Programming Robots to Express Emotions: Interaction Paradigms, Communication Modalities, and Context

Vimitha Manohar and Jacob W. Crandall

***Abstract*—Robots are beginning to be used in many fields, including health care, assistive industries, and entertainment. However, we believe that the usefulness of robots will remain limited until end-users without technology expertise can easily program them. For example, the wide range of situations in which robots must express emotions as well as the differences in people with whom robots interact require that emotional expressions be highly customized. Thus, end users should have the ability to create their own robot behaviors to express emotions in the specific situations and environments in which their robots operate. In this paper, we study the ability of novice users to program robots to express emotions using off-the-shelf programming interfaces and methods for Nao and Pleo robots. Via a series of user studies, we show that novice participants created non-verbal expressions with similar characteristics to those identified by experts. However, overall, the emotions expressed through these non-verbal expressions were not easily discerned by others. Verbal expressions were more discernible, though substantial room for improvement was observed. Results also indicate, but do not definitively show, that procedural mechanisms can improve users' abilities to create good verbal expressions.**

***Index Terms*—Robot programming systems, human-robot interaction, emotions**

I. Introduction

Robotic systems are becoming increasingly prevalent in entertainment (e.g., as tour guides [1, 2] and actors in theatrical plays [3]), health care and assistive robotics (e.g., [4, 5]), search and rescue (e.g., [6, 7]), education, military operations, and other domains. In these systems, a robot interacts with end users and other people in the environment. As in human-to-human communications [8], emotions play an important role in these human-robot interactions. Expressions of emotion by a robot can (1) help the robot communicate its internal state to end users, (2) encourage desired behaviors from those with whom the robot interacts, and (3) help people to connect emotionally with the robot.

Robot emotions are only useful if they are interpreted correctly by the people with whom the robot interacts. Given the wide range of domains in which robots operate as well as the variations in people with whom robots interact, there is no one-size-fits-all set of emotional expressions. For example, an appropriate behavior to express *surprise* may be very different in a theatrical play than in a search and rescue operation. Likewise, a child with autism is likely to interpret emotional expressions differently than typical children their age. Additionally, end users may want to give their robot a unique personality. Thus, end users should be able to customize robot behaviors to express emotions as they desire.

V. Manohar and J. W. Crandall are in the Department of Electrical Engineering and Computer Science at the Masdar Institute of Science and Technology, Abu Dhabi, MA (email: {vmanohar,jcrandall}@masdar.ac.ae).

Given the need to create customized robot behaviors to express emotions, designers of robotic systems should provide end users with the ability to program their own robot behaviors. Since it is not desirable to require end users to have substantial technology expertise, various programming interfaces and methods have been developed for such purposes. Such programming interfaces interact with the robot's hardware and software to allow users to create customized behaviors.

In this paper, we study the ability of people to create robot behaviors that express recognizable emotions using existing programming interfaces and methods for Nao and Pleo robots. In this context, we analyze the impact of interaction paradigms and communication modalities on the ability of users to create recognizable emotions. We also seek to find ways to improve these programming interfaces and methods so that end users can better create behaviors that express recognizable emotions. More specifically, we focus on four questions:

- How well are novice users able to express emotions through verbal and non-verbal robot behaviors?
- What are the common characteristics of the robot behaviors created by novice users, and how do these characteristics compare to those created by experts?
- How can existing programming interfaces and methods be altered to increase users' abilities to create robot behaviors that express recognizable emotions? In this paper, we study how procedural aspects of behavior creation can be altered to improve the *context* in which users create verbal and non-verbal expressions for robots.
- What is the impact of using *direct interaction* (DI) as opposed to *kinesthetic teaching* (KT) to create robot behaviors that express emotions? While one would likely hypothesize that KT offers a more natural interaction than DI, KT requires the robotic system to have additional hardware and software capabilities, which come at a significant cost. Thus, it is useful to quantify the strength of the benefit of using KT rather than DI.

To begin to answer these questions, we describe and discuss a series of user studies wherein participants programmed Nao and Pleo robots to express emotions. Since Nao and Pleo both have very limited facial expressions, we address these question for robotic systems with such limitations.

II. Background

Before describing the user studies, we first provide relevant background information. After discussing previous work in designing robots that express emotions, we overview relevant programming interface methods that have been designed to help end users create robot behaviors.

A. Robots that Express Emotions

A substantial amount of effort has been dedicated to developing robots that effectively express emotions. We briefly discuss six such robots: Kismet [9, 10], Keepon [11, 12], Pleo, KASPAR [13], FACE [14], and KOBIAN [15, 16].

Kismet, a robot with a 15-degrees-of-freedom (DOF) face, was designed to encourage natural infant-caretaker interactions [9, 10]. Kismet's designers programmed it to express anger, disgust, fear, joy, sorrow, and surprise via facial expressions. It was also designed to vocalize emotions via speakers accompanied with mouth movements.

Another robot designed to express emotions was Keepon [11, 12], a simple snowman-shaped robot. Keepon was not designed to have adaptive facial expressions. Rather, it displayed emotions by the way in which it bobbed its body up and down. Interestingly, these simple expressions of emotion were shown to facilitate the exchange and sharing of mental state with children with autism [12].

Pleo, a robotic dinosaur sold by Innvo Labs Corporation, exhibits an impressive array of emotions via body language and sounds. This robot has also been shown to have some appeal to children with autism. For example, in one experiment, autistic children used encouraging tones to encourage Pleo to cross a body of water that it seemingly feared [17].

Robots with more human-like form have also been designed to express emotions. KASPAR [13] was a child-sized humanoid robot that communicated emotions primarily via simple predefined facial expressions. It also had the ability to communicate by gesturing with its head and arms. Another android robot, named FACE, was designed to mimic a bystander's body movements and facial expressions [14]. One study demonstrated that FACE could potentially be used to help autistic children to learn via imitation.

KOBIAN, a humanoid robot, was designed to express emotions using a combination of facial expressions and body movements [15, 16]. Kobian was carefully crafted to express seven emotions (anger, happiness, surprise, disgust, sadness, fear, and perplexity). Each emotional expression was patterned after expressions observed in photographs and by poses performed by actors. In an initial study, people observed only KOBIAN's facial expressions [15]. Subsequently, whole body expressions were added. A second set of experiments showed that the addition of whole body expressions increased the recognizability of emotions by 33.5% on average. Notable improvements in the recognition of anger (61.7%) and surprise (68.5%) were observed [16].

While these and other robots have been designed to express emotions through robot behaviors, these robot behaviors were carefully crafted by experts. We envision robotic systems that allow end users that do not have substantial technological expertise to quickly and easily program their own customized robot behaviors. Thus, in this paper, we study the ability of novice users to quickly program robot behaviors that express recognizable emotions using existing programming methods. In so doing, we hope to identify how these programming methods can be improved to help users more easily create recognizable robot behaviors.

B. Programming Interfaces

Some of the most common programming methods for robots are visual programming, direct interaction, kinesthetic teaching, and human motion capture. Each of these methods is designed to help users to communicate behaviors to robots more easily. We briefly discuss each method.

1) Visual programming: Visual programming methods allow the user to express a behavior using a logical flow of movements constructed in a visual environment [18]. In such methods, the user is given many predefined functions (e.g., specific robot movements), each of which is represented by a graphical icon. Typically, the user drags and drops these icons into a display, and then connects these icons with arrows to specify a sequence of movements. Visual programming has been shown to be a simple and effective way to choreograph robot behaviors [19].

2) Direct interaction: Direct interaction (DI) refers to the direct manipulation of a graphical element in the interface to specify robot movements rather than using a secondary element in the interface (such as a slider) [20]. In this context, DI refers to the manipulation of a 3D image of the robot in a graphical display to specify and record robot movements. In these manipulations, the user drags the joints of the robot in the image to specify the robot's movements in each time frame. This sequence of robot poses forms a behavior. The rate at which the robot transitions from one pose to the next can be adjusted to control the speed of the robot's movements.

3) Kinesthetic teaching: An alternative method to DI for designing new robot behaviors is kinesthetic teaching (KT). KT is similar to DI, except that, in KT, the user physically moves the robot while the interface records the position of the robot's joints in each time frame rather than dragging a 3D image on the screen [21]. The advantage to this technique is that the programmer is better able to control the robot's movements. However, this method can only be used when the user is co-located with the robot, and when the robot possesses sensors and motors that permit its use. KT can be combined with other methods to help users to adjust the speed of the robot's movements.

4) Human motion capture: Another common method for communicating a behavior to a robot is human motion capture (HMC). In HMC, the user acts out the desired behavior while a tracking system records his or her movements. Tracking can be accomplished via tracking suits [22], computer vision [23], depth sensors (such as Kinect [24]), etc. While HMC is a compelling interface method, such methods must overcome tracking errors and retargeting issues to be successful [25]. We do not evaluate this programming method in this paper.

(a) Pleo (b) Nao

Fig. 1: Robots studied in this research.

C. Behavior Generation

DI, KT, and HMC are input methods commonly used in association with *learning from demonstration* [26] (also called, among other names, *programming by demonstration* and *teaching by demonstrations*). In learning from demonstration, a robot behavior is generated in two steps. First, the user provides one or more demonstrations of the desired robot behavior through some input method. Second, the robot computes a behavior from these demonstrations. At the simplest level, the robot copies the behavior demonstrated by the user. Alternatively, a vast array of more sophisticated behavior-generation algorithms have recently been developed (e.g., [27, 28, 29]). Such machine learning algorithms produce a generalized behavior from many demonstrations by distinguishing between effective and ineffective demonstrations.

In our studies, the robot simply copies the user's demonstrated behavior. We made this design choice for several reasons. First, and foremost, Akguns *et al.* observed that users often provide only a single demonstration of the behavior, and rarely more than a few demonstrations [21]. We have observed similar trends in our studies. Given such limited data, machine learning methods are unlikely to have substantial positive impact on behavior generation. Second, in the context of creating behaviors that express customized emotions, distinguishing between effective and ineffective demonstrations is extremely difficult. Third, we believe that studying how people best demonstrate behaviors is interesting on its own without complicating matters with a learning algorithm.

III. Robotic Systems

In this research, we study how participants create robot behaviors to express emotions using three separate robotic systems. These robotic systems use two different robots and two separate programming interfaces. We describe these robots and programming interfaces in this section.

A. Robots

The robots we considered in this research are Nao and Pleo.

1) Pleo: Pleo, depicted in Fig. 1a, is a dinosaur robot sold by Innvo Labs Corporation. It is modeled after a one week old Camarasaurus. Pleo hardware consists of two ARM7 CPUs, two microphones, eight touch sensors, one shake sensor, one IR transceiver and receiver, and 14 motors. It has 15 joints. One motor controls the eyes and the mouth such that the eyes cannot be closed if the mouth is open. Due to its expressiveness, Pleo has been used in several studies where emotions have been deemed necessary [30]. Pleo does not have a high range of facial features; its emotions are primarily expressed through sound and body language.

2) Nao: Nao, depicted in Fig. 1b, is a 60 cm tall humanoid robot developed by Aldebaran Robotics. Nao has tactical sensors in its head, two speakers, two cameras, two lateral microphones, prehensile hands, and 25 DOF (we used the Academic Edition). Due to instability issues, we fixed Nao to a kneeling position in our studies, thus limiting Nao's movements to the upper body. Like Pleo, Nao can potentially express emotions through both verbal and non-verbal means, but has limited facial expressions.

Despite its limited facial expressions, Nao is being used by many researchers (e.g., [31, 32, 33]). Given this popularity, we believe that it is important to study how well people can create behaviors that express emotions for such robots.

B. Programming Interfaces

Standard programming interfaces are available for both Pleo and Nao robots (Fig. 2). For example, MySkit can be used to program custom behaviors for Pleo. Nao comes standard with Choregraphe. We describe relevant details of each of these programming interfaces.

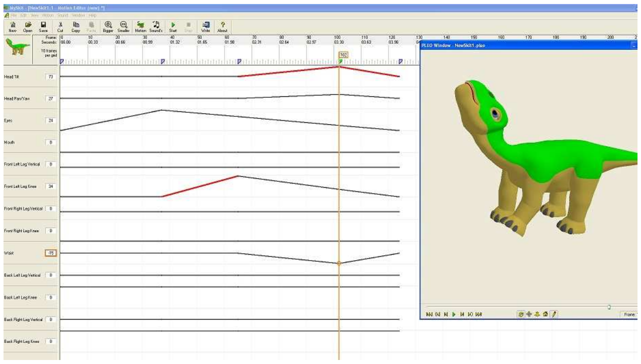
1) MySkit: MySkit (Fig. 2a) has three separate displays: a timeline display, a DI display, and a sound editor display. The timeline display is used by the programmer to specify the position of each of Pleo's joints in each frame. Our observation is that users primarily use the DI display to program Pleo's movements. However, after specifying Pleo's movements with the DI display, users tend to use the timeline display to adjust the speed of movements and to fine-tune the behavior, particularly when defining movements for the eyelids. To do this, the user drags a line on the timeline display to specify the position of a joint and the speed of the joint movements.

The MySkit sound editor display allows users to record and merge audio into the behavior. We observed that this display was convenient for coordinating verbal expressions (recorded audio in our study) with the robot's movements.

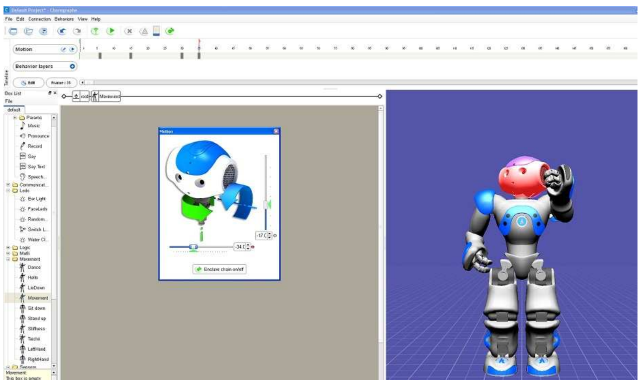
2) Choregraphe: Choregraphe (Fig. 2b) is the programming interface distributed with Nao robots [34]. It uses visual programming to allow users to merge predefined behaviors to create new behaviors. It also allows users to define new behaviors through both DI and KT. The studies reported in this paper focus primarily on these latter functions.

Like MySkit, Choregraphe provides timeline and DI displays through which the user specifies robot movements. We found Choregraphe's DI display to be more cumbersome than MySkit's. Due to Nao's high DOF, Choregraphe requires users to drag sliders to specify joint movements. We note that this is not strictly DI, though the image of the robot on the screen does change as sliders are moved.

Via Choregraph, a user can also specify Nao's movements using KT. In this mode, the user simply moves the robot's



(a) MySkit



(b) Choregraphe

Fig. 2: Programming interfaces studied in this research.

limbs rather than manipulating the 3D image on the screen. The robot's pose is communicated back to the programming interface to record the robot's behavior. The user can record continuous sequences of movements to specify the behavior of each joint over time. As in DI, individual time frames can then be edited to adjust robot poses and the speed of movements.

We did not find built-in sound recording capabilities in Choregraphe. As such, in our studies, participants recorded audio using a separate program. The participants were allowed to record sound clips in this separate program until they were satisfied. They then imported the resulting audio file into Choregraphe. The timeline display in Choregraphe was then used to merge the robot movement (non-verbal expression) and the audio file (verbal expression). Participants defined the point at which the audio file began playing with respect to the robot's movements. They could adjust when the sound began playing or they could adjust the speed of the robot's movements to synchronize movements to sounds.

IV. User Study 1: Experimental Setup

To begin to evaluate how well novice users can program robots to express emotions using off-the-shelf programming interfaces and to begin to better understand what aspects of such systems best allow users to create robot behaviors that express emotions, we conducted two user studies, each carried out in multiple phases. We first describe and discuss the experimental setup and results of the first user study. In Sections VI and VII, we describe and discuss the second user study, which is motivated by results from the first study.

TABLE I: Robotic systems used in the experiment

Name	Robot	Programming Interface	Primary Programming Method
Pleo-DI	Pleo	MySkit	Direct interaction
Nao-DI	Nao	Choregraphe	Direct interaction
Nao-KT	Nao	Choregraphe	Kinesthetic teaching

In the first phase of the first user study (Phase 1.1), participants created robot behaviors targeted to express specific emotions for the three robotics systems overviewed in Table I. In the second and third phases (Phases 1.2 and 1.3, respectively), participants attempted to identify the emotions expressed by the robot behaviors created in Phase 1.1. We describe the setup for each phase of the study in turn.

A. Phase 1.1: Creating Robot Behaviors

The first set of participants were asked to create robot behaviors to express emotions using one of the three robotic systems. Each participant was asked to program behaviors that expressed two emotions out of anger, happiness, sadness, and surprise, which are four of the six emotions identified by Ekman [35, 36]. Each robot behavior consisted of a verbal and non-verbal expression linked together and played in parallel. The non-verbal expression consisted of robot movements, whereas the verbal expression consisted of audio recorded by the participants. In recording verbal expressions, participants were not allowed to use words beyond typical expressions such as "wow," "yippee," and "oh."

The protocol for each participant was as follows:

1) The participant was assigned a robotic system (Pleo-DI, Nao-DI, or Nao-KT).
2) The participant was trained on how to use the designated programming interface to create a robot behavior.
3) The participant was randomly assigned two emotions.
4) The participant created a behavior for the first emotion.
5) The participant created a behavior for the second emotion.
6) The participant answered a post-experiment questionnaire, which was designed to give insight into the participant's experiences in programming the robot.

The length of time it took each participant to program each behavior and the behaviors themselves were recorded and archived for use in subsequent phases of the study.

Twenty-four students with little or no experience programming robots participated in this phase of the study. The average age of the participants was 25.6 years. The subjects were divided equally among robotic systems and emotions; eight participants were assigned to each system. As such, four behaviors for each emotion were created with each system. To minimize ordering effects, the order that behaviors were created to express each emotion was counter-balanced across all participants. For example, behaviors to express anger were created first using each system the same number of times that they were created second.

B. Phases 1.2 and 1.3: Identifying Emotions

Each behavior created in Phase 1.1 was evaluated by two groups of additional participants. In Phase 1.2, a group of participants attempted to identify the emotions expressed by the combined verbal and non-verbal expressions. The other group attempted, in Phase 1.3, to identify the emotions expressed by the verbal and non-verbal expressions separately.

Participants in Phase 1.2 were asked to identify the emotions expressed by the robot behaviors created in Phase 1.1. Each participant was asked to identify eight behaviors from each robotic system. The behaviors were presented to the participants in a random order. Each participant was shown each behavior three times. After each showing, participants noted the emotion that they thought the behavior best resembled out of the four emotions (anger, happiness, sadness and surprise). The participants also recorded the confidence of their assessment on the scale one to four (1 = no idea, 2 = not sure, 3 = sure, 4 = very sure). All results presented in this paper consider only the final assessment made by the participants for each behavior. Twenty people with an average age of 27.4 years participated in Phase 1.2. Each behavior was evaluated by ten participants.

Phase 1.3 was designed to investigate the quality of non-verbal and verbal expressions separately. As in Phase 1.2, participants who took part in this phase were also asked to identify emotions from the behaviors created in Phase 1.1. However, the verbal and non-verbal expressions were presented and evaluated by the participants separately. All other processes of Phase 1.3 were identical to those performed in Phase 1.2, with the exception that only two of the four behaviors (selected randomly) created for each emotion and system were assessed in this phase. Ten participants with an average age of 24.5 years participated in Phase 1.3.

C. Metrics

We used the results from Phases 1.2 and 1.3 to quantify the quality of the emotional expressions present in the robot behaviors created in Phase 1.1. We did this by combining two metrics (*recognizability* and *confidence*) into a single metric called *discernibility*. Let R_i denote the recognizability of behavior i. In our studies, R_i is the ratio between the number of participants that correctly identified the target emotion of robot behavior i and the number of participants that attempted to identify the target emotion in robot behavior i.

While recognizability is an informative metric, it does not account for the confidence that participants had in their assessments. Let $C_{i,j} \in \{1, 2, 3, 4\}$ denote the confidence that participant j had in his/her assessment of behavior i. Our metric of *confidence* (denoted C_i) is given by:

$$C_i = \frac{1}{|S|} \sum_{j \in S} (C_{i,j} - 1)/3, \tag{1}$$

where S is the set of participants that evaluated behavior i, and $|S|$ denotes the cardinality of S. Thus, C_i is a value between 0 and 1, with higher values denoting higher confidence in the assessments of the target emotion.

TABLE II: Attributes identified by Coulson [37].

Emotion	Typical characteristics
Sad	Head lowered towards the chest and arms at the side of the body
Happy	Head backward with raised arms above shoulder level and straight at the elbow
Surprised	Arms raised, with forearms straight, head backward and chest bent
Angry	Arms raised forward and upward, head bent backward

Finally, let D_i denote the *discernibility of behavior* i, given by the product of the recognition rate R_i and confidence C_i:

$$D_i = R_i C_i. \tag{2}$$

Thus, the discernibility of an emotion is in the range 0 to 1, with $D_i = 0$ indicating low discernibility and $D_i = 1$ indicating perfect discernibility.

V. User Study 1: Results

We discuss the results of three aspects of the user study. First, we note common characteristics of behaviors created by our participants, and compare these characteristics to observations made in previous work. Second, we discuss how well the intended emotions were identified by participants. Third, we discuss the impact of the various programming interfaces on the behaviors that were created.

A. Characteristics of Programmed Behaviors

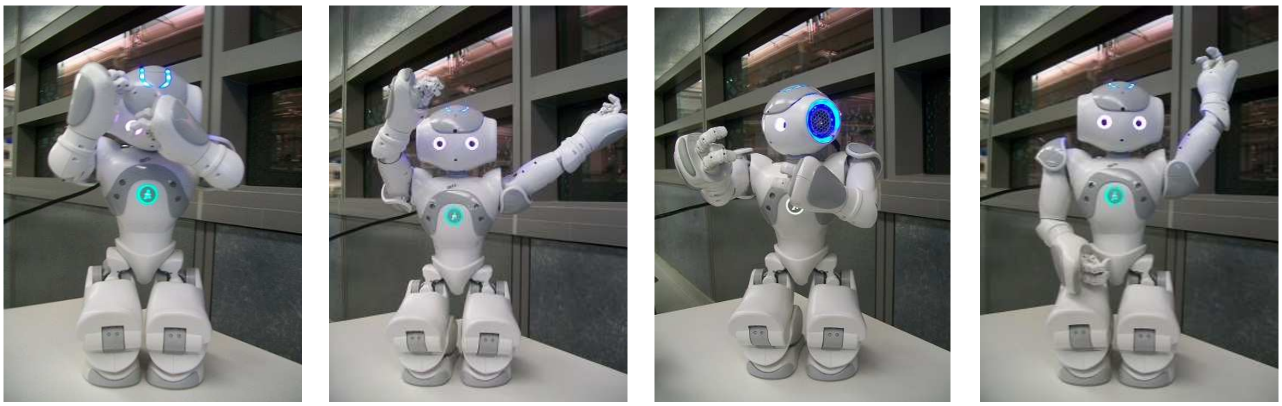
We now summarize common characteristics of robot behaviors created by participants in our studies. In describing Nao's behaviors, we compare our findings with past work by Coulson [37] and Shaarani and Romano [38], who studied the expression of emotion from static and non-static postures of a human body. Table II summarizes observations of non-verbal expressions made by Coulson.

1) Sadness: Robot behaviors created by participants in our study to express sadness were typically characterized by either soft sniffing, sighing, or loud crying. Non-verbal expressions of sadness for Pleo typically consisted of a lowering of the tail and a slow dropping of the head to or near the ground (Fig. 3a). Expressions of sadness created for Nao typically included (1) lowering of the head (Fig. 4a), (2) slow shaking of the head (by four participants), and (3) bringing hands or arms slowly up in front of the face (by six participants) or below the chest (by three participants). Coulson's work typified sadness with lowering of the head with the arms placed at the side of the body. Shaarani and Romano observed that the movement with the highest recognition rate (97%) in their study involved bringing the face down with hands holding the cheeks.

2) Happiness: In our study, robot behaviors to express happiness were typically punctuated by up-beat verbal expressions such as "yeah" and "yippee." For Pleo, three out of four of the participants expressed happiness by programming Pleo to lift one of its front feet (Fig. 3b). All four non-verbal expressions created for Pleo also included quick head movements. Expressions of happiness through Nao often included a raised chin with arms brought into the air (ten of twelve participants),

(a) Sad (b) Happy (c) Surprised (d) Angry

Fig. 3: Sample still images of Pleo in behaviors created by participants for each emotion.



(a) Sad (b) Happy (c) Surprised (d) Angry

Fig. 4: Sample still images of Nao in behaviors created by participants for each emotion.

sometimes with the pumping of an arm (five participants). This behavior is similar to posture characteristics specified by Coulson to express happiness. Shaarani and Romano observed that the expression of happiness that received the highest recognition rate (92%) included the body being held erect and both hands held high and upright, similar to the Nao behavior shown in Fig. 4b, which was created by one of our participants.

3) Surprise: Verbal expressions for surprise varied significantly, including sudden articulations of "wow" (four participants), "huh" (three participants), "yeah" (three participants), and "oh" (two participants). We could not identify any consistent non-verbal attribute for expressions of surprise for Pleo. For Nao, participants typically brought one arm up quickly into the air, or touched the head or chest while raising the chin slightly or looking to the side. The characteristic of a raised arm with the head held backward is similar to posture characteristics specified by Coulson for surprise.

In our study, the behavior expressing surprise that had the highest recognition rate included quick raising of both arms while looking to the side (Fig. 4c). In Shaarani and Romano's work, the two expressions of surprise that received the highest rating (75%) were postures in which the chest was held straight and the body was held erect to the back, with both arms wide open to the side or wide open in front and close to the body.

4) Anger: In our study, most verbal expressions of anger included some form of the expression "aaahhh," "argh," or "hmm." Non-verbal expressions of anger for Pleo were sometimes characterized by stomping or pawing (two participants). However, Nao body movements often included raising one or two arms (similar to the characteristic specified by Coulson) and then dropping the arm or arms quickly. Vigorous shaking of the head and arms was also observed. In Coulson's work, expressions of anger included bending of the head.

The behavior expressing anger in our study with the highest recognition rate for Nao included raising one hand and then dropping it quickly (Fig. 4d). In Shaarani and Romano's paper, the most recognizable posture (92% recognition) to express anger included clinching one hand and then pointing a finger at something. It is difficult for Nao to point since a single motor controls all of its fingers on a hand.

In summary, the non-verbal robot expressions created by our novice participants contained many of the attributes that have been recognized in past work to define emotions.

B. Discernibility of Emotions

Fig. 5 shows the average discernibility of the emotions expressed with each robotics system for each emotion. The figure shows a number of interesting trends. First, the robotic system used to generate and express emotions had no statistically significant effect on the discernibility of the emotions ($F(2, 21) = 0.26$, $p = 0.773$). Emotions from Pleo-DI and Nao-KT were recognized about 64% of the time, while emotions from Nao-DI were recognized 74% of the time.

A second interesting trend illustrated by Fig. 5 is that sadness appears to be the easiest emotion to convey overall. A similar result was observed by Barakova and Louren [39]. In

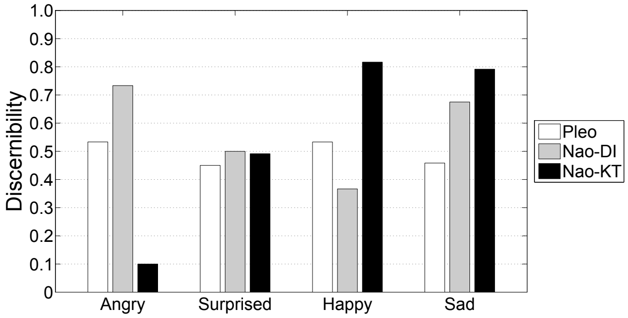


Fig. 5: Discernibility of the emotions created with each system.

our study, sadness was recognized 82% of the time, whereas happiness, surprise, and anger were recognized 69%, 62%, and 57% of the time, respectively. However, the differences in discernibility were not statistical significant ($F(3, 44) = 0.93$, $p = 0.436$). We note that these results are similar to those observed for emotional behaviors created by experts [15, 16].

While behaviors created using Nao-KT were typically quite discernible, the behaviors designed to express anger with this system were not. Participants in Phase 1.2 were only able to identify the four "angry behaviors" created using Nao-KT 18% of the time. On the surface, these results are somewhat puzzling, especially in light of the high discernability of behaviors expressing anger produced with Nao-DI. However, the difference is mostly likely due to chance. A closer inspection of the four behaviors designed to express anger for Nao-KT reveals a potential source of the problem. While the speed of robot movements for expressions of anger were similar for the behaviors created with Nao-KT and Nao-DI, the verbal expressions attached to the Nao-KT behaviors are not easily interpreted as expressions of anger. This result highlights the need to further investigate the relationship between the quality of verbal and non-verbal expressions, and the ease with which emotions are correctly detected. Phase 1.3 of this user study was designed to begin this investigation.

The discernibility of verbal and non-verbal expressions as evaluated in Phase 1.3 are shown in Fig. 6. The figure shows that, when robot movements were presented without audio, discernibility was only about 0.3. The recognition rate of these behaviors was about 40%, which is substantially higher than random guessing (25%). However, the discernibility of the combined verbal and non-verbal expressions was substantially higher (nearly 0.55). Similarly, participants in Phase 1.3 could identify the emotions from just listening to the audio clips (verbal expressions) nearly 67% of the time, with a discernibility of about 0.55 as well. The differences in discernibility depicted in Fig. 6 are statistically significant ($F(2, 21) = 6.66$, $p = 0.002$). Pairwise comparisons show that non-verbal expressions were statistically less discernible than both the combined behavior ($p < 0.001$) and the verbal expressions alone ($p = 0.004$).

In short, though non-verbal expressions created by the novice participants seemed to contain the characteristics of typical emotions, they were not very discernible to others. They were less discernible than emotions recognized by even

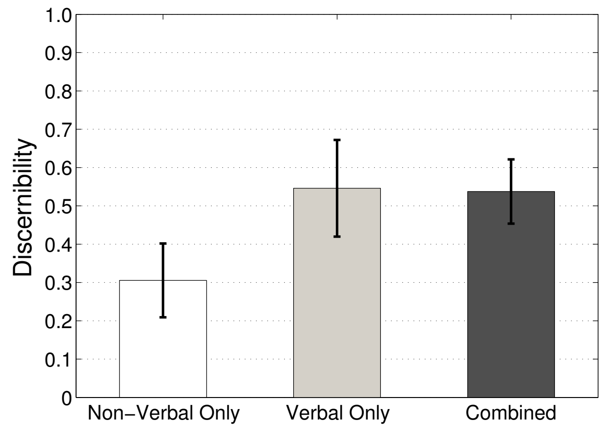


Fig. 6: Discernibility of non-verbal expressions (*Non-verbal Only*), verbal expressions (*Verbal Only*), and the combined behavior (*Combined*). Error bars show a 95% confidence interval on the mean.

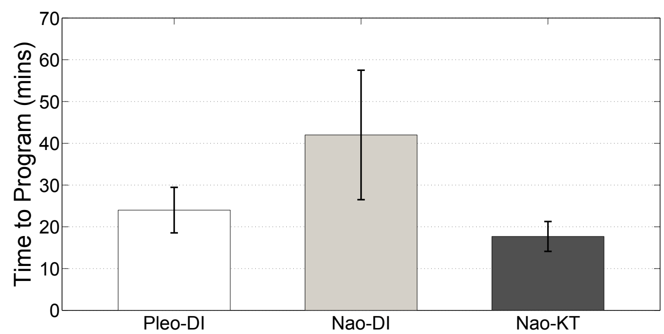


Fig. 7: Time used to program behaviors using each system. Error bars show a 95% confidence interval on the mean.

still images in previous work [38]. Perhaps unsurprisingly, verbal expressions were much more discernible, though even the discernibility of verbal expressions was only about 0.55 on average.

C. Effects of the Programming Interface

While we did not observe any difference in the discernibility of emotions created by the various robotic systems, the robotic system did impact the behaviors participants created, as well as the participants' experiences in programming the robots. As shown in Fig. 7, the robotic system had a statistically significant impact on the amount of time it took users to program robot behaviors ($F(2, 21) = 6.5$, $p = 0.007$). Participants took the longest time creating behaviors using Nao-DI (42.0 minutes), and the shortest time using Nao-KT (17.7 minutes). Creating behaviors for Pleo took 24.0 minutes on average. A pairwise comparison shows a statistical difference between the time users took to program robot behaviors using Pleo-DI and Nao-DI ($p = 0.028$), Nao-KT and Nao-DI ($p = 0.023$), and Pleo-DI and Nao-KT ($p = 0.052$) (marginal).

MySkit offers a simple user-friendly interface that partici-

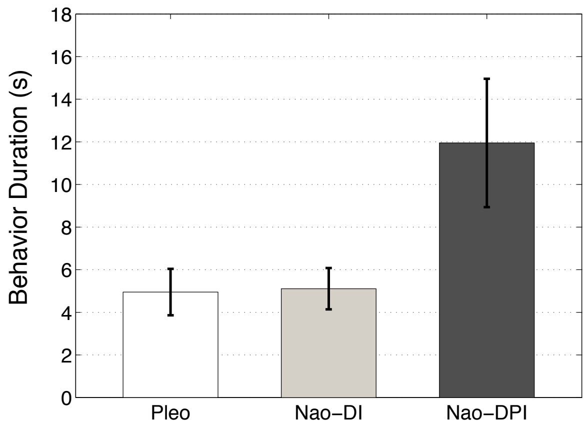


Fig. 8: Length of robot behaviors created using each system. Error bars show a 95% confidence interval on the mean.

pants quickly learned to use. Furthermore, Pleo has less DOF, so it is not overly complicated to program. On the other hand, Choregraphe has a more cumbersome DI display and Nao has more DOF than Pleo. As a result, users took longer to program emotions with Nao-DI. Conversely, the KT method offered in Nao-KT compensates for Nao's high DOF. Thus, users took less time to program behaviors with Nao-KT.

The ease of use offered by KT appears to have caused users to create longer behaviors with Nao-KT than with Pleo-DI and Nao-DI (Fig. 8). A one-way ANOVA shows a statistical difference between the length of the behaviors created with the various robotic systems ($F(2, 21) = 10$, $p < 0.001$). On average, the length of behaviors created using Nao-DI and Pleo-DI were just 5.0 seconds and 4.8 seconds, respectively, whereas behaviors created with Nao-KT were, on average, 12.0 seconds in length. There was not a statistical difference between the length of behaviors created using Nao-DI, and Pleo-DI ($p = 0.88$). However, there was a statistical difference between the length of behaviors created using Nao-KT and Nao-DI ($p = 0.018$) and between the length of behaviors created using Nao-KT and Pleo-DI ($p = 0.003$).

The post-experiment questionnaire corroborates these findings. In this questionnaire, participants were asked to rate the programming interface they used on the scale one to five (with five being "Excellent"). The average ratings for Nao-DI, Pleo-DI, and Nao-KT were 3.0, 4.0, and 4.25, respectively. Neither Pleo-DI nor Nao-KT received a single rating less than 4. A Kruskal Wallis test shows a statistically significant difference across systems ($\chi^2(2, N = 24) = 7.20$, $p = 0.027$). Pairwise comparisons show that Nao-DI was rated statistically lower than Nao-KT ($p = 0.03$) and marginal statistically lower than Pleo-DI ($p = 0.06$). There was no statistical difference between Pleo-DI and Nao-KT in this regard ($p = 0.149$).

Together, these results indicate that these off-the-shelf robotic systems do not allow novice users to quickly program robots to express recognizable emotions. Novice users typically required between 15 and 45 minutes to create robot behaviors of just a few seconds in duration. Furthermore, the target emotions of these behaviors were often indiscernible to others. Thus, we desire to improve these systems so that novice users can more easily produce good behaviors.

VI. User Study 2: Experimental Setup

In the first user study, we observed that participants were better able to express discernible emotions via verbal expressions (self-created audio clips) than via non-verbal expressions (robot body movements). However, neither the verbal expressions alone, the non-verbal expressions alone, nor the combined robot behaviors created by our novice participants were as high as one might desire. With these results in mind, we ran a follow-up user study. The goal of this user study was twofold. First, given the importance of verbal expressions in creating recognizable robot behaviors, we desire to identify how the robotic system can be enhanced to help users create better verbal expressions. Second, we desire to find methods to help users to create better non-verbal expressions.

While there exist many potential methods for achieving these goals, we focus in this paper on the importance of *context* in creating effective robot behaviors. As actors use contextual cues to improve their ability to act [40], we hypothesized that users creating robot behaviors can also benefit from context to express emotions more effectively through these behaviors. The context we provided in this second user study was the robot's expressions themselves. We tested whether the non-verbal expressions created by the user can be used to improve the creation of verbal expressions, and whether the verbal expressions created by the user can be used to improve the creation of non-verbal expressions. To do this, we defined and tested two programming procedures:

1) *Sync-Sound* – In this procedure, the verbal expression is recorded after the non-verbal expression is created. The user first creates the non-verbal expression. Next, the user records the desired verbal expression while viewing the robot as it executes the non-verbal expression already created. The user then merges the verbal expression with the non-verbal expression as before. We hypothesized that this would provide the user with better context when creating the verbal expression, which should lead to a more discernible verbal expression.
2) *Sound-First* – This procedure is the reverse of Sync-Sound. The user first records the verbal expression, after which he or she creates the non-verbal expression. We hypothesized that this procedure could potentially help give the user better context when creating non-verbal expressions, which should lead to more discernible non-verbal expressions.

To test these hypotheses, we conducted a second user study which also consisted of multiple phases. We now describe the experimental protocol for each phase of this study.

A. Phase 2.1: Creating Robot Behaviors

The first phase of this second user study was similar to Phase 1.1 of our first user study. In this phase, each user created robot behaviors to express two target emotions using one of the three procedures specified in Table III. These

TABLE III: Systems and procedures evaluated in the second user study.

Procedure	Robotic System	Order of Creation
Sound-First	Nao, Choregraphe, KT	Verbal then non-verbal
Sync-Sound	Nao, Choregraphe, KT	Non-verbal then verbal
Control	Nao, Choregraphe, KT	None specified

procedures were the Sync-Sound and Sound-First procedures discussed earlier, and the *Control* procedure used in the first user study. In the Control procedure, participants were not given directions on which expression (verbal or non-verbal) they should create first, nor were they instructed to create the non-verbal or verbal expression while replaying the corresponding verbal or non-verbal expression.

As the participants rated the Nao-KT system the highest and took less time to create behaviors using this system, the Nao-KT robotic system was the only system used in this study. Thus, participants in this phase used Choregraphe with KT to create body language for a Nao robot. We did upgrade the sound editor used by the participants in this study in order to provide the users with enhanced ability to modify their recorded audio. Other than these exceptions, the experimental protocol was identical to that of Phase 1.1.

Twenty-four individuals with little or no experience programming robots participated in this phase of the study. The average age of the participants was 25.6 years (coincidentally the same as in Phase 1.1, though there was no overlap in participants). The subjects were divided equally among procedures; eight participants were assigned to each system. As such, four behaviors of each emotion were created with each method. As in the first study, the order that behaviors were created to express each emotion was counter-balanced across all subjects to offset ordering effects.

B. Phases 2.2 and 2.3: Identifying Emotions

In the next two phases of this second user study, participants identified the emotions created in Phase 2.1 using the same experimental protocol as was used in Phases 1.2 and 1.3. In Phase 2.2, participants tried to identify the emotions expressed by the complete (combined verbal and non-verbal expressions) robot behaviors created in Phase 2.1. In Phase 2.3, participants identified the emotions expressed by the verbal and non-verbal expressions separately.

Ten students with an average age of 27.4 years participated in Phase 2.2. Each participant was asked to identify eight behaviors from each of the three procedures. Ten additional students with an average age of 24.7 years participated in Phase 2.3. Each behavior was assessed by five participants in both Phases 2.2 and 2.3. Behaviors were presented to participants in a random order.

VII. User Study 2: Results

We discuss the results of the second user study in two parts. First, we analyze how the various procedures affected

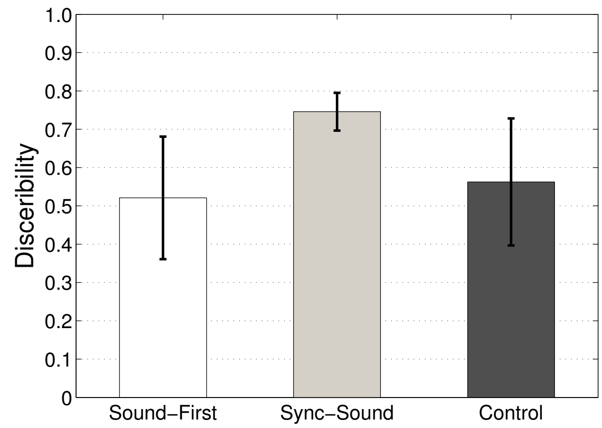


Fig. 9: Discernibility of verbal expressions created with the three procedures. Error bars show a 95% confidence interval on the mean

the discernibility of the target emotions of the robot behaviors created in Phase 2.1. We then discuss the overall salience of verbal and non-verbal expressions in this study.

A. Effects of Procedural Conditions

Recall that this user study was designed with two hypotheses in mind. First, we hypothesized that the Sync-Sound procedure would help users create better verbal expressions. Second, we hypothesized that the Sound-First procedure would help users create better non-verbal expressions. We evaluate each hypothesis in turn.

Fig. 9 shows the discernibility of verbal expressions produced using each of the procedures. The figure shows that participants created robot behaviors that conveyed more discernible verbal expressions in the Sync-Sound condition than in the other two conditions. Furthermore, the verbal expressions created using the Sync-Sound procedure were recognized 78.8% of the time, whereas verbal expressions created using the Sound-First and Control procedures were recognized 62.5% and 72.5% of the time, respectively.

An ANOVA shows that the procedure had a marginally statistically significant effect on the discernibility of emotions expressed by verbal expressions ($F(2, 21) = 2.97, p = 0.073$). Pairwise comparisons show that verbal expressions created using the Sound-First and Sync-Sound procedures ($p = 0.02$) and the Control and Sync-Sound procedures ($p = 0.057$) were statistically different and marginally statistically different, respectively. As expected, there was no statistical difference between the verbal expressions created using the Sound-First and Control procedures ($p = 0.728$). Thus, this data is consistent with the hypothesis that creating the verbal expressions while viewing the previously created robot movements increases the quality of the verbal expressions, albeit the results are only marginally statistically significant.

These results suggest that recording verbal expressions while observing non-verbal expressions might provide a better context for the user to create a more discernible verbal

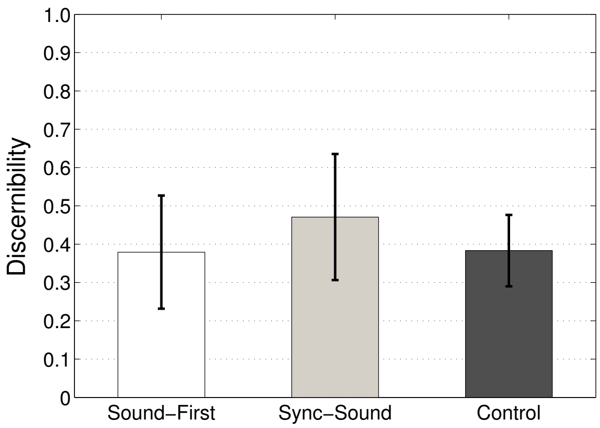


Fig. 10: Discernibility of non-verbal expressions created with the three procedures. Error bars show a 95% confidence interval.

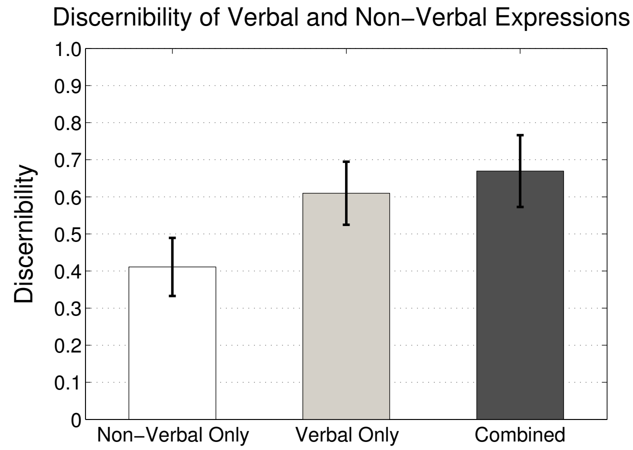


Fig. 11: Discernibility of emotions for verbal and non-verbal expressions viewed separately and for the combined expressions. Error bars show a 95% confidence interval on the mean.

expression. The reason for this improvement is not fully clear to us. We hypothesize that providing the non-verbal expression as context for the verbal expression creates a *feeling* within the user that allows him or her to better express himself or herself. An alternative, but not necessarily opposite, theory is that non-verbal expressions typically precede verbal expressions, and hence the ordering is simply more natural to people. We leave further investigation of this topic to future work.

Fig. 10 shows the discernibility of non-verbal expressions produced using each of the procedures. There was no statistically significant difference across procedures ($F(2, 21) = 0.54$, $p = 0.5924$). Thus, this data is not consistent with the hypothesis that creating verbal expressions before non-verbal expressions improves the quality of the non-verbal expressions. The reason for this result is also not fully clear to us. However, we hypothesize that it is difficult to synchronize body movements to an audio clip since body movements are typically much more difficult (and time consuming) to create than verbal expressions using Nao-KT. This means that it is difficult to create the movement while the sound is playing. As a result, the *feeling* produced by the verbal expression is lost to the user and does not provide a strong context. We leave further investigations of this topic to future work.

B. Discernibility of Emotions

Fig. 11 compares the discernibility of emotions expressed by the combined robot behaviors to the discernibility of emotions from the verbal and non-verbal expressions separately. These results mirror the findings of the first user study, which showed that verbal expressions were much more salient to users than non-verbal expressions. An ANOVA shows a statistically significant effect ($F(2, 21) = 9.27$, $p < 0.001$). Pairwise comparisons show non-verbal expressions to be statistically less recognizable without audio than when combined with audio ($p < 0.001$) and when compared to verbal expressions alone ($p = 0.002$). However, there was no statistical difference between the discernibility of combined behaviors and the verbal expressions alone ($p = 0.3682$). Furthermore, the recognition rate of a complete robot behavior was correlated with the recognition rate of its verbal expression ($r = 0.304; p < 0.001$), but not with that of its non-verbal expression ($r = 038; p = 0.562$).

To more fully investigate this finding, we conducted a fourth phase of the second user study (Phase 2.4) in which participants (again) tried to identify emotions from the behaviors created in Phase 2.1. However, in this phase, we paired non-verbal and verbal expressions together from different behaviors. We evaluated the following three forms of pairings:

1) *Bad Non-Verbal w/ Best Verbal* – Ten poor non-verbal expressions (20–40% recognition rates) paired with verbal expressions with the highest recognition rates.
2) *Best Non-Verbal w/ Wrong Verbal* – The ten non-verbal expressions with the highest recognition rates paired with verbal expressions with high recognition rates (80–100%), but that expressed an alternate emotion then the non-verbal expression.
3) *Best Non-Verbal w/ Right Verbal* – The ten non-verbal expressions with the highest recognition rates each paired with its corresponding verbal expression.

As in Phase 2.2, the participants in Phase 2.4 tried to identify emotions expressed by the full behaviors. From these responses, we computed the recognition rate and discernibility of emotions based on how evaluations corresponded to the emotion intended by the non-verbal expression.

The resulting discernibility measurements are shown in Fig. 12. The figure shows that a behavior with a highly recognizable verbal expression but a less recognizable non-verbal expression is highly discernible (Bad Non-Verbal w/ Best Verbal). Likewise, a behavior with a highly recognizable non-verbal expression is highly discernible (Best Non-Verbal w/ Right Verbal). However, if a verbal expression expresses a different emotion than a highly recognizable non-verbal expression, participants tended to identify the emotion as it was expressed by the verbal rather than the non-verbal expression,

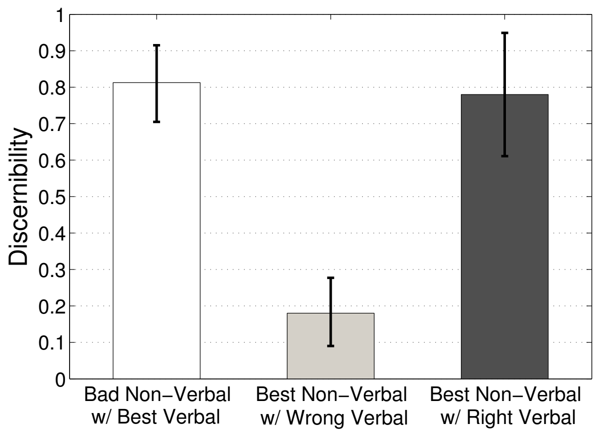


Fig. 12: Discernibility of the emotions expressed by non-verbal expressions when paired with various verbal expressions. Error bars show a 95% confidence interval on the mean.

as witnessed by the low discernibility in the Best-Non-Verbal-w/-Wrong-Verbal condition. These latter results reinforce our assessment that verbal expressions are more salient than non-verbal expressions for the expression of emotion through Nao.

VIII. Conclusions and Future Work

The ability of novice end users to program robots to express emotions via highly customized robot behaviors is an important attribute of robotic systems. In this paper, we addressed several questions related to building such computing systems via a series of user studies. In these user studies, participants created and evaluated robot behaviors designed to express emotions for Nao and Pleo robots.

Among other results, we made several observations from these studies. First, participants were sometimes able to express discernible emotions through robot behaviors created for both Nao and Pleo. Recognition rates for the created emotions mirrored the recognition rates of behaviors created by experts for a humanoid robot [15, 16]. However, the ability of participants to create recognizable behaviors was almost exclusively tied to the ability of users to create good verbal expressions. Non-verbal expressions created by users had low discernibility. Additionally, it typically took users between 15 and 45 minutes to create behaviors that lasted only a few seconds, suggesting that new programming methods are needed to allow novice users to easily create customized robot behaviors.

Finally, novice users appeared to be able to create more discernible verbal expressions when they created these expressions while viewing the non-verbal expressions that they had previously created. While these results were only marginally statistically significant, this suggests that programming interfaces and methods for creating robot behaviors should better utilize *context* to help users create more effective robot behaviors.

This work motivates the need for additional research in this area. Potential solutions could include (among others) leveraging verbal expressions in the creation of non-verbal expressions [41], Laban movement analysis [42], or in using human motion capture (HMC) rather than DI or KT. HMC could allow users to create verbal and non-verbal expressions simultaneously, though tracking errors and morphology concerns would need to be addressed [25]. Furthermore, we note that the studies reported in this paper are for robots with limited facial expressions. Interesting future work would be to conduct similar kinds of studies using robots that have greater facial expressions.

References

[1] N. Tomatis, R. Philippsen, B. Jensen, K. O. Arras, G. Terrien, R. Piguet, and R. Siegwart, “Building a fully autonomous tour guide robot: Where academic research meets industry,” in *Proceedings of the International Symposium on Robotics*, 2002.

[2] S. Thrun, M. Bennewitz, W. Burgard, A. B. Cremers, F. Dellaert, D. Fox, D. HŁhnel, C. Rosenberg, N. Roy, J. Schulte, and D. Schulz, “MINERVA: A second-generation museum tour-guide robot,” in *Proceedings of IEEE International Conference on Robotics and Automation*, pp. 1999–2005, 1999.

[3] C. Breazeal, A. Brooks, J. Gray, M. Hancher, J. McBean, D. Stiehl, and J. Strickon, “Interactive robot theatre,” *Communications of the ACM*, vol. 46, no. 7, pp. 76–85, 2003.

[4] A. Tapus, M. Mataric, and B. Scassellati, “The grand challenges in socially assistive robotics,” *IEEE Robotics and Automation Magazine*, vol. 14, no. 1, pp. 35–42, 2007.

[5] M. A. Goodrich, J. W. Crandall, and E. Barakova, “Teleoperation and beyond for assistive humanoid robots,” in *Reviews of Human Factors and Ergonomics* (D. Kaber, ed.), vol. 9, pp. 175–226, SAGE Publications, 2013.

[6] J. Blitch, R. Murphy, and T. Durkin, “Mobile semiautonomous robots for urban search and rescue,” in *Encyclopedia of Library and Information Science* (A. Kent and J. G. Williams, eds.), vol. 72, pp. 261–275, CRC Press, 2002.

[7] M. A. Goodrich, B. S. Morse, C. Engh, J. L. Cooper, and J. A. Adams, “Towards using unmanned aerial vehicles (UAVs) in wilderness search and rescue: Lessons from field trials,” *Interaction Studies*, vol. 10, no. 3, pp. 455–481, 2009.

[8] D. Keltner and A. M. Kring, “Emotion, social function, and psychopathology.,” *Review of General Psychology*, vol. 2, no. 3, pp. 320–342, 1998.

[9] C. Breazeal and B. Scassellati, “Infant-like social interactions between a robot and a human caretaker,” *Adaptive Behavior*, vol. 8, no. 1, pp. 49–74, 2000.

[10] C. Breazeal and B. Scassellati, “How to build robots that make friends and influence people,” in *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 858–863, 1999.

[11] H. Kozima, C. Nakagawa, and Y. Yasuda, “Interactive robots for communication-care: a case-study in autism therapy,” in *Proceedings of the 14th IEEE International Symposium on Robot and Human Interactive Communication*, pp. 341–346, 2005.

[12] K. Hideki, M. P. Michalowski, and C. Nakagawa, “Keepon: A playful robot for research, therapy, and entertainment,” *International Journal of Social Robotics*, vol. 1, pp. 3–18, 2009.

[13] B. Robins, K. Dautenhahn, and P. Dickerson, “From isolation to communication: A case study evaluation of robot assisted play for children with autism with a minimally expressive humanoid robot,” in *Proceedings of the 2nd International Conference on Advances in Computer-Human Interactions*, pp. 205–211, 2009.

[14] G. Pioggia, R. Igliozzi, M. Sica, M. Ferro, F. Muratori, A. Ahluwalia, and D. De Rossi, “Exploring emotional and imitational android-based interactions in autistic spectrum disorders,” *Journal of CyberTherapy and Rehabilitation*, vol. 1, no. 1, pp. 49–61, 2008.

[15] N. Endo, S. Momoki, M. Zecca, M. Saito, Y. Mizoguchi, K. Itoh, and A. Takanishi, "Development of whole-body emotion expression humanoid robot," in *Proceedings of the IEEE International Conference on Robotics and Automation*, pp. 2140–2145, 2008.

[16] M. Zecca, Y. Mizoguchi, K. Endo, F. Iida, Y. Kawabata, N. Endo, K. Itoh, and A. Takanishi, "Whole body emotion expressions for KOBIAN humanoid robot – preliminary experiments with different emotional patterns," in *Proceedings of the 18th IEEE International Symposium on Robot and Human Interactive Communication*, pp. 381–386, 2009.

[17] B. Scassellati, "Affective prosody recognition for human-robot interaction." Yale University, 2009.

[18] N. Giullian, D. Ricks, A. Atherton, M. Colton, M. Goodrich, and B. Brinton, "Detailed requirements for robots in autism therapy," in *Proceedings of the IEEE International Conference on Systems, Man, and Cybernetics*, pp. 2595–2602, 2010.

[19] J. Atherton and M. Goodrich, "Visual robot choreography for clinicians," in *Proceedings of the International Conference on Collaboration Technologies and Systems*, pp. 186 –189, 2011.

[20] M. Goodrich and D. Olsen, "Seven principles of efficient human robot interaction," in *Proceedings of the IEEE International Conference on Systems Man and Cybernetics*, pp. 3942–3948, 2003.

[21] B. Akgun, M. Cakmak, J. Yoo, and A. Thomaz, "Trajectories and keyframes for kinesthetic teaching: A human-robot interaction perspective," in *Proceedings of the 7th ACM/IEEE International Conference on Human-Robot Interaction*, pp. 391–398, 2012.

[22] C. Ott, D. Lee, and Y. Nakamura, "Motion capture based human motion recognition and imitation by direct marker control," in *Proceedings of the 8th IEEE-RAS International Conference on Humanoid Robots*, pp. 399–405, 2008.

[23] J. Shotton, T. Sharp, A. Kipman, A. W. Fitzgibbon, M. Finocchio, A. Blake, M. Cook, and R. Moore, "Real-time human pose recognition in parts from single depth images," *Communications of the ACM*, vol. 56, no. 1, pp. 116–124, 2013.

[24] R. El-laithy, J. Huang, and M. Yeh, "Study on the use of Microsoft Kinect for robotics applications," in *Proceedings of the IEEE/ION Position Location and Navigation Symposium*, pp. 1280–1288, 2012.

[25] B., Dariush, M. G. B. Jian, C. Goerick, and K. Fujimura, "Whole body humanoid control from human motion descriptors," in *IEEE International Conference on Robotics and Automation*, pp. 2677–2684, 2008.

[26] B. D. Argall, S. Chernova, M. Veloso, and B. Browning, "A survey of robot learning from demonstration," *Robotics and Autonomous Systems*, vol. 57, pp. 469–483, 2009.

[27] M. Vondrak, L. Sigal, J. Hodgins, and O. Jenkins, "Video-based 3D motion capture through biped control," *ACM Transactions on Graphics*, vol. 31, no. 4, pp. 27:1–27:12, 2012.

[28] S. Calinon and A. Billard, "Statistical learning by imitation of competing constraints in joint space and task space," *Advanced Robotics*, vol. 23, no. 15, pp. 2059–2076, 2009.

[29] B. Argall, B. Browning, and M. Veloso, "Learning by demonstration with critique from a human teacher," in *Proceedings of the 2nd ACM/IEEE International Conference on Human-Robot Interaction*, pp. 57–64, 2007.

[30] B. Robins, K. Dautenhahn, R. Boekhorst, and A. Billard, "Effects of repeated exposure to a humanoid robot on children with autism," in *Designing a More Inclusive World* (S. Keates, J. Clarkson, P. Langdon, and P. Robinson, eds.), pp. 225–236, Springer London, 2004.

[31] T. Lourens and E. Barakova, "User-friendly robot environment for creation of social scenarios," in *Proceedings of the 4th International Conference on Interplay between Natural and Artificial Computation*, pp. 212–221, 2011.

[32] C. Breazeal and J. Velsquez, "Toward teaching a robot 'infant' using emotive communication acts," in *Workshop on Socially Situated Intelligence held at SAB'98*, 1998.

[33] M. Diaz, N. Nuno, J. Saez-Pons, D. Pardo, and C. Angulo, "Building up child-robot relationship for therapeutic purposes: From initial attraction towards long-term social engagement," in *Proceedings of the 11th IEEE International Conference on Automatic Face and Gesture Recognition*, pp. 927–932, 2011.

[34] E. Pot, J. Monceaux, R. Gelin, and B. Maisonnier, "Choregraphe: a graphical tool for humanoid robot programming," in *Proceedings of the 18th IEEE International Symposium on Robot and Human Interactive Communication*, pp. 46–51, 2009.

[35] P. Eckman, "An argument for basic emotions," *Cognition and Emotion*, vol. 6, pp. 169–200, 1992.

[36] E. Austin, "Handbook of cognition and emotion: Edited by tim dalgleish and mick power," *The British Journal of Psychiatry*, vol. 176, no. 5, 2000.

[37] M. Coulson, "Attributing emotion to static body postures: Recognition accuracy, confusions, and viewpoint dependence," *Journal of Nonverbal Behavior*, vol. 28, pp. 117–139, 2004.

[38] A. S. Shaarani and D. M. Romano, "Perception of emotions from static postures," in *Affective Computing and Intelligent Interaction* (A. Paiva, R. Prada, and R. Picard, eds.), vol. 4738 of *Lecture Notes in Computer Science*, pp. 761–762, Springer Berlin Heidelberg, 2007.

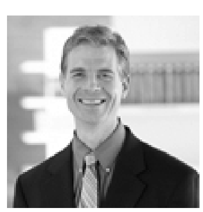
[39] E. Barakova and T. Lourens, "Expressing and interpreting emotional movements in social games with robots," *Personal Ubiquitous Comput.*, vol. 14, pp. 457–467, 2010.

[40] H. G. Schmidt, H. P. Boshuizen, and G. van Breukelen., "Long-term retention of a theatrical script by repertory actors: the role of context," *Memory*, vol. 10, no. 1, pp. 21–28, 2002.

[41] A. Lim, T. Ogata, and H. G. Okuno, "Converting emotional voice to motion for robot telepresence," in *Proceedings of the IEEE International Conference on Humanoid Robots*, pp. 472–479, 2011.

[42] T. Lourens, R. van Berkel, and E. Barakova, "Communicating emotions and mental states to robots in a real time parallel framework using laban movement analysis," *Robotics and Autonomous Systems*, vol. 58, no. 12, pp. 1256–1265, 2010.

Vimitha Manohar Vimitha received the M.Sc. degree from Masdar Institute, Abu Dhabi in 2011. She is currently a Ph.D. candidate at the Masdar Institute of Science and Technology, where she is applying machine learning techniques for classification of bacterial metagenomes with reference to the Earth Microbiome Project in bioinformatics. Her research interests include data-mining, machine learning, and human-machine systems.



Jacob W. Crandall received the B.S., M.S., and Ph.D. degrees in Computer Science from Brigham Young University, Provo, UT, in 2001, 2004, and 2006, respectively. Jacob is currently an Associate Professor in the Department of Electrical Engineering and Computer Science at the Masdar Institute of Science and Technology. Previously, he was a postdoctoral associate in the Department of Aeronautics & Astronautics at the Massachusetts Institute of Technology. His research interests include human-machine systems, multi-agent systems, machine learning, game theory, and robotics.