It's Not the Way You Look, It's How You Move: Validating a General Scheme for Robot Affective Behaviour

Jekaterina Novikova, Gang Ren, Leon Watts

HAL Id: hal-01609384
https://hal.inria.fr/hal-01609384

Submitted on 3 Oct 2017

It's Not the Way You Look, It's How You Move: Validating a General Scheme for Robot Affective Behaviour

Jekaterina Novikova, Gang Ren, and Leon Watts

Department of Computer Science, University of Bath, UK
{j.novikova,g.ren,l.watts}@bath.ac.uk

Abstract. In the emerging world of human-robot interaction, people and robots will work together to achieve joint objectives. This paper discusses the design and validation of a general scheme for creating emotionally expressive behaviours for robots, in order that people might better interpret how a robot collaborator is succeeding or failing in its work. It exemplifies a unified approach to creating robot behaviours for two very different robot forms, based on combinations of four groups of design parameters (approach/avoidance, energy, intensity and frequency). 59 people rated video clips of robots performing expressive behaviours both for emotional expressivity on Valence-Arousal-Dominance dimensions, and their judgement of the successfulness of the robots' work. Results are discussed in terms of the utility of expressive behaviour for facilitating human understanding of robot intentions and the design of cues for basic emotional states.

Keywords: Human-robot interaction; social robotics; nonverbal communication; artificial emotions; body language.

1 Introduction

People tend to treat interactive systems as if they are social agents [20]. McCarthy famously argued that people routinely attribute mental states even to simple systems, such as a thermostat, to make sense of their operational dynamics when their state is otherwise uninspectable [15].When treated as social agents, interactive systems are additionally attributed with social qualities, such as helpfulness or obstinacy, which can influence a person's readiness or ability to make use of them. These qualities could serve to facilitate social coordination in Human-Robot Interaction (HRI) by reflecting aspects of an agent's ability to take action. From a design perspective, this depends upon the creation of cues that can effectively encode relevant social qualities. Researchers in Computer-Supported Cooperative Work have a long-standing interest in the design of systems that provide cues about the status of ongoing collaborator activity. These fall under the general theme of support for awareness [24] and range from the mechanistic articulation of work by coordinating action [18], through to general activity-based indications of social action with a powerful influence on the attitudes and feelings of collaborators [1].

Although awareness is a multifaceted problem, Schmidt argues that its support always depends on combining a selective aspect of the world of work with abstractions

that span the material and computational world [24]. In this paper, we consider the problem of designing a set of emotional cues for robots that could help a person to maintain awareness of relevant states of a robot collaborator. We describe an approach to the systematic design of such cues that depends on qualities of emotional expression. By drawing on affect research in the animal world, we translate a set of design rules for generating expressive behaviours for robots into the design of five basic emotions for two robots of very different construction and behavioural capability. We argue that this general approach could be mapped to any physical robot form, thus advancing the design potential for the use of affect in HRI.

Bodily expression of emotion is an important part of human socio-emotional communication in both humans and animals [2]. Prior studies have shown that affective cues can be interpreted successfully by people when expressed by robots [12,22]. Research on artificial expressions of emotion has made use of a variety of approaches, typically tailored towards particular conceptions of the form of robot on which they will be displayed. In some studies, animators are employed to design emotional movements for a robot [21]. In other studies, body movements of humanoid robots are copied from a human actors' body language [4]. However, humanoid robots represent a highly restricted physical form. Some researchers define more general high-level design patterns for creating an emotionally expressive behavior in robots [12,22,26] but have struggled to establish their general utility [17].

There is a considerable gap between high-level design guidelines for bodily expression of emotion and the implementation of a specific robot with expressive movements. We extrapolate from the design scheme of Novikova and Watts to treat non-humanoid robots in terms of their expressivity [17]. In addition to our goal of establishing the validity of a general framework for designing intelligible emotional cues for social robots, we shall describe in detail how a particular scheme [17] can be implemented into different types of non-humanoid robots. We validate the design scheme with a user study, based on Valence-Arousal-Dominance (VAD) ratings of behavioural expressions as judged by human observers. We thus present data on the consistency of interpretation of expressive behaviours enacted by two non-humanoid robots with very different degrees of expressivity. The contributions of the paper reflect both design and validation considerations:

1. Refinement and generalisation of a design scheme proposed by [17] by presenting a new way of classifying robots based on expressivity, illustrated with five basic emotions as a sequence of VAD parameters.
2. Validation by exposing similarities and differences in the perception of VAD after applying the design scheme to non-humanoid robots of different expressivity.

2   Related Work

2.1   Emotional Body Language (EBL) in Robots

Nonverbal communication through body movements plays an important role in human communication. Expressing emotions is one of the main functions of bodily communication [2]. But people and animals don't only *express* emotional feelings, they also

communicate certain information through their emotional postures and gestures. Thus expressive behaviors can serve as a rich source of information in inter human communication. Heider and Simmel [11] demonstrated in 1944 already that people are biased to interpret moving figures and motion patterns in social or emotional terms. Their experiment showed that it's possible to communicate emotional meaning to people through very basic forms and thus created the base for future work on emotionally expressive robots.

For designing expressive and communicative robot movements it is important to know which features cause the interpretation of intentions and emotions [8]. Up to date, the researchers mostly focused on identification of features related to animacy [25]. However, there exist a small number of studies investigating the relation between robot movements and perceived emotion. The biggest part of these studies use humanoid robots as examples and almost directly transfer human emotive gestures to humanoid robot bodies [4,28].

Karg et al. [12] in their study analyzed if a hexapod robot can express emotion in the way it walks and if these expressions are recognizable. The authors mapped human emotive gait parameters to a hexapod by changing a step length, height and time for one step depending on the emotion. The results of the study revealed that different levels of pleasure, arousal and dominance were recognizable in the way the hexapod walked. Furthermore, higher velocity of a gait resulted in a higher level of preceived arousal, while lower velocity resulted in lower pleasure and lower dominance. Saerbeck and Bartneck [22] also analyzed the relationship between motion characteristics of a robot and perceived affect. They systematically varied two motion characteristics, acceleration and curvature, and found a strong relation between these motion parameters and attribution of affect. Specifically, they found that the level of acceleration is correlated with perceived arousal. They didn't find a direct relationship between acceleration or curvature and perceived valence. Two different robotic embodiments - the iCat robot shaped as a cat with an animated mechanical face and the Roomba robot of a circular shape - were used in this experiment. The authors didn't find significant differences between the embodiments, thus suggested that motion design tools can be used across embodiments.

In a recent study, Singh and Young [26] investigated how a dog-inspired tail interface can be applied to utility robots and communicate high-level robotic states through affect. The study indicated that people were able to interpret a range of affective states from various tail configurations and gestures. As a result, the authors presented a set of guidelines for mapping tail parameters to intended perceived robotic state, e.g. a higher speed projects a higher valence and arousal while a lower speed projects a lower valence and lower arousal, a large horizontal wag results in a higher valence.

There also exist several recent studies on the use of Laban Movement analysis (LMA) for design of emotionally expressive robots with non-humanoid shape. For example, [3] discusses design parameters for designing the movement of a robot with a circular shape using Laban movement analysis which helped to improve the recognition of emotions in the context of a game. Another study [23] develops a computational model for recognizing and generating affective hand movements for display on anthropomorphic and non-anthropomorphic structures. These studies provide an evidence that

even very low actuated robots or non-anthropomorphic structures can make expressive movements based on LMA.

Based on this research, in our recent study recent study [17] we presented an integrated account of the effect of a range of characteristics of robot movement on human perception of affect. We used anatomical body planes as a reference for combining research on animal social behaviour with Shape and Effort dimensions derived from the Laban theory of movements so present a scheme for designing emotionally expressive robotic behaviours. The scheme includes two concepts to define emotionally expressive behaviours for robots: Expressive Shape and Expressive Quality. Expressive Shape defines how the overall posture of a robot should change in terms of its physical form, and relates this change to the emotional significance of approach and avoidance in the animal world. It is associated with ten distinct parameters of body motion (see Table 1). Expressive Quality defines the performative characteristics of robot movement, i.e. strength or frequency, again grounding the meaning of these characteristics in prior work on signals of affective state in animals. It is associated with a further thirteen parameters of motion. The general grounding of the scheme is intended to reflect its generality in application for different types of non-humanoid robots. However, it has not been validated in a design context and does not explain how this design scheme might be implemented with different forms of non-humanoid robots. Thus in this current work we introduce a new concept of a robot *expressivity* that allows us to further generalize earlier proposed scheme for designing expressive behaviour and validate it with two very different types of robots.

It is common for non-humanoid robots to vary greatly in terms of the number of embodied degrees of freedom, and the maximum amplitude, velocity and frequency of motions they are able to perform. However, there are some similarities in the influence of the parameter on perceived dimensions of emotional meaning, e.g. higher speed of expressive movement often increases perceived level of arousal, or that reduction of size (shrinking) can reduce the perceived level of dominance. Thus, it may be that all robots are capable of expressing basic emotional states, regardless of their form factor, as long as their behavioural capabilities are mobilised appropriately. So, from a design perspective, we propose that all robots can be described in terms of their general *expressivity* whilst still being able to convey emotional meaning through their movement. As a property, we argue that expressivity refers to aspects of the construction of a robot that constrain the robot's ability to vary in terms of Expressive Shape and Expressive Quality. This leads us to our first hypothesis:

H1. Perceptions of emotionally expressive movements do not vary as a function of the degree of a non-humanoid robot's expressivity.

We provide a detailed description of expressivity as it applies to this study in the method section, so that its treatment as an independent variable is clear.

2.2 Value of Emotions

Organizational behavior researchers have investigated the social influence of human emotions in a workplace [10]. The way individuals conduct their individual work is often conditioned by inferences about the emotional state of other people. Appraisal theories state that emotions can tell a story about the agent that expresses them. Thus,

Fig. 1. The more expressive Lego robot E4 (left) and less expressive Sphero robot (right).

people may draw inferences about the beliefs and intentions of other expressive agents, given perceptions of their emotional state [9]. Prior research in HRI suggests that the same can be true in human-robot teams. Even the non-emotional non-verbal signals of a robot can improve the performance of tasks and decrease perceived human's workload [14]. Beck et al [4] suggest that emotionally expressive tutoring robots would help people to learn better and faster. The findings of [13] indicate that a social robot with emotional behavior can serve as a better assistant in a gaming situation than functionally equivalent robots that do not have emotional behaviours.

From a collaboration perspective, the value of such inferences is to facilitate social coordination. They define expectations for the action that other agents may take, given a basic common sense frame of reasoning [15], but also with an understanding of the ongoing status of their activity. With social agency, it is highly likely that an observer's perceptions of robot affective state will also be conditioned by their general understanding of the task context. More specifically, emotional expressions reflect beliefs about the progress of the work an agent is carrying out. Task progress is a joint function of changes that result from an individual's actions and changes to the environment in which they operate. Consequently, alongside the contribution of emotional expression, it is important to consider how positive and negative task-related events might influence an observer's situational awareness. We shall also treat task context as an important control variable for an observer's expectations about a robot's next actions, in particular whether or not the robot will continue with its current course of action.

All these findings lead us to the third hypothesis:

H2: An observer's beliefs about the successfulness of a robot's actions varies consistently with the nature of the robot's expressive behaviour.

In this paper, we treat beliefs about successfulness through the two complementary observer ratings: judgement of whether the robot successfully completed its task, and judgement of the robot's intention to continue or abandon its current activity.

3   Method

We designed a mixed-model experiment, in which participants observed and rated video clips of a robot in action. We used a between-subject design for presenting clips of two different robots - a more expressive non-humanoid robot E4 with several limbs for the first group of participants and a less expressive abstract robotic ball Sphero for

SHAPE

Group	ID	Parameter Name	Body Part or Ability
Approach	1	Transfer weight forward	ability to bend or bow forward
Approach	2	Move limbs forward	movable limbs
Approach	3	Move its body forward	wheels, tracks, legs. Roll, fly, swim, drive, go, move the body forward
Approach	4	Move visible appendage(-s) away from the body	movable limbs, visible movable appendage(-s) not used for moving forward/backward, movable head
Approach	5	Extend or expand its body	ability to extend or expand itself
Avoidance	6	Transfer weight backward	ability to bend or bow backward
Avoidance	7	Move limbs backward	movable limbs
Avoidance	8	Move its body backward	wheels, tracks, legs. Roll backward, fly backward, swim backward, drive backward, go backward, move the body backward
Avoidance	9	Attract limbs close to the body	movable limbs, visible movable appendage(-s) not used for moving forward/backward, movable head
Avoidance	10	Reduce its body	ability to reduce itself

QUALITY

Group	ID	Parameter Name	Ability to Program
Energy	11	High strength	motor's speed at high level
Energy	12	Low strength	motor's speed at low level
Intensity	13	Sudden	sudden start/finish
Intensity	14	Not sudden	smooth start/finish
Flow	15	Short duration	movement able to finish in a short time
Flow	16	Medium duration	movement able to finish in a medium time
Flow	17	Long duration	movement able to finish in a long time
Flow	18	High change in tempo	motor's speed change
Flow	19	High frequency	high level of vibration, spinning or frequent movements of the limbs
Flow	20	Medium frequency	medium level of vibration, spinning or frequent movements of the limbs
Flow	21	Low frequency	low level of vibration, spinning or frequent movements of the limbs
Flow	22	Direct trajectory	straight, linear, direct movement of the whole body
Flow	23	Indirect trajectory	curved movement of the whole body

Table 1. Parameters of a *Shape* (left) and *Quality* (right) group with associated robot's programming abilities.
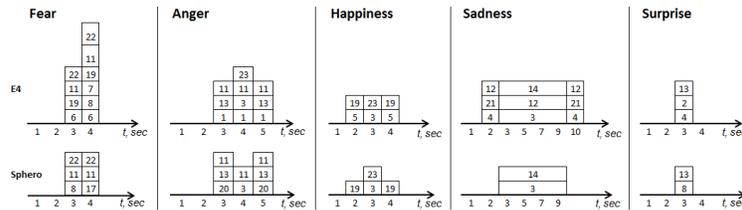
Fig. 2. The combination of design parameters for the emotional expressions of fear, anger, happiness, sadness and surprise, as implemented in a more expressive E4 robot (top) and a less expressive Sphero robot (bottom).

the second group. Within each group, we used a within-subject design for presenting subjects with a sequence of expressive behaviours performed by their respective robot.

3.1 Classifying Robot Expressivity

We used the scheme proposed by [17] for designing emotional body language in our robots. The scheme presents a hierarchical system of design characteristics combined into two large movement groups: *Shape* and *Quality*. The lowest level of the scheme consists of 23 parameters. We linked each parameter of the *Shape* group to the capability of a robot to move its body in a specific way, depending on its construction. We also linked each parameter of the *Quality* group to an ability to program robot actions in a specific way. The list of *Shape* and *Quality* design parameters (DPs) with an associated ability to program robot movements are listed in the right-hand part of Table 1.

The list of expressive parameters, allowed us to define the level of expressivity for any type of robot simply by summing the parameters that can be activated in a specific robot. Thus, the maximum expressivity level for any type of robot is determined by its ability to make use of all 23 parameters. This is a simplistic method for contrasting the base expressivity of any form of robot since it does not privilege any particular parameter. It may be that specific parameters of Expressive Quality or Expressive Shape,

or combinations thereof, invoke higher emotional significance. We shall return to this point in our general discussion.

Each design parameter is associated with one or several emotions, as we'll discuss later in the section on *Emotional Expressions*. Thus, the higher a robot's expressivity, the greater its potential ability to express emotion through body language.

3.2   Emotional Expressions

We created five emotional expressions for the robots, namely: (1) afraid, (2) angry, (3) happy, (4) sad and (5) surprised. The emotions were selected as a subset of commonly known *discrete* or *basic* emotions, as defined by [7]. We used design parameters shown in Table 1 to create emotional expressions in robots, based on the mapping from animal behaviours to general parameters of body movement in [17]. We were able to make use of more design parameters for creating expressions in the high expressivity robot E4 than in the less expressive Sphero because of differences in their construction. One of the contributions of this paper is to demonstrate how a general scheme for designing robot emotional expressions can be mapped to non-humanoid robots with very different expressive possibilities. The precise mappings require the designer to exercise judgement, as it true of all design, but the general scheme does not privilege any particular parameter. Thus, design freedom is preserved for at least basic emotions. Figure 2 presents the combinations of design parameters used for creating each emotional expression in both robots, where block numbers correspond to the ID numbers allocated to design parameters and the horizontal axis represents time of onset and offset in seconds. For example, to create an expression of happiness in the Sphero robot, we used a parameter No. 19 (vibration at a high level) at two seconds, parameters No. 3 and 23 at three seconds (moving forward in a curved trajectory), and parameter No. 19 (fast vibration) at four seconds, creating an expressive behaviour that lasted for three seconds in total. As seen from the Figure 2, both robots use the same initial DPs for expressing each of five basic emotions e.g. parameters No. 8, 11 and 22 for expressing Fear; 3, 11 and 13 for expressing Anger etc. Such a similarity in designing emotional expressions makes the comparison of the movements valid although the capabilities of the actuators are very different in two presented robots.

Some previous studies [4,19] showing an importance of a context in interpreting robot's affective cues encouraged us to include a context in the design of our study. We linked a situational context in which robots were acting to the three emotional dimensions of valence, arousal and dominance in accordance with a VAD space proposed by [16]. For each dimension, we designed a positive and negative context thus getting six contextual environments. For creating the context of a positive valence (V+) something positive happened in the robot's environment, e.g. robot managed to finish its task successfully. For the context of a negative valence (V-), something negative happened in the environment due to robot's fault. Similarly, the context was linked to both positive and negative arousal and dominance, as shown in the Table 2. The neutral context was the same for all the dimensions and ment that nothing happened in the robot's environment.

Type of context (more detailed explanation)	Consistent emotional expression	Inconsistent emotional expression	Neutral emotional expression
A+ (a block suddenly falls down from above in front of a robot)	sad (1)	surprised (2)	none (3)
A- (the task is already finished, robot's help is not needed)	surprised (4)	sad (5)	none (6)
V+ (robot finishes its task successfully)	happy (7)	sad (8)	none (9)
V- (all the blocks fall and scatter due to robot's fault)	sad (10)	happy (11)	none (12)
D+ (a dangerous big obstacle prevents a robot from completing a task)	afraid (13)	angry (14)	none (15)
D- (a harmless small obstacle prevents a robot from completing a task)	angry (16)	afraid (17)	none (18)

Table 2. The combination of a context and a consistent/inconsistent/neutral emotional expression. Here, A+,V+,D+ means a context of a positive arousal, valence, dominance, A-,V-,D- means a context of a negative arousal, valence, dominance. Number in brackets is a number of a test condition.

3.3   Independent Variables

The two main independent variables in our experiment were *expressivity* of robot (high expressivity vs. low expressivity), *Design Parameter group* (approach/avoidance; high/low energy; high/low intensity; high/medium/low frequency). We also varied the influence the occurrence of positive and negative events in the robot's environment to examine the *consistency of emotional ratings* as an indication of the robustness of expressive behaviours (consistent; inconsistent; not emotional).

Robots We used two robots in our experiment: E4 and Sphero (see Figure 1).

Robot with Higher Level of Expressivity. The more expressive robot, E4, was implemented with Lego Mindstorms NXT, was based on a Phobot robot design [6]. The robot had two motors which allowed it (1) to move forwards and backwards on a surface, (2) to move the upper part of its body. The upper body part was constructed such that the robot's hands moved together with its neck and eyebrows. Its neck could move forwards and backwards, and its hands and eyebrows could move up and down. The overall expressivity level of the E4 robot was 19. The RWTH Mindstorms NXT Toolbox for MATLAB [1] was used to program E4's behaviours. This software is a free open source product and is subject to the GPL.

Robot with Lower Level of Expressivity. The less expressive robot, Sphero, is a robotic ball [2] with a ARM Cortex M4 processor, two RGB LEDs and two internal motors that allowed it (1) to roll on a surface at different speeds and directions, (2) to spin or vibrate at different frequencies. Although it is also possible to change Sphero's colour, we did not use this function in our study. The outer shell is made of white polycarbonate. The overall expressivity of the Sphero robot was 12.5. We used the Android SDK provided by Sphero [3] to program Sphero's rolling direction, speed and directional pattern. We used a Samsung TabPRO 8.4 tablet to control Sphero via Bluetooth for creating the video clips.

Design Parameters (DPs)

Four groups of design parameters (DPs) were used as independent variables in our study. For the high-level group of *Shape*, we used *Approach* and *Avoidance* DPs. For

---

[1] `http://www.mindstorms.rwth-aachen.de/`

[2] `http://www.gosphero.com`

[3] `https://github.com/orbotix/Sphero-Android-SDK`

the high-level group of *Quality*, we used low and high *Energy*, low and high *Intensity* and low, medium and high *Frequency*, which is a sub-level of the *Flow* group.

Consistency of Emotional Ratings We recorded a set of videos where an event in the robot's task environment was combined with a specific emotional expression of the same and the opposite level of the appropriate dimension, e.g. an event of a positive valence was recorded with the robot expressing an emotion of a positive valence, of negative valence and a neutral one. If the sign of context's emotional dimension matched the sign of a robot's expressed emotion on the same dimension, we treated the emotion as *consistent*. If a sign of the context was opposite to the sign of a presented robot's emotional expression, we treated it as *inconsistent*. If robot only performed the actions related to its task and didn't perform any emotional expression in addition, we called such an emotion *neutral*.

3.4 Test Conditions

We recorded five emotional expressions performed by each robot in a neutral environmental context. In addition, we recorded eighteen combinations of each context and a consistent, inconsistent and neutral emotion. The combination of a context and a consistent/inconsistent/neutral emotional expression is presented in the Table 2.

Five emotional expressions without context plus eighteen test conditions described in the Table 2 resulted in a list of twenty three emotional expressions of each robot in different contexts, each of the duration of 3-13 sec.

We used video recordings in our study instead of real robot's observations in order to overcome the limitations of live trials. The method of using a real robot has several important limitations for our study: (1) the beginning and end times of an interaction trial are not clearly defined, (2) the context is not clearly defined, and finally, (3) while using a real robot its movements are not exactly the same from trial to trial due to the noise in motors' accuracy. Thus, live HRI trials would make it very difficult to control the conditions and to ensure that statistically valid results are obtained. Videotaped HRI trials, on the other hand, overcome these limitations: the movements of the robot are observed as exactly the same by each participant, there is no ambiguity about the duration of interaction, its beginning and end. There is also no ambiguity about the presented situational context in which the robot operates. Woods et al. verified in their study [27] whether videotaped HRI trials for various scenarios could be used in certain situations instead of live HRI trials and concluded that for certain HRI scenarios including the issues of speed, space and distance videotaped trials are representative and realistic, and do have potential as a technique for prototyping, testing and developing HRI scenarios and methodologies. These are the issues that play a crucial role in the context of robot affective expressions thus the conclusions of the Woods et al. study [27] are applicable to our study and justify the choice of videos over the real robot.

3.5 Dependent Variables

Our dependent variables included emotional ratings of robot expressive behaviours; ratings of robot task intention, and ratings of robot task success. We also collected demographic information on age and gender.

Perceived Emotional Dimensions Participants rated valence, arousal and dominance of robot expressive behaviours with a validated questionnaire called the 'Self assessment manikin' (SAM) [5]. SAM has been used to rate the affective dimensions of valence, arousal and dominance in a wide variety of settings [5].

Judgement of Robot Intentions Judgements of robot intentions were scored on a 5-point Likert scale, where score 1 means *'Definitely not going to continue'* and score 5 means *'Definitely going to continue'*.

Judgement of Robot Task Success Judgement of task success was again scored on a five-point Likert scale, in response to the question *Do you think the robot's task was completed successfully?*. The scale ranged from *Definitely No* to *Definitely Yes*.

3.6   Experimental Procedure and Participants

A between-subject design for the robot expressivity variable. 34 participants (9 females and 25 males; age from 18 to 46, M=23.21, SD=7.42) rated video clips of the high-expressivity E4 robot. 20 participants (7 females and 13 males; age from 23 to 38, M=29.25, SD=3.60) were assigned to the low-expressivity Sphero robot .

A within-subject design was used to assign participants to a specific task condition, i.e. each participant was exposed to all the twenty-three experimental conditions with one of the robots. In order to overcome limitations of a within-subject design and decrease the impact of a learning effect, the videos presented to each participant in pseudorandom order but also ensuring that two expressions of the same type were never presented one after another.

Participants watched the video clips whilst seated in a quiet room, completing ratings after each separate clip. They were recorded the whole way through the experiment and at the end of the experiment participants were invited for a 5-10 minute recorded interview, after which they were debriefed. The duration of the experiment did not exceed thirty-five minutes and though participants were informed that they could leave at any time, none decided to do so.

3.7   Data Analysis

Cronbach's α was used as a measure of internal agreement between subjects. For the videos showing only the context the α value for the ratings was 0.835, and for the videos showing only the emotional expressions the α value was 0.607. The ratings for the videos showing the combinations of the context and emotional expressions, the α value for the ratings was 0.708. All these α values are acceptable, indicating a good level of internal agreement between all subjects across all the scenarios and respective video conditions.

Mixed measures ANOVA was used to examine the relation between each design parameter and the SAM ratings for the two robots. The same test with different factors was used to evaluate the potential influence of context consistency.

4 Results

We conducted several tests of two factor mixed measures ANOVA to analyze an influence of different design parameters on the perception of robot's valence, arousal and dominance. We also analyzed the influence of both between- and within-subject factors on the perceived level of a robot's intention to continue its job.

4.1 Perceived Emotional Dimensions

We only report significant results in this section. The overview of all the ANOVA tests results showing the effect of different DPs on a perceived valence, arousal and dominance are shown in the Table 3.

	SHAPE			QUALITY								
	2 (Robots: Sphero, E4) x 3 (DPs: Approach, Avoidance, Neutral) ANOVA			2 (Robots: Sphero, E4) x 2 (DPs: High Energy, Low Energy) ANOVA			2 (Robots: Sphero, E4) x 2 (DPs: High Intensity, Low Intensity) ANOVA			2 (Robots: Sphero, E4) x 3 (DPs: High, Medium, Low Frequency) ANOVA		
	DPs	DPs xRobots	Robots	DPs	DPs xRobots	Robots	DPs	DPs xRobots	Robots	DPs	DPs xRobots	Robots
Perceived Valence	$F_{1.8,57}$ =15.14, **p<0.001**	$F_{1.8,57}$ =0.54, p=0.566	$F_{1,315}$ =1.69, p=0.195	$F_{1,157}$ =51.02, **p<0.001**	$F_{1,157}$ =2.84, p=0.094	$F_{1,157}$ =0.99, p=0.320	$F_{1,477}$ =1.48, p=0.224	$F_{1,477}$ =2.79, p=0.096	$F_{1,477}$ =1.36, p=0.244	$F_{1.83,2}$ =15.84, **p<0.001**	$F_{1.83,2}$ =8.71, **p<0.001**	$F_{1,158}$ =0.14, p=0.705
Perceived Arousal	$F_{1.9,60}$ =191.33, **p<0.001**	$F_{1.9,60}$ =38.54, **p<0.001**	$F_{1,314}$ =0.04, p=0.835	$F_{1,157}$ =93.57, **p<0.001**	$F_{1,157}$ =0.01, p=0.949	$F_{1,157}$ =3.79, p=0.053	$F_{1,475}$ =261.15, **p<0.001**	$F_{1,475}$ =35.01, **p<0.001**	$F_{1,475}$ =0.44, p=0.507	$F_{2,314}$ =56.73, **p<0.001**	$F_{2,314}$ =4.94, **p=0.008**	$F_{1,157}$ =3.93, **p=0.049**
Perceived Dominance	$F_{2,614}$ =20.14, **p<0.001**	$F_{2,614}$ =5.63, **p<0.005**	$F_{1,307}$ =006, p=0.810	$F_{1,156}$ =31.60, **p<0.001**	$F_{1,156}$ =5.69, **p<0.05**	$F_{1,156}$ =0.05, p=0.831	$F_{1,467}$ =0.02, p=0.877	$F_{1,467}$ =12.97, **p<0.001**	$F_{1,467}$ =0.13, p=0.716	$F_{2,310}$ =1.66, p=0.193	$F_{2,310}$ =6.01, **p<0.005**	$F_{1,155}$ =1.89, p=0.171
Perceived Intention	$F_{2,636}$ =5.11, **p=0.006**	$F_{2,636}$ =2.50, p=0.083	$F_{1,318}$ =0.57, p=0.452	$F_{1,157}$ =9.15, **p<0.005**	$F_{1,157}$ =6.71, **p<0.05**	$F_{1,157}$ =0.08, p=0.781	$F_{1,477}$ =6.35, **p<0.05**	$F_{1,477}$ =0.16, p=0.693	$F_{1,477}$ =4.94, **p<0.05**	$F_{2,314}$ =2.19, p=0.114	$F_{2,314}$ =0.69, p=0.505	$F_{1,157}$ =5.12, **p=0.025**

Table 3. ANOVA results, showing the effect of different design parameters (DPs) on perceived Valence, Arousal and Dominance, using the more expressive E4 and less expressive Sphero robots.
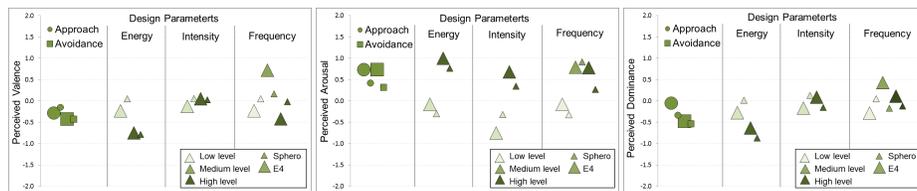
Fig. 3. Plot of the mean values of perceived Valence (left), Arousal (center) and Dominance (right) for the expressions with implemented parameters of approach-avoidance, energy, intensity and frequency, using the more expressive E4 and less expressive Sphero robots.

We found a significant difference in the effect of *Approach* and *Avoidance* design parameters SAM ratings. The first column of the left part of the Figure 3 shows that the mean valence rating for the avoidance behaviours for both robots (mean=-0.43, 95% CI=[-0.54, -0.31]) was lower than approach behaviours (mean=-0.22, CI%=[-0.36,

0.08]). The mean dominance rating for avoidance behaviours (mean=-0.49, 95% CI=[-0.61, -0.37]) was lower than for approach (mean=-0.20, 95% CI=[-0.33, -0.07]), as shown in the first column of the right part of the Figure 3. The effect of interaction between a robot and DP was significant for the perception of arousal and dominance, although the interaction only influenced the observers' ratings when the design factor changed from neutral to not neutral. While changing from approach to avoidance, the interaction effect did not differ significantly.

We found a significant difference in the effect of high and low *Energy* DP on valence, arousal and dominance ratings. The mean valence rating for high-energy behaviours (mean=-0.77, 95% CI=[-0.93, -0.60]) was lower than that of a low energy expression (mean=-0.09, 95% CI=[-0.24, 0.05]). The mean score of arousal for the expression of a low energy (mean=-0.19, 95% CI=[-0.37, -0.02]) was significantly lower than that of a high energy expression (mean=0.88, 95% CI=[0.74, 1.02]). The mean score of dominance for the expression of a low energy (mean=-0.13, 95% CI=[-0.31, 0.05]) was significantly higher than that of a high energy expression (mean=-0.75, 95% CI=[-0.92, -0.58]). The mean scores are presented in the second columns of each plot in the Figure 3. The effect of interaction between a robot and DPs was significant for the perception of dominance: for the more expressive E4 robot the effect of a high-energy DP was stronger than for the less expressive Sphero.

We found a significant difference in the effect of high and low *Intensity* DP on ratings of arousal. The mean arousal rating for the behaviours of low intensity (averaged for both robots; mean=-0.54, 95% CI=[-0.64, -0.43]) was significantly lower (p<0.001) than for those with high intensity (mean=0.51, 95% CI=[0.42, 0.61]). The interaction between Robot and DP was significant for dominance: for E4 robot, the mean rating of valence for low-intensity expressions (mean=-0.13) was lower than that of high-intensity expressions (mean=0.05) although the difference between these two values was not significant. For Sphero, mean valence rating for low-intensity (mean=0.06) was higher than that of high-intensity behaviours (mean=0.03) although this difference was either not significant (see third columns of each plot in Figure 3.

Finally, we found a main effect for the *Frequency* DP on ratings of valence and arousal. Expressive behaviours of medium frequency received the highest valence ratings (mean=0.44, 95% CI=[0.23,0.65]) comparing to those of low (mean=-0.09, 95% CI=[-0.23,0.05]) and high frequency (mean=-0.22, 95% CI=[-0.40, -0.04]). Medium frequency behaviours also received the highest arousal ratings (mean=0.85, 95% CI=[0.69, 1.01]) comparing to those of low- (mean=-0.20, 95% CI=[-0.38, -0.03]) and high-frequency (mean=0.53, 95% CI=[0.39, 0.66]) (see last columns of each plot in Figure 3.

With respect to a *Consistency*, our data suggest that valence, arousal and dominance of a robot's expression are not strongly influenced by positive and negative events in the robot's operational context. However, we found positive context to significantly (p<0.001) increase the mean ratings of both valence (mean=0.58, 95% CI=[0.41, 0.75]) and dominance (mean=0.93, 95% CI=[0.75, 1.10]) when compared to negative contexts. Additionally, the context of a negative arousal significantly (p<0.005) decreased the mean arousal rating (mean=-0.42, 95% CI=[-0.60, -0.25]).

4.2   Value of Emotional Expressions

We treated the value of emotional expressions primarily in terms of their ability to support inferences about a robot's intentions to continue cleaning the room, and the successfulness of its cleaning actions.

Observer Judgement of Robot Intentions   Row four of Table 3 presents ANOVA results for the four types of DP on perceived Intention. We only discuss contrasts that reached statistical significance.

We found a significant difference main effect of *Approach* and *Avoidance* on judgement of robot intention. The mean score of intention for the approach expression (mean = 2.81, 95% CI=[2.67, 2.95]) was significantly higher than either neutral (mean=2.54, 95% CI=[2.40, 2.68]) or avoidance expression (mean=2.59, 95% CI=[2.46, 2.72]). We also found that ratings of intention for differed by *Energy* levels. The mean score of intention for the expression of a low energy (mean=2.81, 95% CI=[2.60, 3.01]) was higher than that of a high energy expression (mean=2.46, 95% CI=[2.28, 2.63]). Although the size of effect is small in both cases, our participants were highly consistent in their ratings on these two measures so confidence in these results is high. The main effect of type of robot did not reach significance for Energy or Approach/Avoidance, but robot type did interact with the Energy DP.

There was a main effect of *Intensity* for judgements of robot intention, with a mean score for low-intensity expressions (mean=2.63, 95% CI=[2.52, 2.75]) significantly lower than that for high-intensity expressions (mean=2.82, 95% CI=[2.71, 2.92]). In this case, scores also varied by type of robot, with both high- and low-intensity behaviours of Sphero rated higher overall than their equivalents for E4.
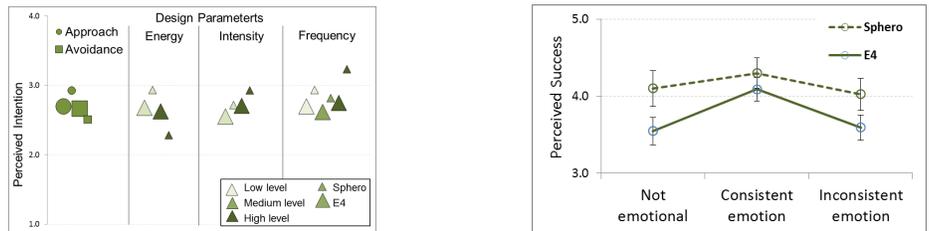
Fig. 4. Left: Plot of the mean values of perceived robot's Intention and standard errors for the expressions of Low, Medium and High Frequency, using the more expressive E4 and less expressive Sphero robots. Right: Plot of the mean values of Success and standard errors for robot expressing emotion consistently, inconsistently and not expressing them, using the E4 and Sphero robots. Based on videos where task was completed successfully.

Observer Judgement of Robot Task Success   Judgement of task success differs from robot intention, as it depends on the interplay between changes in the task environment (its operational context) and the expressive behaviour of the robot. We assume that a

person would jointly assess the robot's behaviour and its operational context to decide whether or not its task was completed successfully. If behavioural and operational context both suggest a positive outcome, they are consistent and thus should present a clear signal of success. Similarly, if both are negative, they should clearly signal failure. This is why we use a *consistency of emotion* factor for analyzing the data using ANOVA test.

A two- (E4 vs. Sphero) x three- (Not emotional, Consistent emotion, Inconsistent emotion) mixed measures ANOVA was used to analyze the influence of expressive behaviour on judgements of task success. In this paper, we limit our analysis to video clips that objectively show that the block-moving task was in fact completed successfully (see Figure 4 Right). The mean rating of success was significantly different for each robot ($F(1.76, 182.98)=3.67$, $p=0.03$, observed power=0.63). Post-hoc tests revealed that observers judge successfulness significantly higher ($p<0.05$) for robots with context-consistent emotional expressions (mean=4.20, 95% CI=[3.95, 4.44]) than for neutral (mean=3.82, 95% CI=[3.54, 4.11]) or context-inconsistent expressions (mean=3.81, 95% CI=[3.54, 4.07]). The difference between two types of robots ($F(1, 104)=4.29$, $p=0.04$) does not interact with this result.

5   Discussion

This paper has reported the implementation of the five basic emotions as robot expressive behaviours in two forms of robot, based on a design scheme for expressing and interpreting emotional body language. The use of two very different robots was intended illustrate the general utility of the design scheme, accompanied by empirical data on human interpretation of the emotional content of these expressive behaviours. Our findings partially support the first hypothesis:

H1. Perceptions of emotionally expressive movements do not vary as a function of the degree of a non-humanoid robot's expressivity.

We found that some design parameters, such as high energy level or avoidance, have a similar influence on observer perceptions of valence, arousal and dominance for both forms of robot i.e. regardless of robot expressivity. These results are consistent with the findings of [26], who showed that (a) high speed of tail movements increased perceived arousal of a robot, and (b) low tail height decreased perceived valence. The latter could be mapped to the *Reduce Yourself* parameter of the *Avoidance* DP group.

Our findings also suggest that some parameters, e.g. approach, high and low intensity or medium and high frequency of movements when implemented into robots of different expressivity level, exert a similar influence on perceptions of a subset of emotional dimensions. For example, high frequency consistently increased ratings of arousal for both types of robots, although its influence on valence differed by robot type. Table 4 presents all the similarities between a more expressive and a less expressive robot revealed by our study. These findings partially supported our first hypothesis.

However, our study also suggests that there are some significant differences in how some parameters influence perceptions of emotion in robot as a function of expressivity, contrary to our expectations:

Group of design parameters	Perceived Valence	Perceived Arousal	Perceived Dominance
Approach	↓ " − "	↑ " + "	
Avoidance	↓ " − "	↑ " + "	↓ " − "
Low intensity		↓ " − "	
High intensity		↑ " + "	
High energy	↓ " − "	↑ " + "	↓ " − "
Medium frequency		↑ " + "	
High frequency		↑ " + "	

Table 4. Similarities in parameters' influence on valence, arousal and dominance between a more expressive robot E4 and a less expressive robot Sphero. Arrows ↑ and ↓ show whether the parameter increased or decreased a perceived value of valence, arousal and dominance. Signs " − " and " + " show whether the value is negative or positive.

– Both types of robots showed that avoidance behaviours were rated as low dominance. However, for the low-expressivity robot, the ratings was significantly lower than for the highly expressive robot.
– Only the high-expressivity robot was rated with a lower level of dominance for low-frequency expressive behaviours than for high-frequency expressions. In addition, the value of dominance ratings in this case was positive for the low-expressivity robot but negative for the high-expressivity robot.
– The high intensity DP increased the level of perceived valence for the highly expressive robot and made it positive, while for the low-expressivity robot the level of perceived valence was decreased and negative.

Table 5 presents all the differences between a more expressive and a less expressive robot revealed by our study. These findings didn't support our first hypothesis. They also add to the current knowledge of the design of emotional expressions in robots, as no previous studies suggested that there could be different consequences of applying expressive movements to different types of robots.

Group of design parameters	Perceived Valence		Perceived Arousal		Perceived Dominance	
	E4	Sphero	E4	Sphero	E4	Sphero
Approach					NA	↓
Avoidance					↓	⇓
Low intensity	↓ "-"	↑ "+"			↓ "-"	↑ "+"
High intensity	↑ "+"	↓ "-"			↑ "+"	↓ "-"
Low frequency	↓ "-"	NA "+"	NA	↓	↓ "-"	NA "+"
Medium frequency	↑	NA			↑ "+"	NA "-"
High frequency	↓	NA	⇑	↑		

Table 5. Differences in parameters' influence on perceived valence, arousal and dominance between a more expressive robot E4 and a less expressive robot Sphero. Arrows ↑ and ↓ show whether the parameter increased or decreased a perceived value of valence, arousal and dominance. Wider arrows ⇓ and ⇑ show a stronger decrease/increase effect. Signs " − " and " + " show whether the value is negative or positive.

In addition to the current knowledge, the Consistency findings of our study revealed that the context of positive valence specifically has a significant effect on perceived

valence and dominance of an expressive robot. With respect to perceived arousal, our findings revealed that the context of a negative arousal decreases it significantly. Other contexts, i.e of a positive or negative dominance, positive arousal or negative valence do not have a significant effect on interpretation of an expressive robot.

In contrast to [19], the results of our study did not provide any evidence that the consistency of context can override the interpretation of emotional expression of a robot. Our findings showed that inappropriate emotional context was not different to the neutral context cases in interpretation of valence, arousal, dominance and robot's intention. However, our findings correspond to the results of [19] in the part stating that alignment of robot's action and affective context enhanced the affective interpretation.

H2. An observer's beliefs about the successfulness of a robot's actions varies consistently with the nature of the robot's expressive behaviour.

The findings of the study revealed that consistent emotional expressiveness increased the rating of a task success and it was significantly different from the cases when a robot completing the task was inconsistently expressive, e.g. expressed sadness after successfully completing the task, or not expressive, e.g. just completed the task and didn't follow it with any emotional expression. Such a result shows that participants' awareness of a situation they observed improved when robot behaved in a consistently emotional way thus supporting our third hypothesis. Our findings conform to those of [13] and [14] by showing an additional value of expressive robot on a neutral one. However, our study also resulted in additional finding that extends the state-of-the-art of HRI and shows that inconsistently expressive robot doesn't create an additional situational understanding in human observers although it doesn't reduce a situational awareness either.

The ratings of robot's Intention varied significantly depending on its expressive movements. This means that emotional expressions of a robot can not only communicate emotional signal but also let people draw additional inferences of that robot. These findings support the second hypothesis and they are consistent with [9] who stated people may presume other things about affective agents based on their expressiveness in addition to how he or she is feeling. However, the study of [9] only made this statement about human agents. Our findings make a first step to generalize this idea to a broader set of agents, including robots.

6   Conclusions

We attempted to address a gap in the literature between high-level design guidelines for robotic emotional expression using a body language and the implementation of expressive movements into specific non-humanoid robots. We have presented a refinement of the general design scheme proposed in [17]. We made this design scheme usable for HRI researchers working with different types of non-humanoid robots in two ways. We presented a new technique for classifying non-humanoid robots based on their expressivity. We also demonstrated representations of five basic emotions of fear, anger, happiness, sadness and surprise as sequence of parameters in accordance with the general design scheme. The results of our validation study show both the similarities and differences in the perception of valence, arousal and dominance after applying the de-

sign scheme to non-humanoid robots of different expressivity. The Energy and Approach/Avoidance group of DPs were robust across the two robot forms. However, our data suggest a need for a more considered mechanism for describing combinations of parameters, especially in terms of the frequency and intensity of expressive behaviours. There is also a need to create a more sophisticated statistical model instead of performing a series of ANOVA calculations, thus reducing the risk of Type I errors.

Although we adopted a very simple model for estimating the general expressivity of any robot, it proved adequate for the questions we posed in this paper. Simple summative models are attractive from a design viewpoint, since they create opportunities for creating equally expressive robots with rather different form factors. They reflect a crude assumption that interpretations depend only on the total number of available cues - a basic bandwidth argument - rather than their choreography. Further work is required to probe the limits of our main finding: interpretations of robot expressive behaviours are consistent, regardless of salient differences in their expressive possibilities. It is hard to imagine non-humanoid form factors of robots that would differ much more than Sphero and E4 but, as we have consistently argued in this paper, it's not the way they look, it's the way they move that counts from the viewpoint of the observer. We have deliberately limited our enquiry to basic emotional states. Were a designer to explore sophisticated robot emotional expressions, such as guilt, regret or schadenfreude, a different picture may emerge. However, there are also ethical considerations which have directed our work away from matters such as these.

References

1. Ackermann, M., and Starr, B. Social activity indicators: Interface components for cscw systems. In *Proc. UIST'95, the 8th ACM Symposium on User Interface Software and Technology, Pittsburgh, USA* (1995).
2. Argyle, M. *Bodily communication*. Routledge, 2013.
3. Barakova, E. I., and Lourens, T. Expressing and interpreting emotional movements in social games with robots. *Personal and ubiquitous computing 14*, 5 (2010), 457–467.
4. Beck, A., Hiolle, A., Mazel, A., and Cañamero, L. Interpretation of emotional body language displayed by robots. In *Proceedings of the 3rd international workshop on Affective interaction in natural environments*, ACM (2010), 37–42.
5. Bradley, M. M., and Lang, P. J. Measuring emotion: the self-assessment manikin and the semantic differential. *Journal of behavior therapy and experimental psychiatry 25*, 1 (1994), 49–59.
6. Cramer, H., Kemper, N., Zwijnenburg, A., and de Rooij, O. Phobot: Hri'08 student design competition winner.
7. Ekman, P. An argument for basic emotions. *Cognition & Emotion 6*, 3-4 (1992), 169–200.
8. Gaur, V., and Scassellati, B. Which motion features induce the perception of animacy? In *Proc. 2006 IEEE International Conference for Development and Learning, Bloomington, Indiana* (2006).
9. Gopinath, R. Employees emotions in workplace. *Research Journal of Business Management 5*, 1 (2011), 1–15.
10. Hareli, S., and Rafaeli, A. Emotion cycles: On the social influence of emotion in organizations. *Research in organizational behavior 28* (2008), 35–59.
11. Heider, F., and Simmel, M. An experimental study of apparent behavior. *The American Journal of Psychology* (1944), 243–259.

12. Karg, M., Schwimmbeck, M., Kuhnlenz, K., and Buss, M. Towards mapping emotive gait patterns from human to robot. In *RO-MAN, 2010 IEEE*, IEEE (2010), 258–263.
13. Leite, I., Pereira, A., Martinho, C., and Paiva, A. Are emotional robots more fun to play with? In *Robot and Human Interactive Communication, 2008. RO-MAN 2008. The 17th IEEE International Symposium on*, IEEE (2008), 77–82.
14. Lohse, M., Rothuis, R., Gallego-Pérez, J., Karreman, D. E., and Evers, V. Robot gestures make difficult tasks easier: the impact of gestures on perceived workload and task performance. In *Proceedings of the 32nd annual ACM conference on Human factors in computing systems*, ACM (2014), 1459–1466.
15. McCarthy, J. Ascribing mental qualities to machines. In *Formalizing Common Sense: Papers*. Ablex, 1990, 93–118.
16. Mehrabian, A. *Basic dimensions for a general psychological theory: Implications for personality, social, environmental, and developmental studies*. Oelgeschlager, Gunn & Hain Cambridge, MA, 1980.
17. Novikova, J., and Watts, L. A design model of emotional body expressions in non-humanoid robots. In *Proceedings of the Second International Conference on Human-Agent Interaction (HAI 2014)* (2014).
18. Pinelle, D., Gutwin, C., and Greenberg, S. Task analysis for groupware usability evaluation: Modeling shared-workspace tasks with the mechanics of collaboration. *ACM Transactions on Computer-Human Interaction* (2003), 281–311.
19. Read, R., and Belpaeme, T. Situational context directs how people affectively interpret robotic non-linguistic utterances. In *Proceedings of the 2014 ACM/IEEE international conference on Human-robot interaction*, ACM (2014), 41–48.
20. Reeves, B., and Nass, C. *How people treat computers, television, and new media like real people and places*. CSLI Publications and Cambridge university press, 1996.
21. Ribeiro, T., and Paiva, A. The illusion of robotic life: principles and practices of animation for robots. In *Proceedings of the seventh annual ACM/IEEE international conference on Human-Robot Interaction*, ACM (2012), 383–390.
22. Saerbeck, M., and Bartneck, C. Perception of affect elicited by robot motion. In *Proceedings of the 5th ACM/IEEE international conference on Human-robot interaction*, IEEE Press (2010), 53–60.
23. Samadani, A.-A., Kubica, E., Gorbet, R., and Kulić, D. Perception and generation of affective hand movements. *International Journal of Social Robotics 5*, 1 (2013), 35–51.
24. Schmidt, K. The problem with 'awareness'. *Computer Supported Cooperative Work. The Journal of Collaborative Computing 11* (2002), 3–4.
25. Scholl, B. J., and Tremoulet, P. D. Perceptual causality and animacy. *Trends in cognitive sciences 4*, 8 (2000), 299–309.
26. Singh, A., and Young, J. E. A dog tail for utility robots: exploring affective properties of tail movement. In *Human-Computer Interaction–INTERACT 2013*. Springer, 2013, 403–419.
27. Woods, S., Walters, M., Koay, K. L., and Dautenhahn, K. Comparing human robot interaction scenarios using live and video based methods: towards a novel methodological approach. In *Advanced Motion Control, 2006. 9th IEEE International Workshop on*, IEEE (2006), 750–755.
28. Xu, J., Broekens, J., Hindriks, K., and Neerincx, M. A. Robot mood is contagious: effects of robot body language in the imitation game. In *Proceedings of the 2014 international conference on Autonomous agents and multi-agent systems*, International Foundation for Autonomous Agents and Multiagent Systems (2014), 973–980.