



A First Step toward the Automatic Understanding of Social Touch for Naturalistic Human–Robot Interaction

Merel M. Jung*, Mannes Poel, Dennis Reidsma and Dirk K. J. Heylen

Human Media Interaction, University of Twente, Enschede, Netherlands

Social robots should be able to automatically understand and respond to human touch. The meaning of touch does not only depend on the form of touch but also on the context in which the touch takes place. To gain more insight into the factors that are relevant to interpret the meaning of touch within a social context we elicited touch behaviors by letting participants interact with a robot pet companion in the context of different affective scenarios. In a contextualized lab setting, participants (n = 31) acted as if they were coming home in different emotional states (i.e., stressed, depressed, relaxed, and excited) without being given specific instructions on the kinds of behaviors that they should display. Based on video footage of the interactions and interviews we explored the use of touch behaviors, the expressed social messages, and the expected robot pet responses. Results show that emotional state influenced the social messages that were communicated to the robot pet as well as the expected responses. Furthermore, it was found that multimodal cues were used to communicate with the robot pet, that is, participants often talked to the robot pet while touching it and making eye contact. Additionally, the findings of this study indicate that the categorization of touch behaviors into discrete touch gesture categories based on dictionary definitions is not a suitable approach to capture the complex nature of touch behaviors in less controlled settings. These findings can inform the design of a behavioral model for robot pet companions and future directions to interpret touch behaviors in less controlled settings are discussed.

Keywords: social touch, human-robot interaction, robot pet companion, multimodal interaction, touch recognition, behavior analysis, affective context

1. INTRODUCTION

Touch plays an important role in establishing and maintaining social interaction (Gallace and Spence, 2010). In interpersonal interaction, this modality can be used to communicate emotions and other social messages (Jones and Yarbrough, 1985; Hertenstein et al., 2006, 2009). More recently, the study of social touch was also extended to interaction with humanoid and robotic animals (e.g., Knight et al., 2009; Kim et al., 2010; Yohanan and MacLean, 2012; Cooney et al., 2015). In order to make these interactions more natural, robots should be able to understand and respond to human touch.

A social robot needs to sense and recognize different touch gestures (e.g., Kim et al., 2010; Silvera-Tawil et al., 2012; Altun and MacLean, 2015; Jung et al., 2015, 2016) and should be able

OPEN ACCESS

Edited by:

Hatice Gunes, University of Cambridge, UK

Reviewed by:

Christian Becker-Asano, Robert Bosch, Germany Maria Koutsombogera, Trinity College Dublin, Ireland

> *Correspondence: Merel M. Jung m.m.jung@utwente.nl

Specialty section:

This article was submitted to Human-Media Interaction, a section of the journal Frontiers in ICT

Received: 31 December 2016 Accepted: 27 February 2017 Published: 17 March 2017

Citation:

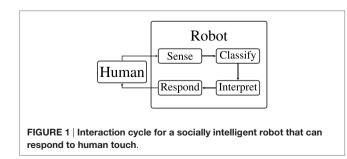
Jung MM, Poel M, Reidsma D and Heylen DKJ (2017) A First Step toward the Automatic Understanding of Social Touch for Naturalistic Human–Robot Interaction. Front. ICT 4:3. doi: 10.3389/fict.2017.00003 to interpret touch in order to respond in an appropriate manner (see Figure 1). Perhaps robot seal Paro is the most famous example of a social robot that responds to touch (Wada and Shibata, 2007). Paro is equipped with touch sensors with which it distinguishes between soft touches (which are always interpreted to be positive) and rough touches (which are always interpreted to be negative) (Wada and Shibata, 2007). This interpretation of touch is oversimplified as the complexity of the human tactile system allows for touch behaviors to vary not only depending on the intensity but also based on movement, velocity, abruptness, temperature, location, and duration (Hertenstein et al., 2009). Moreover, the meaning of touch can often not be inferred from the type of touch alone but is also dependent on other factors such as concurrent verbal and non-verbal behavior, the type of interpersonal relationship (Heslin et al., 1983; Suvilehto et al., 2015), and the situation in which the touch takes place (Jones and Yarbrough, 1985). Although previous research (Heslin et al., 1983; Hertenstein et al., 2006, 2009) indicated that there is no one-to-one mapping of touch gestures to a specific meaning of touch, touch can have a clear meaning in a specific context (Jones and Yarbrough, 1985).

The current study focuses on (touch) interaction with a robot pet companion. According to Veevers, a pet companion can fulfill different roles in the life of humans, a pet can facilitate interpersonal interaction or can even serve as a surrogate for interpersonal interaction, and expensive and/or exotic pets can be owned as a status symbol (Veevers, 1985). Furthermore, interaction with pet companions is associated with health benefits, and more recent studies indicate that these effects also extend to interaction with robot pets (Eachus, 2001; Banks et al., 2008). Although touch is a natural way to interact with real pets, currently commercially available robot pets such as Paro (Wada and Shibata, 2007), Hasbro's companion pets,¹ and JustoCat² are equipped with only a few touch sensors and do not interpret different types of touch within context.

We argue that the recognition and interpretation of touch consists of three levels: (1) low-level touch parameters such as intensity, duration, and contact area; (2) mid-level touch gestures such as pat, stroke, and tickle; and (3) high-level social messages such as affection, greeting, and play. To automatically understand social touch, research focuses on investigating the connection between these levels. Current studies in the domain of social

¹http://joyforall.hasbro.com.

²http://justocat.com.



touch for human-robot interaction focused mainly on highly controlled settings in which users were requested to perform different touch behaviors, one at the time, according to predefined labels (e.g., Cooney et al., 2012; Silvera-Tawil et al., 2012, 2014; Yohanan and MacLean, 2012; Jung et al., 2015, 2016). In this study we focus on the latter two levels as we are interested in the meaning of touch behaviors. To gain more insight into the factors that are relevant to interpret touch behaviors within social context, we opted to elicit touch behaviors by letting participants act out four scenarios in which they interacted with a robot pet companion in different emotional states. Moreover, in contrast to most previous studies, participants could freely act out the given scenarios with the robot pet within the confined space of a living room setting.

In this paper, we present contributions in two areas. First, we explore the use of touch behaviors as well as the expressed social messages and expected robot pet responses in different affective scenarios. Second, we reflect upon the challenges of the segmentation and labeling of touch behaviors in a less controlled setting in which no specific instructions are given on the kinds of (touch) behaviors that should be displayed. We address the first contribution with the following three research questions. (RQ1) What kinds of touch gestures are used to communicate with a robot pet in the different affective scenarios? (RQ2) Which social messages are communicated, and what is the expected response in the different affective scenarios? (RQ3) What other social signals can aid the interpretation of touch behaviors? Furthermore, we reflect upon our effort to segment and label touch behaviors in a less controlled setting with the fourth research question. (RQ4) How well do annotation schemes work in a contextualized lab situation?

The remainder of the article is structured as follows. Related work on the meaning of social touch in both interpersonal and human-robot interaction will be discussed in the next section followed by the description of the materials and methods for the presented study. Then, the results will be provided and discussed in the subsequent sections. Conclusions will be drawn in the last section.

2. RELATED WORK

Previous studies have looked into the meaning of touch in both interpersonal interaction (Jones and Yarbrough, 1985) and human-robot interaction with either a humanoid robot (Kim et al., 2010; Silvera-Tawil et al., 2014; Cooney et al., 2015) or a robot animal (Knight et al., 2009; Yohanan and MacLean, 2012). In a diary study on the use of interpersonal touch, different meanings of touch were categorized based on the participants' verbal translations of the touch interactions (Jones and Yarbrough, 1985). Seven main categories were distinguished: positive affect touches (e.g., support), playful touches (e.g., playful affection), control touches (e.g., attention-getting), ritualistic touches (e.g., greeting), hybrid touches (e.g., greeting/affection), task-related touches (instrumental intrinsic), and accidental touches. Interestingly, there was a lack of reports on negative interpersonal touch interaction. Within these categories, common contextual factors were identified such as the type of touch, any accompanying verbal statement, and the situation in which the touch took place. It was found that depending on the context, a specific form of touch can have multiple meanings and that different forms of touch can have a similar meaning. Furthermore, touch was found to be often preceded, accompanied, or followed by a verbal statement.

In a study on human-robot interaction, participants were asked to indicate which touch gestures they were likely to use to communicate emotional states to a cat-sized robot animal (Yohanan and MacLean, 2012). Gestures that were judged to be likely used were performed sequentially on the robot. Participants expected that the robot's emotional response was either similar or sympathetic to the emotional state that was communicated. The nature of the touch behavior was found to be friendly as no aggressive gestures (e.g., slap or hit) were used even when negative emotions were communicated. Five categories of intent were distinguished based on touch gesture characteristics that could be mapped to emotional states: affectionate, comforting, playful, protective, and restful. Also, video segments of the touch gestures were annotated to characterize the gestures based on the point of contact, intensity, and duration revealing differences between touch gestures and their use in different emotional states. In follow-up research, the touch sensor data recorded in this study (i.e., Yohanan and MacLean, 2012) were used to classify 26 touch gestures and 9 emotional states using random forests (Altun and MacLean, 2015). Between-subjects emotion recognition of 9 emotional states yielded an accuracy of 36%, while within-subjects the accuracy was 48%. Between-subjects touch gesture recognition of 26 gestures yielded an accuracy of 33%. Furthermore, the authors' results indicated that accurate touch gesture recognition could improve affect recognition.

In other work, Kim et al. (2010) instructed participants to use four different touch gestures to give positive or negative feedback to a humanoid robot while playing a game. A model was trained to infer whether a touch gesture was meant as a positive or a negative reward for the robot. It was found that participants used "pat" and "rub" to give positive feedback and "hit" to give negative feedback, while "push" could be used for both although the touch gesture was mostly used for negative feedback. Knight et al. (2009) argued for the importance of body location as contextual factor to infer the meaning of touch. The authors made the distinction between what they called *symbolic gestures*, which have social significance based on the involved body location(s) (e.g., footrub and hug) and body location-independent *touch subgestures* (e.g., pat and poke).

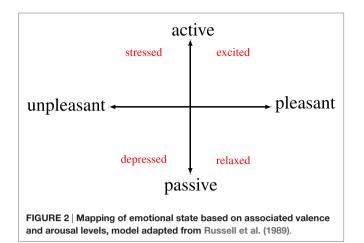
Although previous studies indicate that there is a link between touch gestures and the higher level social meaning of touch, Silvera-Tawil et al. (2014) argued that the meaning of touch could also be recognized directly based on characteristics from touch sensor data and other factors such as the context and the touch location. In their effort to automatically recognize emotions and social messages directly from sensor data without first recognizing the used touch gestures, participants were asked to perform six basic emotions: anger, disgust, fear, happiness, sadness, and surprise on both a mannequin arm with an artificial skin and a human arm. In addition, six social messages were communicated: acceptance, affection, animosity, attention-getting, greeting, and rejection. Recognition rates for the emotions were 46.9% for the algorithm and 51.8% for human classification. The recognition rates for the social messages were found to be slightly higher, yielding accuracies of 49.7 and 62.1% for the algorithm and human classification, respectively.

Some attempts have been made to study touch interaction in a less controlled setting, for example, Noda et al. (2007) elicited touch during the interaction with a humanoid robot by designing a scenario in which participants used different touch gestures to communicate a particular social message such as greeting the robot by shaking hands, playing together by tickling the robot, and hugging the robot to say goodbye at the end of the interaction. Results showed an accuracy of over 60% for the recognition of the different touch behaviors that were performed within the scenario. In another study on the use of touch in multimodal human-robot interaction, participants were given various reasons to interact with a small humanoid robot such as giving reassurance, getting attention, and giving approval (Cooney et al., 2015). The robot was capable of recognizing touch, speech, and visual cues, and participants were free to use different modalities. Also, participants rated videos in which a confederate interacted with the robot using different modalities. Results showed that touch was often used to communicate with the robot and that touch was especially important for expressing affection. Furthermore, playing with the robot and expressing loneliness were deemed more suitable than displaying negative emotions.

To summarize, previous studies illustrate that touch can be used to express and communicate different kinds of affective and social messages (Jones and Yarbrough, 1985; Yohanan and MacLean, 2012; Silvera-Tawil et al., 2014; Cooney et al., 2015). Moreover, touch gestures that were used to communicate were often positive in nature, and their meaning is dependent on the context such as one's emotional state (Jones and Yarbrough, 1985; Yohanan and MacLean, 2012; Cooney et al., 2015). These findings confirm that currently available robot pet companions, such as Paro, which only distinguishes between positive and negative touch, are not sufficiently capable of understanding and responding to people in a socially appropriate way. Furthermore, there are indications that other modalities might be helpful in interpreting the social messages as touch behavior generally does not occur in isolation (Jones and Yarbrough, 1985; Cooney et al., 2015). For the reasons outlined above, we opted to study interactions between a human and a robot pet companion in the context of different emotional states in a contextualized lab situation.

3. MATERIALS AND METHODS

In this study we elicited interactions between a human and a robot pet companion in a lab-build living room setting. Participants were instructed to act as if they would come home in different emotional states (i.e., stressed, depressed, relaxed, and excited). These four emotional states were chosen as they span opposite ends of the valence and arousal scale (see **Figure 2**): stressed (low valence, high arousal), depressed (low valence, low arousal), relaxed (high valence, low arousal), and excited (high valence, high arousal) (Russell et al., 1989). Furthermore, similar emotional states have been used in a more controlled research setting before, and the



results from this study indicate that emotional state influences touch behavior as well as the expected robot response (Yohanan and MacLean, 2012). To gain more insight into the factors that are relevant to interpret touch within a social context, we annotated touch behaviors from video footage of the interactions. Also, a questionnaire was administered and interviews were conducted to interpret the high-level meaning behind the interactions and get insight into the responses that would be expected from the robot pet.

3.1. Participants

In total 31 participants (20 males, 11 females) volunteered to take part in the study. The age of the participants ranged from 22 to 64 years (M = 34.3; SD = 12.8), and 28 were right-handed, 2 left-handed, and 1 ambidextrous. All studied or worked at the University of Twente in the Netherlands. Most (21) had the Dutch nationality; others were Belgian, Ecuadorean, English, German (2×), Greek, Indian, Iranian, Italian, and South Korean. This study was approved by the ethics committee of the Faculty of Electrical Engineering, Mathematics and Computer Science of the University of Twente. All research participants provided written informed consent in accordance with the Declaration of Helsinki.

3.2. Apparatus/Materials

3.2.1. Living Room Setting

The living room setting consisted of a space of approximately 23 m² containing a small couch, a coffee table, and two plants (see **Figure 3**, left). Two camcorders were positioned facing the couch at an approximately 45° angle to record the interactions (50 fps, 1,080 p). To allow participants to interact freely with the robot pet (i.e., no wires) and have a controlled interaction (i.e., no unpredictable robot behavior), a stuffed animal dog was used as a proxy for a robot pet. The robot pet (35 cm; in a laying position) was positioned on the couch at the far end from the door facing the table (see **Figure 3**).

3.2.2. Questionnaire

The questionnaire was divided into two parts. Part one was completed before the interview was conducted and part two after

the interview. Part one consisted of demographics: gender, age, nationality, occupation, and handedness followed by six questions about the reenactment of the scenarios rated on a 4-point Likert scale ranging from 1 (strongly disagree) to 4 (strongly agree). Four questions were about the participants' ability to imagine themselves in the scenarios: "I was able to imagine myself coming home feeling stressed/depressed/relaxed/excited." The other two questions were about the robot pet: "I was able to imagine that the pet was a functional robot" and "I based my interaction with the robot pet on how I interact with a real animal."

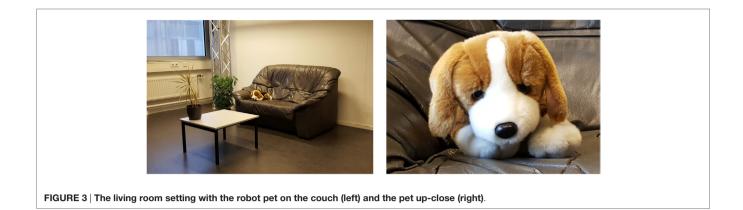
The second part consisted of a questionnaire about the expectations of living with a robot pet, which was based on the 11-item Comfort from Companion Animals Scale (CCAS) (Zasloff, 1996). Participants were asked to imagine that they would get a robot pet like the one in the study as a gift. This robot pet can react to touch and verbal commands. Participants were asked to answer the questions about the role they expect the robot pet would play in their life. The questions from the CCAS were adjusted to fit the purpose of the study, for example, the item "my pet provides me with companionship" was changed to "I expect my robot pet to provide me with companionship." Items were rated on a 4-point Likert scale ranging from 1 (strongly disagree) to 4 (strongly agree), as all items were phased positively a higher score indicates greater expected comfort from the robot pet.

3.2.3. Interview

A semi-structured interview was conducted between the first and the second part of the questionnaire. The video footage of their reenactment of the scenarios was shown to the participants, and they were asked to answer the following questions after watching each of the four scenario fragments: (1) "What message did you want to communicate to the robot?" (2) "What response would you expect from the robot?" (3) "How could the robot express this?" The participant, the interviewer, and the computer screen were recorded during the interview using a camcorder.

3.3. Procedure

Upon entering the room in which the study took place, the participant was welcomed and was asked to read the instructions and sign an informed consent form. Then, participants were taken into the hallway where they received the instructions for the example scenario in which they were asked to act out coming home in a neutral mood. If the instructions were clear, participants were asked to interact with a robot pet by acting out four different scenarios, one by one, in which they would come home in a particular emotional state, feeling stressed, depressed, relaxed, or excited. The study had a within-subject design; instructions for each of the scenarios were given to each of the participants in random order. In each scenario, the participant was instructed to enter the "living room," sit down on the couch, and act out the scenario as he/she sees fit. Participants were instructed to focus on the initial interaction as the robot pet would not respond (\approx 30 s were given as a guideline); however, the duration of the interaction was up to the participant who was instructed to return to the hallway when he or she finished an interaction. When the participant had returned to the hallway at the end of an interaction the next scenario was provided.



After the last scenario, the participant was asked to fill out a questionnaire asking about demographic information and about acting out the scenarios. Then, the video footage of their reenactment of the scenarios was shown to the participant, and an interview was conducted on these interactions. After the interview, the participants completed the second part of the questionnaire about their expectations if they would own a functional robot pet. The entire procedure took approximately 20 min for each participant. At the end of the study, participants were offered a candy bar to thank them for participating.

3.4. Data Analysis

3.4.1. Questionnaire

The questionnaire data were analyzed using IBM SPSS Statistics version 22. The median values and the 25th and 75th percentiles (i.e., Q1 and Q3, respectively) were calculated for the questions about the reenactment of the scenarios. The ratings on the items of the expected comfort from the robot pet scale were summed before calculating these descriptive statistics. Additionally, a Friedman test was conducted to check whether there was a statistical difference between the perceived ability of the participants to imagine themselves in the different scenarios. The significance threshold was set at 0.05, and the exact *p*-value is reported for a two-tailed test.

3.4.2. Annotation of Scenario Videos

The video footage from the two cameras were synced and put together in a split screen video before annotation. Videos were coded by two annotators, which included one of the authors (Merel M. Jung), henceforth "the first coder," using the ELAN³ annotation software.

For the segmentation of touch behaviors we followed a method that is commonly used to segment signs and co-speech gestures into movement units, which in the simplest form consist of three phases: a preparation phase, an expressive phase, and a retraction phase (Kendon, 1980; Kita et al., 1997). The onset of a movement unit is defined at the first indication of the initiation

of a movement that is usually preceded by the departure of the hand's resting position. The end of a movement unit is defined as the moment when the hand makes first contact with a resting surface such as the lap or an arm rest. Similarly, the touch actions were segmented by the first coder from the moment that the participant reached out to the robot pet to make physical contact until the contact with the pet was ended and the hands of the participants returned to the resting position. Per touch action segment, the following information was coded by the two annotators in a single annotation tier: the performed sequence of touch gestures and the robot pet's body part(s) on which each touch gesture was performed (see Figure 4). The touch gesture categories consisted of the 30 touch gestures plus their definitions from the touch dictionary of Yohanan and MacLean (2012), which is a list of plausible touch gestures for interaction with a robot pet. Furthermore, based on observations we added an additional category for puppeteering, which was defined as "participant puppeteers the robot pet to pretend that it is moving on its own" and to reduce forced-choice we added other, which was defined as "the touch gesture performed cannot be described by any of the previous categories." The robot pet's body parts were divided into six categories: head (i.e., back, top, and sides of the head and ears), face (i.e., forehead, eyes, nose, mouth, cheeks, and chin), body (i.e., neck, back, and sides), belly, legs, and tail.

Coding the touch behaviors that were performed during each touch segment proved to be difficult. Both annotators were often unsure when to define the start of a new touch gesture as gestures were often followed up in quick succession. Furthermore, hybrid forms of several touch gestures were often observed resulting in difficulties to categorize the touch behavior into one of the categories. In Table 1 some of the touch gestures are listed that were frequently observed but that were also difficult to distinguish based on their dictionary definitions. These touch gestures are all of relatively long duration compared to quick gestures such as pat and slap, and all include movement across the contact area. The distinguishing features are based on the gesture's intensity, human contact point (e.g., whole hand vs. fingernails), and the movement pattern (e.g., back and forth or seemingly random). An example of commonly encountered confusion was in cases where the hand was moved repeatedly back and forth on the fur of the robot pet, which indicated the use of a rub gesture, while the use of gentle pressure seemed to indicate a stroke-like gesture.

³Max Planck Institute for Psycholinguistics, The Language Archive, Nijmegen, The Netherlands; http://tla.mpi.nl/tools/tla-tools/elan.

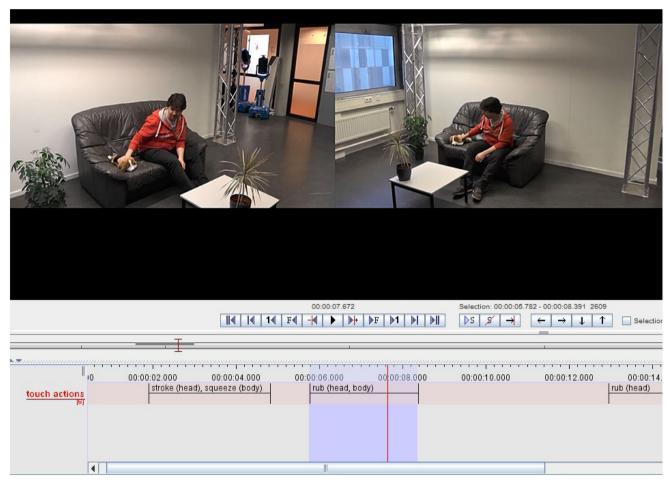


FIGURE 4 | Screenshot of the annotation process showing the annotation tier in which the touch gestures and the body location on the robot pet are annotated.

TABLE 1 | Example touch gesture categories with definitions, adapted from Yohanan and MacLean (2012).

Gesture label	Gesture definition
Massage	Rub or knead the robot pet with your hands
Rub	Move your hand repeatedly back and forth on the fur of the
	robot pet with firm pressure
Scratch	Rub the robot pet with your fingernails
Stroke	Move your hand with gentle pressure over the robot pet's fur, often repeatedly
Tickle	Touch the robot pet with light finger movements

Furthermore, the use of video footage to code touch gestures made it difficult to determine the exact point of contact, which is the only differentiating feature to distinguish between a rub and a scratch gesture based on these definitions.

Even after several iterations of revisiting the codebook in order to clarify what the distinguishing features of several touch gestures are, it was still not possible to reach an acceptable level of agreement. Difficulties were caused by a mixture of touch events that were hard to observe on video and differences in interpretation by the annotators, which included both the segmentation of individual touch gestures (i.e., within the larger predefined segments) and the assignment of labels, despite the commonly developed annotation scheme. Furthermore, as one touch segment could consist of a sequence of touch gestures, it was difficult to calculate the inter-rater reliability (i.e., Cohen's kappa) as the number of touch gestures could differ per coder. The location of the touch gestures on the robot pet's body was related to the coding of the touch gestures themselves, and therefore it was also not possible to reach an acceptable agreement on this part.

Due to the difficulties described above we decided instead to coarsely describe the interactions in the results section based on the modalities that the participants used to communicate to the robot pet. Also, a Friedman test was conducted to check whether there was a statistical difference between the duration of the interactions in the different scenarios. The significance threshold was set at 0.05, and the exact *p*-value is reported for a two-tailed test. The implications of the findings from the annotation process will be explicated in the discussion section.

3.4.3. Interview

The interview answers were grouped per scenario based on common themes. The data were split into two parts. (1) Information on the social messages (and possible behaviors to express those) that were communicated by the participant to the robot pet. (2) Information on the expected messages and behaviors that were expected to be communicated by the robot pet. Themes were labeled, and the number of participants that mentioned the specific topic was counted for each scenario. Furthermore, the communicated social messages for each scenario were mapped to the expected responses from the robot pet to look for frequently occurring patterns.

4. RESULTS

4.1. Questionnaire

Participants' rating of their ability to imagine themselves in the four different scenarios on a scale ranging from 1 (strongly disagree) to 4 (strongly agree) were the following: stressed (*Mdn* (Q1, Q3) = 3 (3, 3)), depressed (*Mdn* (Q1, Q3) = 3 (2, 3)), relaxed (*Mdn* (Q1, Q3) = 3 (3, 3)), and excited (*Mdn* (Q1, Q3) = 3 (2, 4)). There was no statistically significant difference between the ratings of the scenarios ($\chi^2(3) = 3.297$, p = 0.352). Median (Q1, Q3) perceived ability to imagine the pet as a functional robot was 2 (2, 2) and the statement "I based my interaction with the robot pet on how I interact with a real animal" was rated at 3 (2, 4). The total scores for the expected comfort from the robot pet ranged from 18 to 37 (*Mdn* (Q1, Q3) = 30(26, 35)), possible total scores ranged from 11 to 44 where a higher score indicated greater expected comfort.

4.2. Observations from the Scenario Videos

Between the different scenarios there were some differences in the level of interaction with the robot pet (see **Table 2**). Participants often used both touch and speech to communicate with the robot pet. Almost all participants used at least the touch modality to communicate, few exceptions occurred in the low valence scenarios (i.e., stressed and depressed). Examples of touch behaviors that were observed were participants sitting next to the robot pet on the couch while touching it using stroking-like gestures, hugging the pet, and having the robot pet sit on their laps while resting a hand on top of it. Speech was most prevalent in the excited scenario, while it was least prevalent in the

TABLE 2 | Number of participants that engaged in different levels of interaction with the robot pet per scenario.

Interaction type	Emotional state				
	Stressed	Depressed	Relaxed	Excited	
No interaction	3	3	0	0	
Speech only	2	0	0	0	
Touch only	8	12	13	7	
Touch + speech	18	16	18	24	
Sum	31	31	31	31	

depressed scenario. Observed behaviors included participants using speech to greet the robot pet when entering, talk about their day, express their emotional state, and show interest in the pet. Some instances of pet-directed speech were observed as well. Another notable observation was that participants oriented the robot pet to face them indicating that they wanted to make eye contact. Furthermore, some participants incorporated the use of their mobile phone in the scenarios, for example, to indicate that they would be preoccupied with their own activities (e.g., sending text messages to friends), to take a picture of the robot pet or to watch online videos together. Others engaged in fake activities with imaginary objects such as playing catch or watching TV together.

The duration of an interaction was measured as the time in seconds between the start of the interaction (i.e., opening the door to enter the living room) and the end (i.e., closing the door after leaving the room). Overall, the duration of the interactions ranged between 17 and 112 s. There was a statistically significant difference in the duration of interaction between the four scenarios ($\chi^2(3) = 16.347$, p = 0.001). A *post hoc* analysis with Wilcoxon signed-rank tests was conducted with a Bonferroni correction applied, resulting in a significance level set at p < 0.008. The median (Q1, Q3) duration in seconds for each of the scenarios was stressed 41 (29, 55), depressed 42 (32, 55), relaxed 42 (32, 53), and excited 35 (28, 45). The duration of interaction in the excited was significantly shorter compared to the other scenarios: stressed (Z = -2.968, p = 0.002), depressed (Z = -3.875, p < 0.001), and relaxed (Z = 3.316, p = 0.001). The other scenarios did not differ significantly (all p's >0.008).

4.3. Interview

In general, participants mostly watched the whole scenario before answering the questions, while others commented on their behavior right away. Also, some participants mentioned at the beginning that they felt a bit awkward to watch themselves on video. The social messages that were communicated by the participant to the robot pet and messages that were expected to be communicated by the robot pet are listed for each scenario in **Tables 3** and **4**, respectively. **Table 5** shows the mapping between the two most frequently communicated social messages for each scenario and the most common expected responses from the robot pet to these messages. We will further discuss the interview results based on these mappings.

4.3.1. Stressed

In the stressed scenario, most participants wanted to communicate that they were stressed by indicating to the robot pet that they had lots of things to do or that they were preoccupied with something (n = 17). Notably, some of these participants involved the robot pet as a way to regulate their emotions by touching the pet as a means of distraction. In response, some of these participants wanted company from the robot pet by staying close and through physical interaction (n = 6). Importantly, the pet's behavior should be calm, and the robot should not be too demanding. Other participants wanted support from the robot pet by calming them down and providing comfort (n = 6).

TABLE 3 | Social messages that were communicated to robot pet for each scenario.

Emotional state			
Stressed	Depressed	Relaxed	Excited
Express emotional state (17)	Seek emotional support (11)	Enjoy company (14)	Express emotional state (15)
Do not want to interact (6)	Express emotional state (8)	Express emotional state (8)	Actively seek interaction (11)
Acknowledge (3)	Do not want to interact (6)	Acknowledge (7)	Enjoy company (4)
Seek emotional support (3)	Want to interact (3)	No expectations (2)	Do not want to interact (1)
Actively seek interaction (2)	Acknowledge (1) Enjoy company (1) No expectations (1)		

The number of participants is in parentheses.

TABLE 4 | Social messages that the robot pet is expected to communicate for each scenario.

Emotional state			
Stressed	Depressed	Relaxed	Excited
Keep company (8)	Keep company (12)	Keep company (12)	Pick up the mood (24)
Provide emotional support (7)	Provide emotional support (11)	Pick up the mood (7)	Engage in interaction (6)
Focus on own needs (6)	Engage in interaction (4)	Engage in interaction (5)	Show appreciation (1)
Understand the situation (5)	Focus on own needs (2)	Focus on own needs (5)	
Engage in interaction (3)	Ask for attention (1)	No interaction (1)	
Do not understand (2)	Show appreciation (1)	Do not understand (1)	

The number of participants is in parentheses.

TABLE 5 | Breakdown of the most frequently communicated messages to the robot pet and the expected responses for each scenario.

Emotional state	Communicated message		Expected robot pet response
Stressed	Express emotional state (17)	\rightarrow	Keep company (6) Provide emotional support (6)
	Do not want to interact (6)	\rightarrow	Understand the situation (3) Focus on own needs (2)
Depressed	Seek emotional support (11)	\rightarrow	Provide emotional support (6) Keep company (3)
	Express emotional state (8)	\rightarrow	Provide emotional support (4) Keep company (3)
Relaxed	Enjoy company (14)	\rightarrow	Keep company (10)
	Express emotional state (8)	\rightarrow	Pick up the mood (5)
Excited	Express emotional state (15)	\rightarrow	Pick up the mood (15)
	Actively seek interaction (11)	\rightarrow	Pick up the mood (6)
			Engage in interaction (5)

The number of participants is in parentheses.

In contrast, some participants did not want to interact with the robot pet at all as they preferred to be alone in this situation or did not want to be distracted by the pet (n = 6). In response, most participants wanted that the robot pet showed its understanding of the situation by keeping its distance (n = 3). Others mentioned that the robot pet should have its own personality and should behave accordingly, which might result in the robot pet asking for attention even if this behavior is undesirable in this situation or that the pet would mind its own business (n = 2).

4.3.2. Depressed

In the depressed scenario, participants often communicated to the robot pet that they were looking for comfort in order to feel less depressed (n = 11). In response, these participants often wanted comfort from the robot pet (n = 6). They wanted the pet to do this by sitting on their lap or right next to them and making sounds. Also, participants specified that the robot pet should not approach them too enthusiastically. Others indicated that the robot pet should keep them company (n = 3) by staying close and showing its understanding of the situation.

Other participants just wanted to express how they felt (n = 8), for example, by telling the pet why they were feeling depressed. In response most of these participants also expected that the robot pet would either provide emotional support (n = 4) or would keep them company (n = 3).

4.3.3. Relaxed

In the relaxed scenario, participants often wanted to communicate that they enjoyed the pet's company (n = 14), for example, by having the pet sit on their lap or right next to them, touching the robot and talking to it. In response, these participants often wanted company from the robot pet (n = 10), for example, by staying close, listen, and engage in physical interaction. Furthermore, the pet's behavior should be calm and should reflect that it enjoys being together with the human (e.g., wagging tail or purring).

Other participants mentioned that they wanted to express that they were feeling relaxed (n = 8) such as by telling the pet about their day and that everything was alright. In response most of these participants wanted that the robot pet picked up on their mood by displaying relaxed behavior as well such as by lying down (n = 5).

4.3.4. Excited

In the excited scenario, participants often wanted to communicate their excitement to the robot pet (n = 15), for example, by touching and talking to the robot. In response, all these participants wanted that the robot pet picked up on their mood by becoming excited as well (n = 15). The robot pet could show its excitement by actively moving around, wagging its tail, and making positive sounds.

Other participants wanted to actively interact with the robot pet (n = 11) by playing with it or going out for a walk together. In response, most of these participants wanted that the robot pet picked up on their mood as well (n = 24) or preferred that the robot pet would actively engage them in play behavior (n = 6).

5. DISCUSSION

5.1. Categorization of Touch Behaviors

In this study, we observed participants that freely interacted with a robot pet companion. As a consequence, we observed an interesting but complex mix of touch behaviors such as the use of multiple touch gestures that were alternated, hybrid forms of prototypical touch gestures and combinations of simultaneously performed touch gestures (e.g., stroking while hugging). A previous attempt to annotate touch behaviors was limited to the coding of characteristics of touch gestures that were performed sequentially, which completely eliminates difficulties regarding segmentation and labeling that were encountered in this study (Yohanan and MacLean, 2012). Segmentation and labeling of individual touch gestures based a method borrowed from previous work on air gestures proved not to be straightforward. Although air gestures and touch gestures both rely on the same modality (i.e., movements of the hand(s)) their communicative functions are different. Air gestures, especially sign language, are a more explicit form of communication compared to communication through touch in which there is no one-to-one mapping between touch gestures and their meaning. Furthermore, in less controlled interactions it proved to be difficult to categorize touch behaviors into discrete touch gesture categories based on dictionary definitions, such as the gestures defined in Table 1. These results indicate that this approach might not be suitable to capture the nature of touch behavior in less controlled settings.

In accordance with previous findings from Yohanan and MacLean (2012) we frequently observed the use of massage, rub, scratch, stroke, and tickle-like gestures to communicate to the robot pet. As a result valuable information would be lost if these gestures would be collapsed into a single category to bypass the difficulties to clearly distinguish between these gestures. Some of the difficulties were due to the use of video footage to observe touch behavior. For example, the intensity level can only be roughly estimated from video [see also Yohanan and MacLean (2012)], and some details such as the precise point of contact were lost because of occlusion. However, confusions in identifying touch gestures with similar characteristics were also observed in studies where touch behaviors were captured by pressure sensors and algorithms were trained to automatically recognize different gestures (e.g., Silvera-Tawil et al., 2012; Jung et al., 2015, 2016). Moreover, segmentation and categorization of touch behavior based on touch sensor data would still remain challenging.

As the segmentation and categorization of touch behaviors into touch gestures might not be that straight forward in a less controlled setting, it might be more sensible to recognize and interpret social messages directly from touch sensor data as was previously suggested by Silvera-Tawil et al. (2014). Moreover, processing techniques from other modalities such as image processing, speech, and action recognition proved to be transferable to touch gesture recognition (Jung et al., 2015). Therefore, the existing body of literature on the transition toward automatic behavior analysis of these modalities in naturalistic settings might provide valuable insights for the understanding of touch behavior as well (e.g., Nicolaou et al., 2011; Gunes and Schuller, 2013; Kächele et al., 2016).

5.2. Observed Multimodal Behaviors

The following coarse descriptions of interactions with the robot pet from two different participants illustrate the use of multimodal cues in the depressed and excited scenario, respectively.

Participant walks into the living room and sits down on the couch next to the robot pet. Immediately she picks up the pet and holds it against her body using a hug-like gesture. While holding the pet she tells to the pet that she had a bad day while she makes eye contact from time to time. Then she sits quietly while still holding the pet and making eye contact. Finally, she puts the pet back on the couch and gets up to leave the room.

Participant runs into the living room and slides in front of the couch. He picks up the robot pet from the couch and then sits down on the couch with the pet resting on his leg. Then he talks to the pet using pet-directed speech: 'How are you? How are you? Yes! You're a good dog! Good doggy!'. Meanwhile he touches the pet using stroke-like gestures and looks at it. He then puts the robot pet back on the couch again while he still touches the pet using stroke-like gestures. Finally, he gets up from the couch and leaves the room.

As illustrated in the descriptions above, participants often talked to the robot pet while touching it (see also **Table 2**) indicating that the combination of speech (emotion) recognition and touch recognition might aid the understanding of touch behavior. Although we observed forms of speech that had characteristics of pet-directed speech (e.g., short sentences, repetition, and higher pitched voice), it should be noted that no analysis of the prosodic features of the speech was performed. However, the use of pet-directed speech has been observed previously, for example, Batliner et al. (2006) found that children used pet-directed speech when interacting with Sony's pet robot dog AIBO. A limitation of the current setup is that it did not allow for a detailed analysis on the added value of other social cues such as facial expression, posture, and gaze behavior for the interpretation of touch behavior.

By allowing the participants to freely interact with the robot pet within the confined space of a living room setting we were able to observe behavior that might otherwise not be observed. Social interaction involving objects such as taking pictures of the robot with a mobile phone were also observed by Cooney et al. (2015) who argued that these factors should be investigated to enable rich social interaction with robots. However, it is important to keep in mind that although participants in this study were able to freely interact within the given context, the results are confined to the given interaction scenarios. Furthermore, as the study relied on acted behaviors, participants might have displayed prototypical behaviors to clearly differentiate between the scenarios. However, although participants indicated in the questionnaire and during the interview that they had some difficulties acting out the scenarios with a stuffed animal, social behaviors such as making eye contact while talking (e.g., see descriptions above) were observed indicating that at least most participants treated the pet as a social agent. Additionally, it should be noted that touch was not only used to communicate to the robot pet but also often used to move/puppeteer the robot pet as it was unable to move on its own.

Surprisingly, interactions in the excited scenario were shorter despite the fact that all participants engaged in some form of interaction with the robot pet (see **Table 2**). A possible explanation is that participants often only quickly wanted to convey their excitement compared to other scenarios where they were seeking comfort or quietly sat down together with the robot pet to enjoy each others company (see **Table 3**). Furthermore, previous studies indicate that some emotions are more straightforward than others, for example, anger was found to be easier to express through touch than sadness (Hertenstein et al., 2009). Similarly, excitement might have been easier to convey than the other emotional states in this study.

5.3. Communicated Social Messages and Expected Robot Pet Responses

The interview results showed that the communicated messages and expected robot pet responses differed depending on the affective scenario and individual preference (see Tables 3 and 4). Moreover, Table 5 shows that there is no one-to-one relation between communicated messages and expected responses. For example, variation in expectations from the robot pet in the stressed scenario ranged from actively providing support to staying out of the way meaning that in order to respond in a socially appropriate manner, a robot pet should be able to judge whether the user wants to be left alone and when to engage in interaction. From the interviews it became clear that this is not always clear-cut, in the depressed and stressed scenarios some participants indicated that they did not want to initiate interaction but that they might be open to the robot pet approaching them (sometimes after a while). Participants often wanted to communicate their emotional state to the robot pet, especially in the high arousal scenarios (see Table 3); however, it should be noted that the focus on emotional states in the scenarios provided in the study might have biased participants toward expressing this emotional state.

Whether a robot pet should completely adapt its behavior to the user is dependent on the role of the pet. In this study the nature of the bond between the participant and his/her robot pet was not specified. Some participants argued that a robot pet should mimic a real pet with its own personality and needs, which might conflict with the current needs of the user. In contrast, other participants proposed that the robot pet could take the role of therapist/coach, which would focus on the user's needs. Mentioned abilities that such a robot pet should have included cheering you up, providing comfort, talking about feelings, and communicating motivational messages. In the role of a friend the robot pet should also take the user's needs into account, albeit to a lesser extent.

In this study we observed how various people, in this case males and females from the working-age population, interacted with a robot pet companion. However, it should be noted that individual factors such as previous experience with animals, personality, gender, age, and nationality might play an important role in these interactions. Interestingly, even though the robot pet's embodiment clearly resembled a dog, some participants treated the robot pet as a cat. Whether participants treated the robot pet as a dog or a cat seemed to depend on their preference and history with real pets. Additionally, it should be noted that the participants studied or worked in the computer science department and that all were at least to some extent familiar with social robots. As a result some participants took the current state of technology into account when suggesting possible robot behaviors, for example, one participant mentioned that it is non-trivial to build a robot dog that would be able to jump on the couch. The use of a stuffed animal dog as a proxy for a functioning robot pet allowed for a more controlled setup. However, the lack of response from the robot pet resulted in less realistic interactions as the participant had to puppeteer the pet or imagine its response. Furthermore, it is important to note that participants were asked to act as if they were coming home in a particular emotional state. Although this is a common approach in studies on touch behavior (e.g., Hertenstein et al., 2006, 2009; Yohanan and MacLean, 2012; Silvera-Tawil et al., 2014), it is unclear whether the same results would have been found if the emotional states were induced in the participants. Despite the above mentioned considerations we observed an interesting range of interactions and were able to find patterns in the social messages that were communicated and the responses that were expected from the robot pet.

6. CONCLUSION

To gain more insight into the factors that are relevant to interpret touch within a social context we studied interactions between humans and a robot pet companion in different affective scenarios. The study took place in a contextualized lab setting in which participants acted as if they were coming home in different emotional states (i.e., stressed, depressed, relaxed, and excited) without being given specific instructions on the kinds of behaviors that they should display.

Results showed that depending on the emotional state of the user, different social messages were communicated to the robot pet such as expressing one's emotional state, seeking emotional support, or enjoying the pet's company. The expected response from the robot pet to these social messages also varied based on the emotional state. Examples of expected responses were keeping the user company, providing emotional support, or picking up on the user's mood. Additionally, the expected response from the robot pet was dependent on the different roles that were envisioned such as a robot that mimics a real pet with its own personality or a robot companion that serves as a therapist/coach offering emotional support.

Findings from the video observations showed the use of multimodal cues to communicate with the robot pet. Participants often talked to the robot pet while touching it and making eye contact confirming previous findings on the importance of studying touch in multimodal interaction. Segmentation and labeling of touch gestures proved to be difficult due to the complexity of the observed interactions. The findings of this study indicate that the categorization of touch behaviors into discrete touch gesture categories based on dictionary definitions is not a suitable approach to capture the nature of touch behavior in less controlled settings. Additional research will be necessary to determine whether direct recognition and interpretation of higher level social messages from touch sensor data would be a viable option in less controlled situations. Moreover, as the current results are based on acted scenarios, it is important to verify in future research whether similar behaviors occur in a naturalistic setting in which people would interact with a fully functioning robot pet in their own home. A first step could be to induce emotions in participants and observe their interactions with a responding robot pet in a lab setting. The use of verbal behavior that coincides with touch interaction seems another interesting direction for future studies into the automatic understanding of social touch.

REFERENCES

- Altun, K., and MacLean, K. E. (2015). Recognizing affect in human touch of a robot. Pattern Recognit. Lett. 66, 31–40. doi:10.1016/j.patrec.2014.10.016
- Banks, M. R., Willoughby, L. M., and Banks, W. A. (2008). Animal-assisted therapy and loneliness in nursing homes: use of robotic versus living dogs. J. Am. Med. Dir. Assoc. 9, 173–177. doi:10.1016/j.jamda.2007.11.007
- Batliner, A., Biersack, S., and Steidl, S. (2006). "The prosody of pet robot directed speech: evidence from children," in *Proceedings of Speech Prosody* (Dresden, Germany), 1–4.
- Cooney, M. D., Nishio, S., and Ishiguro, H. (2012). "Recognizing affection for a touch-based interaction with a humanoid robot," in *Proceedings of the International Conference on Intelligent Robots and Systems (IROS)* (Vilamoura-Algarve, Portugal), 1420–1427.
- Cooney, M. D., Nishio, S., and Ishiguro, H. (2015). Importance of touch for conveying affection in a multimodal interaction with a small humanoid robot. *Int. J. Humanoid Robot.* 12, 1550002. doi:10.1142/S0219843615500024
- Eachus, P. (2001). Pets, people and robots: the role of companion animals and robopets in the promotion of health and well-being. *Int. J. Health Promot. Educ.* 39, 7–13. doi:10.1080/14635240.2001.10806140
- Gallace, A., and Spence, C. (2010). The science of interpersonal touch: an overview. Neurosci. Biobehav. Rev. 34, 246–259. doi:10.1016/j.neubiorev.2008.10.004
- Gunes, H., and Schuller, B. (2013). Categorical and dimensional affect analysis in continuous input: current trends and future directions. *Image Vis. Comput.* 31, 120–136. doi:10.1016/j.imavis.2012.06.016
- Hertenstein, M. J., Holmes, R., McCullough, M., and Keltner, D. (2009). The communication of emotion via touch. *Emotion* 9, 566–573. doi:10.1037/a0016108
- Hertenstein, M. J., Keltner, D., App, B., Bulleit, B. A., and Jaskolka, A. R. (2006). Touch communicates distinct emotions. *Emotion* 6, 528–533. doi:10.1037/1528-3542.6.3.528
- Heslin, R., Nguyen, T. D., and Nguyen, M. L. (1983). Meaning of touch: the case of touch from a stranger or same sex person. J. Nonverbal Behav. 7, 147–157. doi:10.1007/BF00986945
- Jones, S. E., and Yarbrough, A. E. (1985). A naturalistic study of the meanings of touch. Commun. Monogr. 52, 19–56. doi:10.1080/03637758509376094
- Jung, M. M., Cang, X. L., Poel, M., and MacLean, K. E. (2015). "Touch challenge '15: recognizing social touch gestures," in *Proceedings of the International Conference on Multimodal Interaction (ICMI)* (Seattle, WA), 387–390.
- Jung, M. M., Poel, M., Poppe, R., and Heylen, D. K. J. (2016). Automatic recognition of touch gestures in the corpus of social touch. J. Multimodal User Interfaces. 11, 81–96. doi:10.1007/s12193-016-0232-9
- Kächele, M., Schels, M., Meudt, S., Palm, G., and Schwenker, F. (2016). Revisiting the EmotiW challenge: how wild is it really? J. Multimodal User Interfaces 10, 151–162. doi:10.1007/s12193-015-0202-7
- Kendon, A. (1980). Gesticulation and speech: two aspects of the process of utterance. Relat. Verbal Nonverbal Commun. 25, 207–227.
- Kim, Y.-M., Koo, S.-Y., Lim, J. G., and Kwon, D.-S. (2010). A robust online touch pattern recognition for dynamic human-robot interaction. *Trans. Consum. Electron.* 56, 1979–1987. doi:10.1109/TCE.2010.5606355

AUTHOR CONTRIBUTIONS

MJ designed the study with assistance from MP and DH; collected the data and analyzed the results with assistance from MP, DR, and DH; and wrote the manuscript with contributions from MP, DR, and DH. All the authors reviewed and approved the manuscript.

ACKNOWLEDGMENTS

The authors would like to acknowledge Josca Snippe for participating in the annotation process. This publication was supported by the Dutch national program COMMIT.

- Kita, S., van Gijn, I., and van der Hulst, H. (1997). "Movement phases in signs and co-speech gestures, and their transcription by human coders," in *International Gesture Workshop* (Bielefeld, Germany), 23–35.
- Knight, H., Toscano, R., Stiehl, W. D., Chang, A., Wang, Y., and Breazeal, C. (2009). "Real-time social touch gesture recognition for sensate robots," in *Proceedings* of the International Conference on Intelligent Robots and Systems (IROS) (St. Louis, MO), 3715–3720.
- Nicolaou, M. A., Gunes, H., and Pantic, M. (2011). Continuous prediction of spontaneous affect from multiple cues and modalities in valence-arousal space. *Trans. Affect. Comput.* 2, 92–105. doi:10.1109/T-AFFC.2011.9
- Noda, T., Ishiguro, H., Miyashita, T., and Hagita, N. (2007). "Map acquisition and classification of haptic interaction using cross correlation between distributed tactile sensors on the whole body surface," in *International Conference on Intelligent Robots and Systems (IROS)* (San Diego, CA), 1099–1105.
- Russell, J. A., Weiss, A., and Mendelsohn, G. A. (1989). Affect grid: a single-item scale of pleasure and arousal. J. Pers. Soc. Psychol. 57, 493–502. doi:10.1037/0022-3514.57.3.493
- Silvera-Tawil, D., Rye, D., and Velonaki, M. (2012). Interpretation of the modality of touch on an artificial arm covered with an EIT-based sensitive skin. *Robot. Res.* 31, 1627–1641. doi:10.1177/0278364912455441
- Silvera-Tawil, D., Rye, D., and Velonaki, M. (2014). Interpretation of social touch on an artificial arm covered with an EIT-based sensitive skin. *Int. J. Soc. Robot.* 6, 489–505. doi:10.1007/s12369-013-0223-x
- Suvilehto, J. T., Glerean, E., Dunbar, R. I., Hari, R., and Nummenmaa, L. (2015). Topography of social touching depends on emotional bonds between humans. *Proc. Natl. Acad. Sci. U.S.A.* 112, 13811–13816. doi:10.1073/pnas. 1519231112
- Veevers, J. E. (1985). The social meaning of pets: alternative roles for companion animals. *Marriage Fam. Rev.* 8, 11–30. doi:10.1300/J002v08n03_03
- Wada, K., and Shibata, T. (2007). Living with seal robots its sociopsychological and physiological influences on the elderly at a care house. *Trans. Robot.* 23, 972–980. doi:10.1109/TRO.2007.906261
- Yohanan, S., and MacLean, K. E. (2012). The role of affective touch in human-robot interaction: human intent and expectations in touching the haptic creature. *Int. J. Soc. Robot.* 4, 163–180. doi:10.1007/s12369-011-0126-7
- Zasloff, R. L. (1996). Measuring attachment to companion animals: a dog is not a cat is not a bird. *Appl. Anim. Behav. Sci.* 47, 43–48. doi:10.1016/0168-1591(95)01009-2

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2017 Jung, Poel, Reidsma and Heylen. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.