one thermal expression for anger, joy, fear, sadness, and neutral state) were evaluated by 15 participants. Thermal expressions were designed based on the metaphors of emotions related to temperature. The results show people tend to base their judgment of the robot's emotional expression exclusively in its facial expression, regardless of the change of its body temperature. However, there are combinations where the thermal expression predominates over the facial expression when judging the robot's emotional state. Thus, these combinations may produce the perception that the robot's emotional expression is, in fact, genuine, simulated, masked, or neutralized.

CCS Concepts: • **Human-centered computing** → **User studies**; *Haptic devices*; • **Computer systems organization** → *Robotics*;

Additional Key Words and Phrases: Robot expressions, emotion regulation, face, body temperature, social robot, and genuine/simulated/masked/neutralized emotions

ACM Reference format:

Denis Peña and Fumihide Tanaka. 2020. Human Perception of Social Robot's Emotional States via Facial and Thermal Expressions. *ACM Trans. Hum.-Robot Interact.* 9, 4, Article 26 (May 2020), 19 pages.
<https://doi.org/10.1145/3388469>

1 INTRODUCTION

The design of social robots commonly considers equipping them with channels to express their emotional state. These emotional expressions are important to improve their social interaction with humans. Moreover, its implementation is usually inspired and adapted from human beings and even other biological systems. Thus, the most commonly used channels are facial expressions, body movements, and tone of voice. These are modalities that stimulate the human sense of sight or hearing. Similar to the way humans express their emotions through various channels [4], several

This work was supported by JSPS KAKENHI Grant Number 17K19993 and 19H01112.

Authors' addresses: D. Peña, University of Tsukuba, 1-1-1 Tennodai, Tsukuba, Ibaraki, 305-8573, Japan; email: penia@ftl.iit.tsukuba.ac.jp; F. Tanaka, University of Tsukuba, 1-1-1 Tennodai, Tsukuba, Ibaraki, 305-8573, Japan; email: fumihide.tanaka@gmail.com.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

© 2020 Copyright held by the owner/author(s). Publication rights licensed to ACM.

2573-9522/2020/05-ART26

<https://doi.org/10.1145/3388469>

studies revealed that the perception and interpretation of the emotional expression of a social robot improve when it is multimodal [15, 30, 33, 45]. Even then, the use and study of modalities that stimulate the other human senses, such as the sense of touch, are less numerous [16, 41, 45]. Despite the various approaches and modalities mentioned above, Jung criticizes the fact that the perspective of emotional expression in most of the works in HRI is limited [18]. The reason is that emotional expressions and emotional state are assumed to be congruent.

In the case of humans, it is not only essential to express their emotions, but also how to regulate their expression in a way that is socially tolerable [9, 31]. Thus, the emotion expressed by a person is not always consistent with his/her emotional state. For instance, a person tries to suppress any expression of nervousness to have a better impression in a job interview. Therefore, social robots may also need to have the ability to regulate the expression of their emotions according to what a specific context requires. This may be crucial especially when it comes to negative emotions. For example, it may be rude to the user if a therapeutic robot, that is being stroked, expresses anger directly. In fact, humans have voluntary and involuntary reactions when feeling an emotion [13, 31, 42]. Examples of voluntary expressions are facial expressions, gaze, and body movements. Among them, facial expressions have the particularity of a high readability of emotions label. However, as a voluntary reaction, facial expressions may not always be consequent with the emotional state, as we discussed in the previous paragraph. Thus, by using facial expression, it is not only possible to express a genuine emotion, but also to simulate, mask, or neutralize it [9]. In contrast to voluntary expressions, an involuntary reaction is inherent to a physiological reaction. Thus, it is difficult to control it; moreover, it may be a cue or a reflector to perceive the emotional state. Examples of involuntary expressions are body temperature, heart rate, sweating. Among them, in this study, we chose the body temperature. We will explain the reason in detail in Section 3.

We expect that the perception of the voluntary/involuntary characteristic of the facial expression and the body temperature can be similarly interpreted when performed by a social robot. Therefore, this work aims to study how the emotional expression and the emotional state of the robot will be perceived when the robot simultaneously uses its face and body temperature to express its emotional state. To address this research question, a robot, TherMoody, that has the ability to control changes in body temperature was used. We then developed emotional expressions for TherMoody by using its face and body temperature. Finally, a user study was carried out to investigate participants' perception on this robot.

2 PREMISES

In this article, we use *emotional expression* (EE) to refer to the any perceptible sign of an *emotional state* (ES) [14, 21]; while *emotion regulation* to refer to the ability to regulate one's emotional expressions [31]. Then, we investigate the human perception of (1) the EE of a robot expressed by its face and the body temperature and (2) the ES of the robot. In particular, we discuss if combining the two channels can be a way to create the perception of the *emotion regulation* for the robot.

It has been discussed that there is no consensus in the literature on the definition of emotion [22], and this article also does not argue it. However, we need to mention that the human perception that is studied in this article can be affected by cultural and contextual factors [3, 10, 18, 32], which will be discussed in detail in Section 8.3.

3 THERMAL EXPRESSION AS AN EMOTIONAL CHANNEL FOR SOCIAL ROBOTS

The change of temperature in the human body when experiencing an emotion is a natural and involuntary physiological reaction [23]. For instance, a person who feels anger cannot avoid revealing his emotion through an increase in his body temperature. Even though the change of temperature is visually less evident, Nummenmaa et al. [36] mapped the human perception of body

Fig. 1. (a) Isometric view and (b) sections of the robot's body and the 3 DoF on the neck [38].

activation associated with different emotions. Although these maps are based on human perception rather than an explicit measurement, the physiological activation of an area of the body is directly related to the increase of its temperature. It was also reported that thermal stimuli could be interpreted as emotional information by humans [28, 40, 44]. Wilson et al. [44] mapped thermal stimuli onto models of emotion, providing information on how thermal stimuli are interpreted emotionally in the dimensions of arousal and pleasure.

The use of thermal expression for the purpose of an emotional channel in social robot has the following benefits: (1) *unaltered robot shape*: Applying temperature to the skin of the robot does not alter its form; (2) *privacy*: Only the person who interacts physically with the robot can feel the thermal stimulus; and (3) *non-disturbing*: Because people who do not physically interact with the robot will not perceive the thermal stimuli, it will not disturb them.

Despite these benefits, there have been made not many attempts in the use of thermal stimuli in HRI settings: Nie et al. [35] suggested that the human perception of friendship toward a robot increases with a warm temperature; Park and Lee [37] revealed that the human perception of a robot as a companion varies according to the temperature level of the robot's skin; Block and Kuchenbecker [1] evaluated human responses in robot hugs controlling its temperature and softness; and Willemsen et al. [43] studied a physical warmth factor in a remote communication setting by introducing an interactive teddy bear having a thermal stimulus. On the other hand, in the field of haptics research, thermal stimuli are used in haptic applications [2, 5, 6, 11, 12, 34]. However, no study has been conducted to investigate the emotion regulation (Section 2) in social robots using facial and thermal stimuli.

4 ROBOT PLATFORM

The platform used is TherMoody, a social robot specially developed to study facial and thermal expressions in a robot. TherMoody has the particular capability of dynamically changing the temperature of its body, specifically, in its head, from cold to hot temperatures. Aesthetically, the body base gives the robot an abstracted form of a pet (see Figure 1). Its design intends that a human intuitively caresses the head of the robot, as with a dog or a cat. To increase this likelihood, the design of the lower and upper limbs of the robot is excluded. The basic design of TherMoody was verified in a pilot study reported in [39]. Considering that the robot was designed so that its head is the main area to be caressed, its thermal system is located in it. Via this system, the head of the robot can achieve temperatures ranging from 10°C to more than 55°C. It is composed of a set of 12 thermoelectric modules and heat sinks (see Figure 2). In addition, there is a display acting as the robot's face, where different facial expressions can be rendered. The details of the design process

Fig. 2. The heat sink LAM4 by Fisher Elektronik (a) is cut into pieces (b), then, assembled to form the head structure (c). Finally, Peltier cells and a display are added (d) [38].

of the robot are described in [38], while a pilot study to validate the design is reported in [39]. For this study, we newly developed a closed-loop control for the temperature of the thermal system. To this end, a precision temperature controller TEC-1123 from Meerstetter Engineering was used, while the temperature measurement was conducted through the use of infrared thermometers.

5 DESIGN OF EMOTIONAL EXPRESSIONS

This section describes the design process of emotional expressions for TherMoody via two modalities: facial expressions and thermal expressions.

5.1 Target Emotional Expressions

To decide which emotional expressions to study, we take into account the categorical approach of the classification of emotional states. This approach divides emotional states into discrete primary and secondary classes. Primary emotional states are evolved adaptations, innate, and cross-culturally recognized. The most widely accepted classification is the one proposed by Ekman, which has joy, sadness, anger, fear, surprise, and disgust [8]. Because of their intuitive perception, the primary emotional states were well studied in the field of HRI [15, 30, 41]. Besides, Jack et al. [17] suggested that, based on the discrimination of dynamic facial expressions, the primary emotional states can be reduced to four categories: joy, sadness, fear/surprise, and anger/disgust. Among them, we establish out target emotional states to anger, fear, joy, and sadness. In addition to them, a neutral emotional state is also considered as a target. This state is perceived as one in which the robot is not performing an emotional expression.

5.2 Design of Facial Expressions

The initial design of the candidates for the facial expressions of each emotional state was based on the following literature. As a general guideline, this study took into account the distinctive facial clues that represent an emotional state. These are described by Ekman [9]. Similarly, to decide which features the face of the robot would have, such as mouth, eyebrow, iris, pupils, and so on, the impact on the human's perception of each of them, reported in [19], was considered. In addition, the design of the robot's face was made in such a way that it is in harmony with the hexagonal head of the robot as much as possible. Thus, it is possible to see in Figure 3 and Figure 4 that the eyebrows and the outline of the eyes are mainly composed of straight lines. Finally, since most of the robot's body has a white hue, it was decided that the color of the face would be white, while the features would be black to contrast. Furthermore, we intended that the facial expression was ideally the only channel to perceive the emotional states of the robot. Therefore, as there is

Fig. 3. The evolutionary design process of the angry facial expression.

Fig. 4. Final design of facial expressions for the five target emotional states.

literature that indicates that colors can be interpreted as emotional information [20, 25, 29], we avoided using or adding other colors in the design of facial expressions. Considering these aspects, one facial expression for each target emotional state was developed.

Continuing with our design work, we focused on improving the initial designs of facial expressions in such a way that they express the target emotional states as clearly as possible. To this end, we carried out an informal evolutionary design process with the sequential collaboration of

Table 1. Type and Intensity of Thermal Stimulus
Selected for Each Emotional Expression

Emotional Expression	Thermal Stimulus	
	Type	Intensity
Anger	Above	Hot
Joy	Neutral	Warm
Neutral	Neutral	Neutral
Sadness	Below	Cool
Fear	Neutral	Cold

ten graduate students (collaborators) who were recruited at the university that the authors of this article were affiliated with. First, the set of facial expressions was shown one-by-one to the initial collaborator. For each design, the collaborator was asked in an open-ended question to describe the perceived emotional state, detailing what features caused that perception. Second, once all the facial expressions were shown, it was revealed to the collaborator that each design intended to express a specific emotional state. Then, the collaborator was asked for recommendations on how to improve the expressiveness of each design. Thus, according to the description and the suggestions, each design evolved by maintaining, adjusting, or adding features to the facial expression. For some facial expressions, it was decided to evolve them into two candidates. This entire process of receiving feedback was iterated for the next collaborators. Thus, the set of facial expressions shown to the following collaborator included the outcome of the previous one.

For instance, Figure 3 shows the evolutionary design process of the angry facial expression. The labels above each image represent the images that were shown to different collaborators. Thus, only the image with the ID Number 1 was shown to the first collaborator, while the images with the ID Number 2 were shown to the second collaborator. The final candidates for the facial expressions, developed by the end of the evolutionary design process, are shown in Figure 4.

5.3 Design of Thermal Expressions

Temperature, compared to the facial expression, required a higher degree of abstraction to be interpreted as an emotional state. As reported in previous works, the interpretation of thermal stimuli as emotional information may be limited only to some emotional states [28, 44]. However, to design the thermal expressions, the cognitive-linguistics theory of conceptual metaphors was considered as a general guideline [23, 26, 29]. This framework proposes that a user's mental models can be expressed through figurative explanations. Specifically, Kövecses summarizes a set of emotion metaphors related to temperature [23]. Thus, anger is metaphorically conceptualized as "HOT FLUID IN A CONTAINER." Some casual expressions are: "he is boiling with anger" or "making one's blood boil." In the case of fear, it is conceptualized as "A FLUID IN A CONTAINER." In addition, Kövecses details that in some languages and cultures, fear is associated with cold rather than warm [24]. Thus, some derived expressions are: "to be frozen in one's tracks," "cold shivers running down one's spine," or "to get cold feet." On the other hand, joy is associated with "WARM" and sadness with a "LACK OF HEAT."

Based on the previous metaphors, it was decided to set a hot thermal stimulus for the thermal expressions of joy and anger, with anger's temperature higher than joy's temperature. The thermal expressions for fear and sadness were set as a cold thermal stimulus, with the temperature for fear colder than for sadness. In the case of a neutral expression, this will be established as a temperature that cannot be distinguished as cold or hot. Table 1 shows the thermal stimuli selected for each emotional expression. They are organized in descending temperature from top to bottom.

Fig. 5. Profile temperature for the final design of each thermal expressions.

As a starting point for the design of the temperature profile for each expression, we took into account the thermal stimuli mapped in the circumplex model of emotions by Wilson et al. [44]. Specifically, those stimuli with close values of arousal and valence to the target emotional states of this study. However, we made modifications to these temperature profiles once they were implemented in Thermody's thermal system. Specifically, we opted for a trapezoidal profile, composed by an attack (A → B), sustain (B → C), and release section (C → E); instead of the semi-trapezoidal profile used in [44], composed by an attack and sustain section (see Figure 5). This decision was taken because we intend to indicate the end of the emotional expression through the release section.

Despite the difference in the temperature profile, according to [44], a slow and small temperature change in the attack section is perceived as a low level of arousal. However, if it is a warm change, then it is generally "interpreted as slightly positive/pleasant (valence)." This guideline was used to set empirically the attack section for joy temperature profile. Conversely, if it is a cool change in the attack section, then it is interpreted as negative/unpleasant. Therefore, we use this information to establish the attack section for the sadness temperature profile. Moreover, regarding the type and intensity of the thermal stimulus for joy and sadness, there is a concordance between what was reported in [44] and the metaphors associated with these emotions.

On the other hand, a rapid and large temperature change are perceived with high arousal and as negative/unpleasant. This information was a guide to establish the rate of change and the extent of the attack section for the temperature profile of fear and anger. Then, the metaphors associated with these emotional states harmoniously complemented the information to set the direction (warming or cooling) of the attack section of each profile. Although we empirically perceived that a faster rate of change seems to be more suitable for the temperature profile of anger and fear, we should point out these values could only be established under the maximum values of warming (6.5°C/s @ 27°C) and cooling (-3.6°C/s @ 27°C) for the thermal system. Another case is that below 18°C , the cooling rate decreases dramatically. Therefore, it is not possible to set a greater change of temperature in the attack section of the fear temperature profile even if this might be more suitable for its design. Consequently, it will be inevitable that these technical limitations affect, especially the design of the fear temperature profile.

Similar to the design of the facial expressions, the initial candidates for thermal expressions passed through an evolutionary design process. The objective was also to increase the correct perception of each intended target emotional state. The collaborators, the given explanation, and the procedure used were the same as for the design work of the facial expressions. Although the

answers to the open-ended question to describe the perceived emotional state had high variability, suggestions to improve the expressiveness of the intended emotional states had a similar tendency. Through this design work, we noticed empirically that: (1) the gradient and the final temperature value of the attack section help in large proportion to discern the target emotional states; (2) a tiny gradient provides a better perception of the profile sequence than a constant temperature value; and (3) by subdividing the sustain section into two segments ($C \rightarrow D$ and $D \rightarrow E$), the end of the temperature profile is better anticipated.

Additionally, it was decided to set the same duration for all candidates. The purpose of this is to standardize the length of the subsequent combination between facial and thermal expression. Then, by the end of the evolutionary design process, the duration was set to 7 seconds. Collaborators often responded that a shorter duration hindered the emotional interpret of the stimuli. It should be emphasized that the temperature profile of each candidate does not pretend to be the ideal parameterization to express the target emotion, but rather intends to express it sufficiently to be recognized. Figure 5 shows the profile temperature for each final candidate of the thermal expressions. The temperature of the neutral expression was set at 27.5°C .¹ This value was chosen because, in this, the collaborators could not define the thermal stimulus as warm or cold.

5.4 Manipulation Check

Ten students, different from those who participated in the evolutionary design process, were recruited for a test to validate the five candidates for facial expressions and the five candidates for the thermal expressions. The participant received a brief explanation of the experiment. Here, with the robot powered off and its face covered with a piece of fabric, the experimenter provided the following scenario to the participant:

We, as human beings, express our emotions in different ways. For example, we realize that a person is anxious or nervous about the expression on his face. But his emotion can also be expressed in other parts of his body, such as shaking his leg continuously or being sweating on his forehead or in the palm of his hand. That is in the case of humans. Now, let me introduce you to TherMoody. It is a robot capable of expressing its emotions by facial expressions and also by changing the temperature of its body. I know, it may seem strange that a robot expresses its emotion by the temperature of its body, but in fact, TherMoody can do it somehow. So, this time we will play a variation of the social game of the charades: TherMoody will express an emotion through its face or its body temperature, and your objective will be to discover which emotion it was. First, TherMoody will express different emotions only through its body temperature, and then it will do it only through its face.

The experimenter absolved any doubt from the participant regarding the experiment. Once everything was clear, the participant was exposed to the maximum and minimum temperature that the robot can achieve. Thus, they were aware of the extreme temperature values before starting the evaluation of the thermal expression. Besides, to ensure that the whole thermal expression was perceived, participants received instructions on when to touch and stop touching the robot.

¹The value of 27.5°C was empirically chosen. However, it has a scientific basis. Heat transfer always occurs in a thermal gradient from hot to cold. Since humans are often the hottest objects in a given environment, the normal direction of heat transfer is from the body to the environment. Even at the same room temperature (for example, 20°C), touching a piece of metal (which has excellent thermal conductivity) can be perceived as a cold object, while touching a piece of wood (which has a low thermal conductivity) as less cold or as a neutral temperature. This perception is due to the rate of heat transfer to each object. In our case, people described the surface of the thermal system as a cold surface when it was set at room temperature, while when it was set at the temperature of their hands, it was described as slightly warm.

Fig. 6. (a) Setup to evaluate the thermal expressions. Participants touch the robot's head but cannot see its face. Thus, the robot's facial expression does not influence this assessment. (b) Setup to evaluate the facial expressions. Participants see the robot's facial expression but cannot touch its body. Thus, the robot's thermal expression does not influence this assessment.

Table 2. Average Perception of the Emotional Expression (min. = 0 (0% agree), max. = 4 (100% agree))

Modality	Emotional Expression	Emotional Perception						
		Fear	Sadness	Neutral	Joy	Anger	$F(4, 36)$	η_p^2
Facial Expression	Fear	3.7	0.8	0.1	0.1	0.3	70.80	0.89
	Sadness	0.6	3.6	0	0	0.1	212.47	0.96
	Neutral	0.7	0.4	3.7	0.2	0.2	121.42	0.93
	Joy	0.1	0.1	0.3	3.7	0.1	172.36	0.95
	Anger	0.3	0.2	0.1	0	3.8	290.28	0.97
Thermal Expression	Fear	3.4	1.7	0.1	0.1	0.4	46.09	0.84
	Sadness	1.6	3.4	0.1	0.1	0.4	84.80	0.90
	Neutral	0.4	0.2	3.7	0.2	0.2	106.40	0.92
	Joy	0.0	0.0	0.3	3.3	0.9	94.14	0.91
	Ange	0.1	0.2	0.2	1	3.5	63.61	0.88

Averages statically significantly different ($p < 0.05$) from others are displayed in bold.

Then, all thermal stimuli were presented to the participant in random order. During this process, the robot did not move its head while its face was covered with a piece of fabric, as shown in Figure 6(a). The objective was that the assessment of the thermal expressions was isolated from the influence of the facial expression of the robot. After each stimulus, the participant was asked to interpret what emotional state the robot expressed. The question had five forced-choice options (anger/fear/joy/sadness/neutral). The answer was collected through a 5-point Likert scale ranging from 0 = "0% agree" to 4 = "100% agree." The participant could choose multiple answers for each stimulus. Once all the thermal expressions were evaluated, the face of the robot was uncovered.

Subsequently, each facial expression was presented in a random order to the participant. During this process, the participant did not touch the robot in order to be stimulated only by its facial expression (see Figure 6(b)). The question and Likert scale were the same as the evaluation of the thermal expressions.

The average perception of each emotional expression in each stimulus is shown in Table 2. A repeated-measures ANOVA was conducted. In addition, a post-hoc Turkey test was applied; those averages statistically significant are displayed in bold. This test showed that each of the candidates in both modalities was perceived with its respective target emotional expression. Thus, we concluded that all the candidates were correctly validated.

Fig. 7. Combination over the time of the facial expression of fear with each of the thermal expressions.

6 USER STUDY

A lab-based user study was carried out to investigate the human perception of the robot's facial and thermal expressions that are used simultaneously to realize the emotion regulation that was explained in Section 2. To provide further analysis of the results, the following terms that were introduced by Ekman [9] to define the relationship between one's emotional expression and one's emotional state are taken as reference.

Neutralized Emotion: The emotional state is avoided to be expressed. For instance, a person in a job interview avoiding expressing nervousness.

Simulated Emotion: There is an emotional expression without feeling any emotional state. For instance, a mother who, without being angry, expresses anger only as a signal to prevent or disapprove of her son's misbehavior.

Genuine Emotion: The emotional state is expressed. For instance, a happy person expressing his happiness.

Masked Emotion: The emotional state is covered by expressing a different emotional expression. For instance, a terrified child expressing fury as a technique to stop being bullied.

6.1 Participants

Fifteen participants from different nationalities studying for a Master's degree at the University of Tsukuba were recruited: five Peruvians, four Japanese, three Chileans, two Mexicans, and one Chinese. The average age was 27.07 years old ($SD = 3.17$). All of them were right-handed and without a physical disability (60.0% male). Before the experiment, the temperature of their right hand was measured ($M = 30.53^{\circ}\text{C}$, $SD = 1.62$). The study was approved by the ethical committee of the University of Tsukuba (2017R166-1), and consent was obtained from all the participants.

6.2 Combination of Unimodal Expressions

The two emotional modalities, facial and temperature, were combined for the five target emotional states: anger, fear, sadness, joy, and neutral state. This produces a total of 25 combinations ($5 \text{ facial} \times 5 \text{ thermal expressions}$). The duration of the combination of both emotional modalities was set as the same as the duration of the thermal expressions. To this end, a 7-second animation was implemented for each facial expression. These animations have three sections: (1) attack: begins on a neutral face and ends at the maximum facial expressiveness of the target emotional state; (2) sustain: keeps the facial expression; and (3) release: returns to the neutral face (see Figure 7).

Fig. 8. Experimental setup to evaluate the combination of both modalities.

To define the duration of the attack and release sections, the rounded average duration of the first and second line of the temperature profile of each thermal expression was considered. Thus, the attack section is from 0 to 2.45 seconds; release, from 2.45 to 4.8 seconds; and sustain, from 4.8 to 7 seconds. This means that, although both unimodal expressions start and end at the same time, they do not have a perfect fit in their profile. Figure 7 illustrates the combination over time of the facial expression of fear with all the thermal expressions.

6.3 Procedure

The experiment was conducted in a location similar to a sitting room (see Figure 8). The ambient conditions of the experiment room were: 20°C of temperature and 40% relative humidity. The participant was welcomed and an explanation of the purpose and the procedure of the study was given. The experimenter then provided an introductory explanation to emphasize the difference between a emotional expression and the emotional state. An example for each strategy of emotional regulation was given; however, the experimenter intentionally avoided the use of the terms: genuine, neutralize, mask and simulate; as well as the definition of these terms related to the emotion regulation. The instruction text was as follows:

We, as human beings, express our emotions in different ways: through our face, perhaps shaking the legs when one is anxious, a fast heart rate when we are afraid, and so on. However, what we express is not necessarily what we feel. Just imagine a person in a job interview. He may be trying to hide his nervousness by showing a relaxed or neutral body posture, so that he can generate a good impression. A similar case is when a mother who, without being angry, expresses anger only as a signal to prevent or disapprove of his son's misbehavior. Another example could be a person who feels sad but behaves in a happy way to prevents his family from worrying about him. As you see, a person can express an emotion, but that expression is not necessarily the emotion this person feels.

After this initial explanation, the experimenter introduced the participant to the robot. It was explained that the robot is able to express its emotional state through two modalities. Then, the experimenter provided the following scenario to the participant:

Now, let me introduce you to TherMoody. It is a robot capable of expressing its emotions simultaneously by facial expressions and also by changing its body temperature

on its head. I know, it may seem strange that a robot expresses its emotion by the temperature of its body, but in fact, TherMoody somehow can do it. So, this time we will play a variation of the social game of the charades: TherMoody will express an emotion through its face and its body temperature simultaneously on 25 occasions. Your objective is to discover which emotional state TherMoody is in and which emotional expression it uses in each case. For that, you have the following questionnaire.

In each trial of the 25 combinations (5 facial expression $\times$ 5 thermal expression), the participant was asked to answer two main questions to evaluate the perception of the emotional expression and emotional state of the robot:

Q1: What emotion was expressed by the robot?

Q2: What was the feeling of the robot?

Both questions had the following forced-choice options: a) It was neutral; b) It was anger; c) It was fear; d) It was joy; and, e) It was sadness. The level of agreement with each option was collected via a 5-point Likert scale from 0 = “0% agree” to 4 = “100% agree”. In addition, it was optional to write a comment for each stimulus. After explaining the questionnaire, the experimenter absolved any doubt from the participant regarding the experiment, emphasizing the difference between Q1 and Q2. Once everything was clear, the stimuli were presented in a random order for each participant. Before starting the evaluation of the 25 combinations, participants were exposed to the maximum and minimum temperature that the robot can achieve. This step was carried out to make participants aware of the extreme temperature values. To ensure that the whole thermal emotional expression was perceived, participants received instructions on when to touch and stop touching the robot. The robot maintained the same posture during the experiment. Although each stimulus has a duration of 7 seconds, the time between stimuli was 3 minutes. In this period, the participant not only evaluated and gave the comments related to the current stimulus, but it was also a prudent time to ensure that the participant’s thermoreceptors recover from the thermal excitation, and also a time for the experimenter to configure by software the next stimulus in the thermal system. Besides, there were two main breaks of 15 minutes after evaluating the tenth and twentieth combination. The entire experiment, including the introductory explanation, lasted about 2.5 hours.

7 RESULTS

Table 3 shows the average perception of the emotional expression (EE) and the emotional state (ES) of the robot in the 25 combinations. In the first column, an ID was assigned to each combination from C1 to C25. The last column shows our interpretation of those results according to the definition of genuine, simulated, masked, and neutralized emotion. The data of each combination were analyzed by using repeated-measures ANOVA. Cells in a row with the same color (white, light yellow, or dark yellow) do not have a statistically significant difference.

In general, for the 25 combinations, the results show that subjects based their perception of the emotional expression of the robot on the facial expression independently of any thermal expression. However, the perception of the emotional state of the robot is not always the same as the emotion expressed via the face. In fact, in some cases, their perception is directly judged according to the thermal emotional expression. Taking as reference the definition of the display rules (neutralized, simulated, genuine, and masked) mentioned in Section 6, we will discuss the results in the following sections:

Table 3. Average Perception of the Emotional Expression (EE) and the Emotional State (ES) of the Robot (min. = 0 (0% agree), max. = 4 (100% agree))

ID	COMBINATION		EMOTIONAL EXPRESSION (EE)							EMOTIONAL STATE (ES)							Interpretation of EE vs ES
	Face	Body	Fe	Sa	Ne	Jo	An	F(4, 56)	η_p^2	Fe	Sa	Ne	Jo	An	F(4, 56)	η_p^2	
C1	Fear	Fear	3.73	0.53	0.00	0.00	0.00	100.64	0.88	3.73	1.07	0.00	0.00	0.00	60.08	0.81	Genuine
C2	Fear	Sad	3.53	1.07	0.13	0.00	0.00	59.71	0.81	3.47	1.13	0.00	0.00	0.00	58.89	0.81	Genuine
C3	Fear	Neu.	3.47	0.60	0.20	0.00	0.07	85.85	0.86	1.73	0.27	2.80	0.00	0.00	25.78	0.65	Simulated
C4	Fear	Joy	3.33	0.27	0.00	0.13	0.13	118.45	0.89	2.07	0.07	0.00	0.40	1.47	13.57	0.49	Gen/Mask
C5	Fear	Angry	3.60	0.47	0.00	0.20	0.20	117.33	0.89	2.07	0.20	0.20	0.27	1.8	8.60	0.38	Gen/Mask
C6	Sad	Fear	0.73	3.87	0.00	0.00	0.07	213.93	0.94	1.87	3.87	0.00	0.00	0.13	56.17	0.8	Genuine
C7	Sad	Sad	0.60	3.80	0.00	0.00	0.00	247.29	0.95	1.40	3.87	0.00	0.00	0.07	163.72	0.92	Genuine
C8	Sad	Neu.	0.40	3.47	0.00	0.00	0.13	208.35	0.94	0.20	1.60	2.67	0.20	0.20	34.70	0.71	Simulated
C9	Sad	Joy	0.53	3.40	0.00	0.07	0.33	58.74	0.81	0.40	1.73	0.00	0.20	1.93	8.74	0.38	Gen/Mask
C10	Sad	Angry	0.67	3.33	0.07	0.33	0.47	28.93	0.67	0.53	2.07	0.00	0.33	1.80	8.92	0.39	Gen/Mask
C11	Neu.	Fear	1.40	1.00	3.27	0.00	0.00	30.60	0.69	3.33	1.73	0.33	0.07	0.00	57.42	0.8	Neutralized
C12	Neu.	Sad	0.73	0.93	3.20	0.13	0.07	21.60	0.61	1.73	2.80	0.53	0.07	0.00	23.12	0.62	Neutralized
C13	Neu.	Neu.	0.20	0.13	3.80	0.07	0.13	232.25	0.94	0.27	0.27	3.93	0.07	0.07	354.54	0.96	Emotionless
C14	Neu.	Joy	0.20	0.13	3.13	0.67	0.53	39.21	0.74	0.13	0.00	0.27	2.93	1.53	28.19	0.67	Neutralized
C15	Neu.	Anger	0.07	0.00	3.40	0.13	1.07	89.01	0.86	0.00	0.00	0.67	0.73	3.47	48.13	0.77	Neutralized
C16	Joy	Fear	0.53	0.33	0.40	3.00	0.00	24.03	0.63	2.60	0.87	0.00	1.00	0.13	11.29	0.45	Masked
C17	Joy	Sad	0.67	0.53	0.07	3.33	0.00	41.83	0.75	2.87	1.73	0.07	0.60	0.07	22.67	0.62	Masked
C18	Joy	Neu.	0.00	0.07	0.60	3.47	0.00	152.82	0.92	0.07	0.07	2.87	1.87	0.00	34.72	0.71	Simulated
C19	Joy	Joy	0.00	0.00	0.20	3.80	0.00	455.54	0.97	0.00	0.07	0.00	3.47	0.93	46.68	0.77	Genuine
C20	Joy	Angry	0.13	0.00	0.13	3.60	0.27	152.90	0.92	0.07	0.00	0.00	2.73	1.47	19.45	0.58	Genuine
C21	Angry	Fear	0.73	0.47	0.00	0.13	3.47	48.20	0.77	2.87	1.47	0.00	0.07	1.40	14.44	0.51	Masked
C22	Angry	Sad	0.60	0.27	0.00	0.00	3.27	84.41	0.86	2.53	1.33	0.00	0.00	1.27	12.99	0.48	Masked
C23	Angry	Neu.	0.27	0.20	0.20	0.13	3.27	117.93	0.89	0.00	0.07	2.60	0.07	1.67	40.15	0.74	Simulated
C24	Angry	Joy	0.07	0.07	0.00	0.07	3.60	315.17	0.96	0.07	0.07	0.13	0.07	3.67	227.94	0.94	Genuine
C25	Angry	Angry	0.07	0.00	0.13	0.13	3.93	330.31	0.96	0.00	0.00	0.13	0.27	3.87	277.92	0.95	Genuine

The first column indicate the ID of the combination, while the last one shows the interpretation of the results according to the definition of genuine, simulated, masked, and neutralized emotion [9]. Fe, Sa, Ne, Jo, An and η_p^2 denote Fear, Sadness, Neutral, Joy, Anger, and the effect size, respectively. Cells in a row with the same color do not have a statistically significant difference ($p < 0.05$).

7.1 Combinations Potentially Perceived as Neutralized Emotions

C11, C12, C14, and C15 have a neutral facial expression combined with a thermal expression. The results revealed that the facial expression was used to judge the emotional expression (EE) of the robot. Nevertheless, the thermal expression was decisive in the judgment of the emotional state (ES) of the robot. For instance, in C15, participants perceived the robot as angry even though it did not express it facially. In these cases, it appeared that the participants perceived ES even with the neutral facial expression. Therefore, these combinations were potentially perceived as cases in neutralized emotions.

7.2 Combinations Potentially Perceived as Simulated Emotions

C3, C8, C18, and C23 are combinations of a facial expression with a neutral thermal expression. The results revealed that in order to perceive the ES of the robot, the facial expression had a significant role. Nevertheless, the perception with the highest values for the ES were those of a neutral state. This indicates that the judgment of the ES was based more on the thermal expression than on the facial expression. Here, it appeared that the participants perceived a neutral ES on non-neutral EEs. Therefore, these combinations were potentially perceived as cases in simulated emotions.

7.3 Combinations Potentially Perceived as Genuine Emotions

C1, C2, C6, C7, C19, C20, C24, and C25 are the cases where both EE and ES were perceived as having the same emotion label. Among them, C2, C6, C20, and C24 are combinations with their modalities inducing different emotion labels; however, the type of their thermal expression is compatible metaphorically with their facial expression (see Table 1). This could reflect the strong presence of the metaphors in the participants' minds when they judged the ES of the robot. As a result, since all those combinations can be interpreted as the robot was expressing its ES, they were potentially perceived as cases in genuine emotions.

7.4 Combinations Potentially Perceived as Masked Emotions

C16, C17, C21, and C22 are the cases where EE and ES were perceived with one predominant emotion label, although different from each other. C16 and C17 are combinations of the facial expression of joy with cold and cool thermal stimulus, respectively. Both of them were perceived as if the robot expressed joy but actually felt fear. This can be interpreted as the robot was covering its fear emotion by expressing joy. Making an analogy with the human case, these combinations might have been perceived as a scared smile or a nervous smile. Then, it is also possible to interpret that C21 and C22 were perceived as the robot was hiding its fear under an angry expression. In fact, a study reported that a person who feels fear due to provocations or threats might produce an expression of anger to face and cope with the situation [9]. For example, a person who feels fearful because he is the victim of bullying could express anger to mask his vulnerability. In summary, since all those combinations can be interpreted as the robot was covering its ES by a different EE, they were potentially perceived as cases in masked emotions.

7.5 Other Combinations

C4, C5, C9, and C10 are the cases where the perception of ES has two emotional categories with a high rating. We speculate it was due to: (1) inconsistencies between the modalities made the participant's interpretation confusing or meaningless. Consequently, because the questions were mandatory, some participants could have judged the robot's ES based on the facial expression, the body temperature, or both; or (2) the participant could interpret these combinations but with an emotion label that was not included in the forced-choice options. Therefore, the participant tried

Fig. 9. Possible emotional regulations concerning anger.

to represent his interpretation by decomposing it into the given options. In fact, it is reflected in the comments of some participants. For instance, some participants interpreted C4 and C5 as a nervous emotion, while C9 and C10 as frustration. Lastly, C13 is the combination where the robot was perceived as being emotionless or expressing nothing.

7.6 Summary of the Results

Overall, we observed that facial expression was perceived as the robot's EE in all the combinations. In contrast, the thermal expression was perceived as the robot's ES in the following cases:

- *Combinations with a neutral facial expression:* Potentially perceived as neutralized emotions.
- *Combinations with a neutral thermal expression:* Potentially perceived as simulated emotions.
- *Combinations with both modalities inducing the same emotion label:* Potentially perceived as genuine emotions.

The thermal expression was not perceived as the robot's ES in combinations with both modalities inducing different emotion labels. Here, it is necessary to consider the type and intensity of the thermal expression described in Table 1.

- If the type of thermal expression was metaphorically compatible with the emotion expressed by the face, then the perception of ES was that expressed through the face. Consequently, these combinations were potentially perceived as genuine emotions.
- If the type of thermal expression was metaphorically incompatible with the emotion expressed by the face, then the combination was potentially perceived as masked or blended emotion.

As a way to visualize the results of Table 3, Figure 9 illustrates the possible emotional regulations concerning anger.

8 GENERAL DISCUSSION

8.1 Location of the Thermal System in TherMoody

The main motivation to locate the thermal system of TherMoody in its head was that this area, in the case of a pet, is one of the main to be caressed. Thus, the transmission of thermal expressions was limited exclusively to this area. However, this may not be the best place to interpret thermal expressions. In fact, it remains to be explored whether the emotional interpretation of these expressions varies according to the part of the body of the robot that produced them. A case of

special interest is the robot's chest, a place where its heart would be located. This is because, at least colloquially, the heart is strongly linked to emotions.

8.2 Contextual Factors in Emotional Expression

The user evaluation is focused on the perception of 25 emotional expressions of a robot Ther-Moody. To perceive them, participants were instructed when to touch the robot's head. This physical interaction was forced; therefore, it was not generated naturally by an emotional reason. Besides, the location where this interaction occurs is the ideal to perceive the Thermoody's thermal stimulus because it is always forced to occur in the robot's head. Lastly, we directly told the participants that the temperature is an emotional expression. With this, we tried to prevent the participants from interpreting this stimulus as a robot's physical condition (fever, alive, feeling cold), robot's malfunction (over-heating), and the like. Recent literature in HRI research [10, 18] discuss the context- and culture-dependency of emotional expression. These are surely the cases with the robot that we discussed in this article. Therefore, we need to interpret the results considering these limitations.

8.3 Cultural Factors in Emotional Expressions

Cultural factors have also been discussed for the emotional expression of social robots [10, 27, 32]. Our design of thermal stimuli used conceptual metaphors related to emotional states. Since metaphors are linguistic expressions, they can be shaped by language. This factor can limit the universality in the conceptualization of emotions. On the other hand, Kövecses suggests that some metaphors emerged "from universal bodily experience" and are potentially universal [24]. Similarly, many cross-cultural metaphors are reported in different languages [23, 24].

In our study, participants were native speakers of Spanish, Japanese, or Chinese. Among these languages, there are similar expressions related to a type of temperature to describe, at least, the emotions evaluated in this work: anger, fear, joy, sadness. To mention some examples for the emotional state of anger in Spanish: "Me hierve la sangre de ira" (my blood boils), in Japanese: "Ikari ga hara no soko wo guragura saseru" (anger boils the bottom of stomach), and in Chinese: "Nu huo zhong shao" (anger fire inside burn, which can be interpreted as anger is burning inside me). In the case of fear in Spanish: "Quedarse helado/congelado de miedo" (got frozen in fear), in Japanese: "Kimo ga hietta" (liver gets cold, expression after a fear situation), and in Chinese: "Shou jiao bing liang" (hand feet ice cold, which can be interpreted as: both my hand and foot get frozen).

8.4 Other Limitations

The results in the validation of the design of each emotional expression, as well as those obtained in the main user study, could have been affected by the forced-choice paradigm in the questionnaire. Since participants must respond their perception in each emotional expression, they could have used the exclusion principle in case their perception did not clearly match with our target emotion labels.

In this study, we did not use explicit subjective scales to measure the participants' perception of the robot's emotion regulation. We did so because the measurement seemed to be less intuitive for participants. However, instead, the methodology we took made the interpretation of the results speculative.

8.5 Practical Use Cases

By leveraging the capability of emotional regulation, a robot can react to an event by performing a complex emotional behavior. This could be useful especially for those robots designed to interact over the time with the same user, such as a pet robot, a companion robot, or a therapeutic robot.

For instance, a pet robot that is being brusquely caressed needs to react in some way so that the human realizes its annoyance (lowest intensity of anger [7]). Considering that a physical HRI is occurring, an unregulated angry face can be interpreted in a rude and hostile way. Consequently, the empathy between the human and the robot could be diminished. In contrast, a sympathetic and gentle way could be achieved by expressing neutralized anger (C15: Neutral face + Anger body).

In fact, TherMoody could regulate its emotional reaction to the same event in different ways. In other words, emotional behavior could be consistent with the personality of the robot. For example, consider the context of the first physical interaction between a pet robot and a new user. As the user is still a stranger, a shy robot could express a neutralized fear (C11: Neutral face + fear body) as a sign of its cowardice to interact. However, an unsociable robot that is unwilling to be touched could express its hostility by performing a neutralized anger (C15: Neutral face + Anger body).

Another use case is with a companion robot capable of imitating some human behaviors. For example, some parents resort to simulating an expression of anger as a warning or disapproval towards the misbehavior of their children. In this case, a robot can recreate this action using an expression of simulated anger (C23: Angry face + Neutral body). This behavior could be useful as a sign of prohibition when, for example, the user is involuntarily stroking near the 'intimate part' of the robot.

9 CONCLUSION

This work focused on the investigation of facial and thermal expressions as emotional modalities for a social robot. Specifically, we attempted to study the perception of the emotional expression and the emotional state of the robot that simultaneously used its face and body temperature. To address this research question, the robot TherMoody, able to change the temperature of its body from 10–55°C, was developed. After the validation of the interpretation of each unimodal expression, the 25 combinations of both emotional modalities were evaluated by a user study. The results showed that the facial expression took prominence on the judgment of the emotional expression of the robot for all the combinations, which suggests that people rely more on visual stimuli than thermal stimuli to assess the emotional expression of the robot. However, in the case of judging the emotional state of the robot, the change of body temperature played an important role. Therefore, by using the combination of facial and thermal expressions, we could design some *emotion regulation* capability for a social robot to let its user have his/her perception of the robot's emotion to be genuine, simulated, masked, or neutralized. This capability may contribute to the design of a social robot that interacts with the user physically for a sustainable period of time.

ACKNOWLEDGMENTS

We thank the editors and the reviewers for their valuable comments and suggestions on the prior draft of this work. A special thanks to Kosmech Workshop of the University of Tsukuba for the support in the manufacture of the robot.

REFERENCES

- [1] Alexis E. Block and Katherine J. Kuchenbecker. 2018. Emotionally supporting humans through robot hugs. In *Proceedings of the Companion of the 2018 ACM/IEEE International Conference on Human-Robot Interaction*. ACM, 293–294. DOI : <https://doi.org/10.1145/3173386.3176905>
- [2] John-John Cabibihan and Sushil Singh Chauhan. 2017. Physiological responses to affective tele-touch during induced emotional stimuli. *IEEE Transactions on Affective Computing* 8, 1 (2017), 108–118. DOI : <https://doi.org/10.1109/TAFFC.2015.2509985>
- [3] Carlos Crivelli and Alan J. Fridlund. 2019. Inside-out: From basic emotions theory to the behavioral ecology view. *Journal of Nonverbal Behavior* 43 (2019), 161–194. DOI : <https://doi.org/10.1007/s10919-019-00294-2>
- [4] Charles Darwin. 1965. *The Expression of the Emotions in Man and Animals*. The University of Chicago Press.

- [5] Carl DiSalvo, Francine Gemperle, Jodi Forlizzi, and Elliott Montgomery. 2003. The hug: An exploration of robotic form for intimate communication. In *Proceedings of the 12th IEEE International Workshop on Robot and Human Interactive Communication*. IEEE, 403–408. DOI : <https://doi.org/10.1109/ROMAN.2003.1251879>
- [6] Chris Dodge. 1997. The bed: A medium for intimate communication. In *Proceedings of the CHI'97 Extended Abstracts on Human Factors in Computing Systems*. ACM, 371–372. DOI : <https://doi.org/10.1145/1120212.1120439>
- [7] Paul Ekman [n.d.]. *The Ekman's Atlas of Emotions*. <http://atlasofemotions.org>.
- [8] Paul Ekman. 1992. An argument for basic emotions. *Cognition and Emotion* 6, 3–4 (1992), 169–200. DOI : <https://doi.org/10.1080/02699939208411068>
- [9] Paul Ekman and Wallace V. Friesen. 1975. *Unmasking the Face: A Guide to Recognizing Emotions from Facial Clues*. Prentice-Hall.
- [10] Kerstin Fischer, Malte F. Jung, Lars Christian Jensen, and Maria Vanessa aus der Wieschen. 2019. Emotion expression in HRI: When and why. In *Proceedings of the 14th ACM/IEEE International Conference on Human-Robot Interaction*. IEEE, 29–38. DOI : <https://doi.org/10.1109/HRI.2019.8673078>
- [11] Daniel Gooch and Leon Watts. 2010. Communicating social presence through thermal hugs. In *Proceedings of the 1st Workshop on Social Interaction in Spatially Separated Environments*.
- [12] Daniel Gooch and Leon Watts. 2012. YourGloves, hothands and hotmits: Devices to hold hands at a distance. In *Proceedings of the 25th Annual ACM Symposium on User Interface Software and Technology*. ACM, 157–166. DOI : <https://doi.org/10.1145/2380116.2380138>
- [13] Mitchell S. Green. 2007. *Self-Expression*. Oxford University Press.
- [14] James J. Gross. 1999. Emotion regulation: Past, present, future. *Cognition and Emotion* 13, 5 (1999), 551–573. DOI : <https://doi.org/10.1080/026999399379186>
- [15] Markus Häring, Nikolaus Bee, and Elisabeth André. 2011. Creation and evaluation of emotion expression with body movement, sound and eye color for humanoid robots. In *Proceedings of the 20th IEEE International Symposium on Robot and Human Interactive Communication*. IEEE, 204–209. DOI : <https://doi.org/10.1109/ROMAN.2011.6005263>
- [16] Yuhan Hu, Zhengnan Zhao, Abheek Vimal, and Guy Hoffman. 2018. Soft skin texture modulation for social robotics. In *Proceedings of the 2018 IEEE International Conference on Soft Robotics*. IEEE, 182–187. DOI : <https://doi.org/10.1109/ROBOSOFT.2018.8404917>
- [17] Rachael E. Jack, Oliver G. B. Garrod, and Philippe G. Schyns. 2014. Dynamic facial expressions of emotion transmit an evolving hierarchy of signals over time. *Current Biology* 24, 2 (2014), 187–192. DOI : <https://doi.org/10.1016/j.cub.2013.11.064>
- [18] Malte F. Jung. 2017. Affective grounding in human-robot interaction. In *Proceedings of the 2017 ACM/IEEE International Conference on Human-Robot Interaction*. ACM, 263–273. DOI : <https://doi.org/10.1145/2909824.3020224>
- [19] Alisa Kalegina, Grace Schroeder, Aidan Allchin, Keara Berlin, and Maya Cakmak. 2018. Characterizing the design space of rendered robot faces. In *Proceedings of the 2018 ACM/IEEE International Conference on Human-Robot Interaction*. ACM, 96–104. DOI : <https://doi.org/10.1145/3171221.3171286>
- [20] Naz Kaya and Helen H. Epps. 2004. Color-emotion associations: Past experience and personal preference. In *Proceedings of the AIC 2004 Color and Paints, Interim Meeting of the International Color Association*. AIC, 31–34.
- [21] Eileen Kennedy-Moore and Jeanne C. Watson. 2001. *Expressing Emotion: Myths, Realities, and Therapeutic Strategies*. Guilford Press.
- [22] Paul R. Kleinginna and Anne M. Kleinginna. 1981. A categorized list of emotion definitions, with suggestions for a consensual definition. *Motivation and Emotion* 5 (1981), 345–379. DOI : <https://doi.org/10.1007/BF00992553>
- [23] Zoltán Kövecses. 2000. *Metaphor and Emotion: Language, Culture, and Body in Human Feeling*. Cambridge University Press.
- [24] Zoltán Kövecses. 2005. *Metaphor in Culture. Universality and Variation*. Cambridge University Press. DOI : <https://doi.org/10.1017/CBO9780511614408>
- [25] Rikard Küller, Byron Mikellides, and Jan Janssens. 2009. Color, arousal, and performance – A comparison of three experiments. *Color Research & Application* 34, 2 (2009), 141–152. DOI : <https://doi.org/10.1002/col.20476>
- [26] George Lakoff and Mark Johnson. 1980. *Metaphors We Live By*. University of Chicago Press.
- [27] Hee Rin Lee and Selma Sabanović. 2014. Culturally variable preferences for robot design and use in South Korea, Turkey, and the United States. In *Proceedings of the 2014 ACM/IEEE International Conference on Human-Robot Interaction*. ACM, 17–24. DOI : <https://doi.org/10.1145/2559636.2559676>
- [28] Wonjun Lee and Youn-Kyung Lim. 2012. Explorative research on the heat as an expression medium: Focused on interpersonal communication. *Personal and Ubiquitous Computing* 16, 8 (2012), 1039–1049. DOI : <https://doi.org/10.1007/s00779-011-0424-y>
- [29] Diana Löffler. 2017. *Color, Metaphor and Culture - Empirical Foundations for User Interface Design*. Ph.D. Dissertation. Julius-Maximilians-Universität Würzburg, Germany. <https://nbn-resolving.org/urn:nbn:de:byb:20-opus-153782>.

- [30] Diana Löffler, Nina Schmidt, and Robert Tscharn. 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*. ACM, 334–343. DOI: <https://doi.org/10.1145/3171221.3171261>
- [31] David Matsumoto and Linda Juang. 2016. *Culture and Psychology* (6th Edition). Cengage Learning.
- [32] Peter E. McKenna, Ayan Ghosh, Ruth Aylett, Frank Broz, and Gnanathusharan Rajendran. 2018. Cultural social signal interplay with an expressive robot. In *Proceedings of the 18th International Conference on Intelligent Virtual Agents*. ACM, 211–218. DOI: <https://doi.org/10.1145/3267851.3267905>
- [33] François Michaud, Paolo Pirjanian, Jonathan Audet, and Dominic Létourneau. 2000. Artificial emotion and social robotics. In *Distributed Autonomous Robotic Systems 4*. Springer, Tokyo, 121–130. DOI: https://doi.org/10.1007/978-4-431-67919-6_12
- [34] Hideyuki Nakanishi, Kazuaki Tanaka, and Yuya Wada. 2014. Remote handshaking: Touch enhances video-mediated social telepresence. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 2143–2152. DOI: <https://doi.org/10.1145/2556288.2557169>
- [35] Jiaqi Nie, Michelle Pak, Angie Lorena Marin, and S. Shyam Sundar. 2012. Can you hold my hand? Physical warmth in human-robot interaction. In *Proceedings of the 7th ACM/IEEE International Conference on Human-Robot Interaction*. ACM, 201–202. DOI: <https://doi.org/10.1145/2157689.2157755>
- [36] Lauri Nummenmaa, Enrico Glerean, Riitta Hari, and Jari K. Hietanen. 2014. Bodily maps of emotions. *Proceedings of the National Academy of Sciences* 111, 2 (2014), 646–651. DOI: <https://doi.org/10.1073/pnas.1321664111>
- [37] Eunil Park and Jaeryoung Lee. 2014. I am a warm robot: The effects of temperature in physical human-robot interaction. *Robotica* 32, 1 (2014), 133–142. DOI: <https://doi.org/10.1017/S026357471300074X>
- [38] Denis Peña and Fumihide Tanaka. 2018. Touch to feel me: Designing a robot for thermo-emotional communication. In *Proceedings of the Companion of the 2018 ACM/IEEE International Conference on Human-Robot Interaction*. ACM, 207–208. DOI: <https://doi.org/10.1145/3173386.3177016>
- [39] Denis Peña and Fumihide Tanaka. 2018. Validation of the design of a robot to study the thermo-emotional expression. In *Social Robotics. ICSR 2018. Lecture Notes in Computer Science*, vol. 11357, S. Ge et al. (Ed.). Springer, Cham, 75–85. DOI: https://doi.org/10.1007/978-3-030-05204-1_8
- [40] Katri Salminen, Veikko Surakka, Jukka Raisamo, Jani Lylykangas, Roope Raisamo, Kalle Mäkelä, and Teemu Ahmaniemi. 2013. Cold or hot? How thermal stimuli are related to human emotional system? In *Haptic and Audio Interaction Design. HAID 2013. Lecture Notes in Computer Science*, vol. 7989, I. Oakley and S. Brewster (Eds.). Springer, Berlin, 20–29. DOI: https://doi.org/10.1007/978-3-642-41068-0_3
- [41] Sichao Song and Seiji Yamada. 2017. Expressing emotions through color, sound, and vibration with an appearance-constrained social robot. In *Proceedings of the 2017 ACM/IEEE International Conference on Human-Robot Interaction*. ACM, 2–11. DOI: <https://doi.org/10.1145/2909824.3020239>
- [42] Marjorie Fink Vargas. 1986. *Louder Than Words: An Introduction to Nonverbal Communication*. Iowa State University Press.
- [43] Christian J. A. M. Willemse, Dirk K. J. Heylen, and Jan B. F. van Erp. 2018. Communication via warm haptic interfaces does not increase social warmth. *Journal on Multimodal User Interfaces* 12 (2018), 329–344. DOI: <https://doi.org/10.1007/s12193-018-0276-0>
- [44] Graham Wilson, Dobromir Dobrev, and Stephen A. Brewster. 2016. Hot under the collar: Mapping thermal feedback to dimensional models of emotion. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. ACM, 4838–4849. DOI: <https://doi.org/10.1145/2858036.2858205>
- [45] Steve Yohanan and Karon E. MacLean. 2011. Design and assessment of the haptic creature's affect display. In *Proceedings of the 6th International Conference on Human-Robot Interaction*. ACM, 473–480. DOI: <https://doi.org/10.1145/1957656.1957820>

Received June 2019; revised December 2019; accepted March 2020