International Journal of Humanoid Robotics

PREPRINT OF AN ARTICLE PUBLISHED IN INTERNATIONAL JOURNAL OF HUMANOID ROBOTICS (IJHR) 12(1), 1550002 1-22. DOI: 10.1142/S0219843615500024 C [COPYRIGHT WORLD SCIENTIFIC PUBLISHING COMPANY] HTTP://WWW.WORLDSCIENTIFIC.COM/DOI/PDF/10.1142/S0219843615500024?SRC=RECSYS

IMPORTANCE OF TOUCH FOR CONVEYING AFFECTION IN A MULTIMODAL INTERACTION WITH A SMALL HUMANOID ROBOT

MARTIN D. COONEY[*, †, ‡], SHUICHI NISHIO[*, ¶], and HIROSHI ISHIGURO[*, †, §]

**Hiroshi Ishiguro Lab, Advanced Telecommunications Research Institute International (ATR), 2-2-2 Hikaridai, Keihanna Science City, Kyoto 619-0288, Japan*

†Dept. of Systems Innovation, Graduate School of Engineering Science, Osaka University, 1-3 Machikaneyama, Toyonaka, Osaka 560-8531, Japan,

‡mcooney@atr.jp
¶nishio@botransfer.org
§ishiguro@atr.jp

Received 31 July 2013
Revised 26 December 2013, 8 April 2014
Accepted Day Month Day

To be accepted as a part of our everyday lives, companion robots will require the capability to communicate socially, recognizing people's behavior and responding appropriately. In particular, we hypothesized that a humanoid robot especially should be able to recognize affectionate touches conveying liking or dislike, because the humanoid form elicits expectations of a high degree of social intelligence, providing affection can contribute to people's quality of life, and people will seek to show affection by touching a robot because such behavior is fundamental and crucial for human bonding. This hypothesis needed to be verified because robots are typically not soft or warm like humans, and people can communicate through various other modalities such as vision and sound. The main challenge faced was that people's social norms are highly complex, involving behavior in multiple channels. To deal with this challenge, we adopted an approach in which we analyzed free interactions and also asked participants to rate short video-clips depicting human-robot interaction. As a result, we verified that touch plays an important part in the communication of affection from a person to a humanoid robot considered capable of recognizing cues in touch, vision, and sound. Our results suggest that designers of affectionate interactions with a humanoid robot should not ignore the fundamental modality of touch.

Keywords: Affection; humanoid robot; touch; multimodal; human-robot interaction.

1. Introduction

This article concerns itself with investigating the role of touch in people's affectionate behavior directed toward a humanoid robot in a multimodal interaction. It extends our previous work in regard to touch[11] (and proprioception)[12] to consider the much broader case in which people are free to interact as they wish.

We use the term "behavior" to refer to both intended and unintended signals,[1] and describe behavior in terms of specific "actions", general "channels", and overall "modalities" (e.g., a nodding action involves the channel of head movement, which is typically perceived via the modality of vision). We also seek to associate with each behavior a fundamental meaning indicating the degree to which it is typically perceived as conveying an "affectionate" attitude of like or dislike. The term, "affection", is used in its usual sense[2-5] to connote liking, loving, positivity, and gentleness; it does not signify affect or emotion in general, but rather the important quality of valence[6] used to interrelate emotions[7] and gauge the perceived impact of conflicting signals.[8]

Motivation for our investigation was five-fold. (1) Humanoid robots are expected to possess an advanced capability to recognize various human behavior.[10] (2) Such social intelligence will facilitate enjoyment and intent to use,[13] and avoid some demerits of simple alternative approaches to realizing interactions. For example, proactive scripts, direct imitation, arbitrary recognition, and tele-operation could respectively reduce enjoyment by undermining a person's control over the course of an interaction, [14] fail when a robot should not or cannot reciprocate a person's behavior, target behavior which is never encountered, or not be practical for all companion robots in terms of cost. For this a specific task is not required; playing, which is its own reward, can provide enjoyment,[16] which is a requirement for technological acceptance.[13, 17-18] (3) Recognizing and providing affection in particular will contribute to people's quality of life. Humans feel a pressing need for affection,[4] which unmet in many individuals[3] has been linked to deleterious health outcomes;[21-22] Several pioneering studies have already shown that affectionate interactions with a robot can be therapeutic.[19-20] We imagine that a great opportunity exists within this context for robots to contribute to improving quality of life in lonely individuals. (4) People will seek to show affection by touching a robot because such behavior is fundamental and crucial for human bonding.[2, 15] As well, spontaneous exhibitions of affection have been observed also toward robots.[5, 23] (5) In our previous study we investigated how people touch a humanoid robot when forced to communicate haptically, but we did not know what would happen in the unconstrained case.

Thus our goal was to investigate the role of touch in conveying affection to a humanoid robot. The challenge was that it is difficult to model people's perceptions because human social norms are complex[24] and non-verbal behavior is highly varied, multi-channeled, and polysemic.[1, 9] Addressing not only such "extremely difficult" perceptual aspects of behavior, but furthermore the reasons which motivate, which we seek to begin to do, involves a foray into a "virtually unexplored area of research".[25] Also, costs in time and effort of closely analyzing behavior are not negligible.

To address the challenge, we employed a combined approach in which we first observed free interactions[26] with a small humanoid robot and then asked participants to rate "thin-slices"[27] of videotaped interaction. Observation was required to gain overall insight into this complex phenomenon; we coded *typical* behavior to deal with the high variance, used a bag-of-words approach in our initial analysis of multiple channel data, and associated behavior with typical reasons to gain insight into motivation. These results were refined by testing the perceived effects of mixed signals via short video-clips depicting interactive behavior. The reported knowledge is relevant for preparing robots for engaging in affectionate interactions.

The rest of this article is structured as follows. Section 2 discusses some related work. Section 3 describes a new platform we built, Penumbra. Section 4 explains methodology we applied in our investigation; results are then reported in Section 5 for questionnaires, observed interactions, and "thin-sliced" video-clip evaluations. Section 6 summarizes contributions and presents discussion.

2. Related Work

A number of previous studies have investigated touch, vision, and sound in human-robot and human-human interaction. In particular, how people seek to communicate affection through touch toward a robot has been well studied. Noda et al. first reported using an observational approach to identify play categories.[28] Knight and her colleagues also observed playful interactions to identify categories, gestures and sub-gestures.[29] Stiehl et al. first described an approach toward recognizing the affectionate meanings of touches based on softness and type of touch.[30] François, Polani, and Dautenhahn also described an approach to distinguish strong and gentle forms of playing in order to adapt a robot's behavior.[31] Yohanan and MacLean investigated how people touched a robot to communicate various emotions.[32] Our previous studies on this topic reported first some playful proprioceptive actions,[12] and then some typical touches toward a humanoid robot and their affectionate meanings in a more generic framework.[11] What was not investigated was the case in which people are free to communicate not only through touch but also through visual and aural means.

Some other studies have partially explored audiovisual behavior directed toward a robot. A pioneering work by Breazeal and Aryananda reported that people used four typical kinds of child-directed speech toward a humanoid robot,[33] including affectionate approval and comforting. Broekens predicted that a person's smiling would be useful for a robot to recognize as an affectionate reinforcement signal.[39] Some other studies reported differences and similarities between human-robot and human-human interactions: gaze behavior differed, [34] whereas preferred interaction distances and perception of gestures were similar.[35, 36] These studies revealed some important findings with regard to people's perceptions toward a robot, but did not clarify how people typically communicate affection to a robot.

Some insight can be inferred from reports of human-human interactions. Affectionate behavior between humans includes mutual and frequent gaze,[40] close distance,[41] forward-inclined[42] open posture,[43] physical warmth,[44] smiling,[45] nodding,[9] gestures such as waving,[46] courtship signals such as hair-stroking,[47] pupil dilation,[48] and mimicking (also known as the chameleon effect or accommodation).[49] Likewise, reasons for interacting via touch,[50] vision[51-52] and sound[53] have been identified, which could also apply to robots. The degree to which affectionate attitude is conveyed by some conflicting signals has been suggested by Mehrabian and colleagues, who studied the interplay of vocal, verbal, and facial signals;[8, 54] also, Gada et al. found that thin slices of observed behavior sufficed to accurately predict the degree of affection felt between interacting persons.[55] Such results provide a basis for forming some predictions but do not address the context of human-robot interaction.

Thus, the main novel contribution of the current work is investigating the role of touch for conveying affection in a multi-modal interaction with a humanoid robot. The acquired knowledge will help inform the development of recognition capability for engaging in affectionate interactions, toward facilitating successful integration of companion robots into everyday human environments and contributing to people's quality of life.

3. Our Robot, Penumbra

To investigate affectionate human-robot interaction, we desired a platform with a "robot-like" and interaction-friendly embodiment (safe and light), and typical sensory-behavioral capabilities. Since such a robot was not available to us, we built a new robot prototype, Penumbra, shown in Fig. 1. Penumbra was given a mechanical appearance with typical features for a humanoid robot such as a head with a face, arms, and wheels. To invite interaction, sensing capability was implied by attaching reflective eyes, ear-like surfaces, and touch pads; to further entice touching and holding, Penumbra was fashioned to be light (0.91kg). In order to be able to interact with either a sitting or a standing person, Penumbra was designed to be a short robot which moves on a tall stage; it was small enough to be held, and tall when standing on its stage. To ensure safety, Penumbra was set to move along a track, its motions were designed to be moderate in speed, and low-torque actuators were used which can be easily halted without hurting a human.

Penumbra was also designed to be capable of recognizing some typical signals conveyed by touch, vision and sound and performing some typical behavior for a humanoid robot: speech, gaze, locomotion, nodding, head-shaking, and gesturing. A simple strategy was employed in which Penumbra was made to respond in kind to detected behavior when possible. Thus, the robot was set to look toward a person when a frontally-oriented face was detected and away otherwise, a person's motion beside the robot triggered forward movement, and, if the robot detected speech followed by a brief pause, it responded by playing back a generic utterance in a flat "robotic" voice. Because imitating touches is a difficult open challenge, Penumbra simply acknowledged touches via speech by stating where it was touched and if it had been picked up or not.

To realize the design, a program running on an external computer was configured to receive and process sensor data, and send commands to the robot via Bluetooth. We used an off-the-shelf wireless camera with an in-built microphone to acquire audiovisual data. Empirically determined thresholds were applied to sound intensity via the library "FMOD" to determine if speech was occurring at each time step; additional thresholds ensured that speech persisted for a few seconds preceding a short pause. To capture participants' facial and body information, the camera was tilted upward. The Open CV computer vision library was then used to detect faces in the robot's camera view and motion in the area in front of the robot; thresholds on detection results per time step and over a window of data were again used to reduce false positives. Touch sensing was realized by building and attaching seven capacitive sensors to the robot's hands, cheeks,

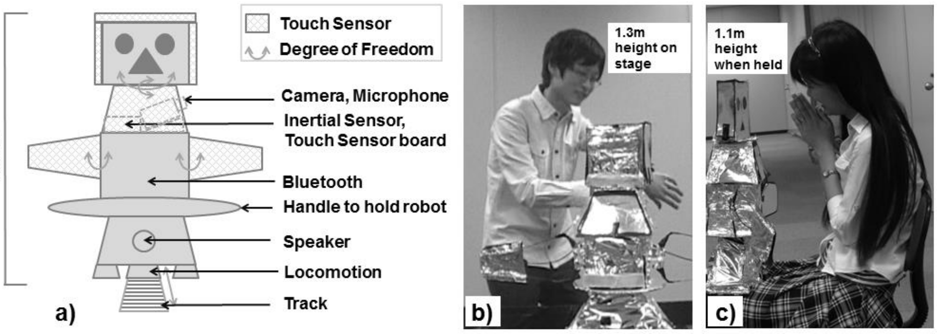

Fig. 1. Penumbra: a) hardware and b-c) interaction scenario involving standing or sitting.

head-top, back, and front. Picking up the robot was detected by thresholding raw z-axis values from an accelerometer.

Motion capability for Penumbra's head and arms, involving flexion/extension and rotation, and flexion/extension respectively, was realized using four light-weight servo actuators. Locomotion was triggered by controlling a gear-motor attached to toy train wheels. Speech playback was enabled by triggering pre-recorded audio. We expected that the robot's appearance and capabilities, spanning three modalities (touch, vision, and sound), would convince participants the robot could recognize a range of haptic, visual, and aural signals, which was verified in the experiment described next.

4. Investigating People's Affectionate Multimodal Behavior

To investigate the role of touch in conveying affection to a humanoid robot in which people are free to employ any means of communication they desire, new data were obtained from 20 young adult Japanese participants (9 females and 11 males; average age=20.4 years, SD=1.7 years). The ethics committee at ATR approved the experiment (13-601-04), and participants gave informed consent. This section describes our approach, and Section 5 presents results.

4.1. *General scenario, goals, and approach*

To enable comparison with our previous study[11], we again restricted our investigation to a simplified dyadic scenario in a controlled environment in which a single human and robot interacted without objects. A young adult Japanese interacted for various typical reasons such as greeting or thanking (discussed in Section 4.3.1) with the robot, which was capable of recognizing multimodal signals.

Within this scenario our goal was to determine the importance of touch for conveying affection, but it was not self-evident how to define "important". We predicted (1) that frequently observed cues will be important to recognize because they can be expected to offer opportunities to interact; conversely, behavior which is never observed does not need to be recognized. (2) Likewise, highly affectionate and unaffectionate behavior will be useful to recognize because occasions when a person is happy or angry with a robot constitute opportunities for bonding or avoiding dissatisfaction which should not be ignored. (3) Furthermore, behavior which is "monosemic", or highly associated with a single motivating reason, is important because estimating intentions could help a robot to behave appropriately and predict a person's future behavior. (4) Also, signals which "dominate", deciding how much affection will be perceived when multiple signals co-occur, will be important because signals which have little effect do not need to be recognized. For example, hugging a robot could convey much affection regardless of whether a person is smiling or crying. We emphasize that using these four descriptors—frequency, degree of affection, monosemy, and dominance—to evaluate importance is our own idea.

To assess these factors, our approach combined four methods: compiling knowledge from the literature, observing people's behavior then asking how much affection should be perceived, acquiring subjective measurements of the degree of affection conveyed in short video clips, and questionnaires to provide additional checks. All methods involved accepted methodology. Both the observational and histological approaches involve standard methodology[26, 27] which has been applied previously in robotics.[29, 32; 38] In our case, both these approaches were required because without observation, we could not gain insight into the basic structure of affectionate behavior; without the latter we could not systematically compare effects of conflicting signals.

4.2. *Procedure*

Experiment sessions lasted approximately one hour, comprising initial instructions (~five minutes), an introduction of our robot (~five minutes), free interactions (~30 minutes), observing videos and completing questionnaires (~15 minutes), and short interviews (~five minutes). First, participants entered a large well-lit room (~15 by 8 meters; ~500 lux brightness) after which the door was closed for privacy. Seated at a desk, participants read a short handout which indicated that they would interact for various reasons with a robot capable of moving and uttering speech; and that they would be free to touch, speak, move around the room, and behave in any fashion they desired. The instructions indicated that they would be asked to describe how much affection was communicated by their behavior to the robot, defined as an attitude of like/dislike. Also, a list of typical reasons for interacting such as greeting or thanking the robot was presented (described in Section 4.3.1).

After having a chance to ask questions, our robot was introduced. Participants moved in front of Penumbra, which greeted them, saying, "Hello, I'm looking forward to working with you" while looking toward them and nodding. Next, Penumbra's capabilities were introduced one by one; this provided an opportunity to demonstrate that the robot could perceive cues in various modalities, and also to help elicit more natural behavior by reducing novelty effects or possible anxiety. Participants were asked to freely touch and pick up the robot, which stated which parts of its body were touched, and when lifted said, "Hey, please don't drop me" and, "I'm scared". Penumbra's visual capabilities were demonstrated by looking toward detected faces and moving forward when a participant moved. Penumbra also responded with generic utterances such as "Understood" and "That's right" when a person spoke and paused. The robot spoke and moved in all conditions when introduced.

Next, participants interacted freely for the typical reasons listed on the handout. Each reason was first read aloud by the experimenter. After having a chance to think, participants performed some behavior, touching, speaking, or moving as they wished. They were then asked by the experimenter to judge how much affection their behavior communicated, and to describe what they had done. Graphic user interface (GUI) programs were used by the experimenter to control Penumbra, record sensor data for future use, and display reasons for interacting (loaded from files in which they had been generated for each participant in random order).

After interacting, participants sat in front of a computer screen and watched short video clips, which they were asked to evaluate. At the end, participants filled out questionnaires and short interviews were conducted.

4.3. *Materials*

As described above, the experimenter read aloud typical reasons for interacting, and participants watched short video clips. The list of reasons and sequence of videos were prepared as follows.

4.3.1. *Reasons for interaction and typical channels*

In our earlier work focused on touch,[10] participants indicated a desire to know why they were interacting, for which we compiled a list of reasons for touching. To ease their task in the current work as well, we again compiled a list, which involved combining reasons for touching obtained previously[50, 56] with reasons for visual and aural expression. Showing a referent was added as a reason for deictic gestures;[51] as well as causing another to continue or stop talking, and expressing power.[52] For sound, the following reasons were added: protest (no!), answer (state, describe, or persuade), request something, and attract the robot's attention,[53] as well as approval (yes).[33] Some instructions were reworded to avoid specifying a modality which the participant would have to use: e.g., "Get the robot to continue" was used in place of "Get another to continue talking". "Protest to the robot" was also combined with "Get the robot to stop" as "Stop". This resulted in 22 reasons, shown in Table 1.

For channels, we decided to test the perceived degree of affection conveyed by gaze, proxemics, posture, head, facial, hand, touch, and vocal actions, based on the sources listed in Section 2. Courtship signals were considered unlikely to be performed toward a robot, warmth and pupil dilation could not be clearly shown, mimicking is not linked to any one channel, and the verbal channel is unique in its variety and could not be tested in the scope of the current work.

Table 1. Reasons for interacting with a humanoid robot.

Reason	Example (Verbal)	Reason	Example (Verbal)
Reassure[53, 58]	"Poor robot, it's okay"	Anger[58]	"I hate you"
Thank[53]	"Thanks!"	Disgust[58]	"Ew, gross"
Accept[53]	"We're friends"	Fear[58]	"I'm scared"
Attraction[53]	"How handsome/cute!"	Stop[33, 50, 51]	"No! Stop!"
Show Love[53,58]	"I love you"	Respond[50]	"Okay, let me tell you…"
Tired (Seek Reassurance)[53]	"I'm tired, my head/belly/throat hurts"	Request[50]	"Could you …?"
Sad (Seek Love)[53]	"I'm lonely/sad"	Attention[50]	"Hey!"
Play[53]	"Let's play! Hey!"	Approve[33]	"Yes!"
Move[53a]	"Move your body/arm/head"	Show[52]	"Look!"
Greet[53]	"Hello/Goodbye"	Continue[51]	"Go on"
Inspect[53a]	"Inspect the robot"	Dominate[51]	"I'm in command"

[a]Interpreted for the context of behavior directed toward a humanoid robot.

4.3.2. *Videos*

To create videos showing conflicting signals, we recorded a confederate pretending to interact with our robot. We found that our confederate typically made direct eye contact at her preferred distance with an open posture; this behavior served as the basis for each video, which was varied based on what modality was shown. For example, when smiling or frowning at the robot, the confederate did so while making direct eye contact at her preferred distance with an open posture.

It was deemed impractical to test all combinations of positive and negative examples for all channels of interest, which would require many videos; we decided to test each modality of touch, vision, and sound, using only six videos. Each video shows two modalities together: one positive and one negative (e.g., positive touch and negative vision). Three different actions were used to represent each modality, to avoid confusion associated with any one example. These actions were chosen based on some pre-trials to find signals with a strong affectionate meaning (positive or negative) and are listed below:

- Touch+: the confederate hugged the robot, rubbed its head, and stroked its cheek.
- Touch-: the confederate slapped the robot, pushed its chest, and covered its face.
- Vision+: the confederate smiled, nodded, and waved toward the robot.
- Vision-: the confederate scowled, shook her head, and shook her finger at the robot.
- Sound+: the confederate said, "I love you", "You're great", and, "It's okay".
- Sound-: the confederate said, "I hate you", "You're the worst", and, "Bad!".

4.4. *Conditions*

Video clips and reasons for interacting were presented in random order. Two factors were used for free interactions, human posture and robot behavior: a participant sat or stood, and Penumbra did, or did not, move and speak. In the behavior condition, the robot played back an utterance and performed one of three motions (a positive nod, a neutral motion swinging its arms, and a negative head-shake). Neutral utterances were selected: “Yeah”, “Really?”, “Is that true?”, “Understood”, and “Of course”.

4.5. *Measurements*

During the free interactions, behavior was videotaped and recorded by an experimenter in the form of written notes; when unclear, participants were asked to describe their behavior. Involved analysis of highly subtle movements (e.g., as in the Facial Action Coding System (FACS) for facial expressions, which involves agreement by certified judges[57]) was outside of the scope of the current work; therefore only large pronounced movements and those described by participants themselves were noted. Participants were also asked after each time they performed a behavior toward the robot for a typical reason to rate their own behavior on a seven-point scale (one: conveying an attitude of dislike, four: neither, seven: conveying liking).

After the free interactions, four questionnaires were also completed, to assess our robot (one page), likelihoods for each reason (one page), and degrees of affection transmitted in the thin-sliced videos (one page) and by typical behavior channels in the participants' opinions (one page). Participants' perceptions of the robot's capabilities were evaluated to check that the system had successfully presented an impression that the robot could perceive behavior in various channels; this was desired because a poor opinion of the robot's perceptual capabilities could alter people's behavior. Using a five-point scale (one: I disagree completely, two: I disagree, three: neither, four: I agree, five: I agree strongly), the following four statements were rated: “The robot can perceive visual information”, “The robot can perceive aural information”, “The robot can perceive touch information”, and “The robot can perceive various kinds of information”. To assess the likelihood of interacting for each typical reason, a seven-point scale was selected, in the form, “I would interact with a robot for this reason” (one: completely unlikely; two: somewhat unlikely, three: slightly unlikely, four: neither, five: slightly likely, six: somewhat likely, seven: very likely). Next, participants played back the videos described in Section 4.3.2, and evaluated affectionate attitudes conveyed using a seven-point scale. Participants were also asked to describe the content of each clip to encourage them to pay attention; however these responses were not analyzed further.

4.6. *Predictions*

Although we wished to keep an open mind, we had some expectations.

- P1: Our setup would be considered appropriate: (a) Penumbra would be perceived as capable of recognizing various signals communicated via touch, vision and sound, (b) our list of reasons for interacting would be considered likely, and (c) channels we selected would be considered important for conveying affection.

- P2: Touches would play a fundamental role in communicating affection even toward a robot capable of recognizing visual and aural signals: (a) touches would be observed, (b) some of these would be perceived as communicating high degrees of affection, (c) some would be useful for predicting an affectionate reason for interacting, and (d) touch would have a strong influence on how conflicting signals are perceived.

The reason for P1 was our expectation that a person's behavior will change if they think their actions will be recognized or not; for example, a person might say goodbye instead of waving when speaking over the phone, or conversely wave instead of speaking if their interlocutor is far away or behind a glass door. P2 stemmed from the importance of touch for communicating affection in human-human and animal interactions.

These predictions required verification. P1 was checked to determine if there was a problem in our robot or instructions which could prevent participants from behaving freely and which might invalidate our results. P2 was checked to gain insight into what affectionate behavior could be important for a robot to recognize.

5. Results and Analysis

To verify our predictions, we checked questionnaire results then analyzed free interactions and thin-sliced video evaluations.

5.1. *Perception of our robot's capabilities, reasons for interacting, and channels*

To check that some important parts of our set-up would be perceived as expected, questionnaire scores were analyzed. Most participants perceived our robot as being capable of perceiving visual information (average score: 4.1 ± 0.51), aural information (4.2 ± 0.62), touch information (4.6 ± 0.60), and various information (3.9 ± 0.85) based on a five-point scale with four being "I agree". A one-tailed t-test indicated a significant difference for all items from a neutral opinion: $t(19) = 9.20$, $p < .001$; $t(19) = 8.72$, $p < .001$; $t(19) = 11.46$, $p < .001$; and $t(19) = 4.72$, $p < .001$. Some participants who provided low scores explained that it would not be possible to recognize everything and that they did not believe our robot could, for example, understand a news program. Thus, our Prediction P1a from Section 4.6 was supported, suggesting that the system in general successfully conveyed an impression which allowed participants to interact freely.

For the reasons for interacting described in Section 4.3.1, a one-tailed t-test indicated that participants perceived the reasons overall as likely compared to a neutral response (likelihood: 4.6 ± 0.94; $t(439) = 7.36$, $p < .001$). A repeated measures ANOVA with a Greenhouse-Geisser correction revealed that participants considered some reasons likelier than others: $F(7.61, 144.57) = 8.82$, $p < .001$, $\eta^2 = .32$. Post hoc tests using the Bonferroni correction revealed significant differences in the likelihood of seeking to play (5.9 ± 0.94) compared with showing fear (p = .001) or approval (p = .002), and seeking attention (p = .015), or reassuring (p = .034), with a trend toward a significant difference for causing a robot to move (p = .087); thanking (5.6 ± 1.2) was more likely than reassuring (p = .014) and seeking love (5.5 ± 1.2) was more likely than approving (p = .040). By contrast, Disgust (2.1 ± 1.4) was significantly less likely than all other reasons except showing love, anger, fear, and approving; anger (2.8 ± 2.0) was significantly less likely for all but these reasons and seek love, move, disgust, attention, show, continue, and dominate. Thus, Prediction P1b, that people would consider the compiled reasons for interacting as likely, was mostly supported.

For the channels we thought would be important for conveying affection described in Section 4.3.1, a repeated measures ANOVA with a Greenhouse-Geisser correction showed a significant difference in participants' perceptions: $F(3.79, 71.97) = 4.29$, $p = .004$, $\eta^2 = .18$. Post hoc tests using the Bonferroni correction revealed that touch (6.5 ± 0.69) was considered significantly more important for conveying affection than distancing (5.5 ± 0.97, p = .037), posture (5.5 ± 1.2, p = .019), and arm gestures (5.4 ± 1.0, p = .001), with no other differences arising. However, a one-tailed t-test indicated that all channels were judged by participants to be at least somewhat important for conveying affection to a robot (lowest value for gestures: $t(19) = 4.78$, $p < .001$; average of all scores: 5.8 ± 1.0). Therefore, Prediction P1c, that people would consider the compiled channels as important for conveying affection, was supported; participants also thought touches would be significantly more important in conveying affection than some other channels, in accordance with Prediction 2.

5.2. *Affectionate behavior in free interactions*

Our coding process followed the basic spirit of the grounded theory method which develops a model from data.[58] During sessions, the experimenter asked participants to describe their behavior. Afterwards, notes were inputted into database tables for each channel and the video footage was reviewed to ensure data completeness and accuracy. Next, notes were replaced by "codes" such as "hug". We found that head and body motions could not be easily separated as nodding sometimes involved body motion, and bowing and leaning involved both head and body motion. Therefore, nodding, bowing, and forward leaning were represented as a single "concept", which we called "nod-like", and the head and body channels were combined into a single "category". We also divided touch into two categories based on whether actions were mostly characterized by manner of contact or how the robot's body was moved (touch vs. inertial). Next, a bag of words approach was taken to split rows in which a participant performed several actions, such as hugging and patting, into separate rows; our assumption was that affection conveyed depends mostly not on the order of actions, but rather on which actions were performed. This resulted in 2170 coded actions, each associated with affection scores and reasons for interacting. Finally, infrequent outliers were removed. This allowed us to show our conceptualization of typical behavior in Fig. 2. Some labels used are defined in Table 2.

Behavior was assessed in terms of frequency, degree of affection, and monosemy:

Frequency Of 2170 coded actions, touch and proprioceptive actions were observed in 223 cases, visual actions were numerous at 1575, and sound comprised 372 speech samples (1686 words); speech was very common, with participants speaking in 85% of cases, five participants speaking for each reason (22 times) and all participants speaking at least nine times (a little less than half the time). A Chi-squared test confirmed that some touches were observed significantly more frequently than some other behavior: for example, hand-shaking was more frequent than waving ($\chi^2(1, N = 438) = 26.45$, $p < .001$) and shoulder-patting was more frequent than head-shaking ($\chi^2(1, N = 438) = 8.16$, $p < .005$. Thus, Prediction P2a, that touches would be observed, was supported.

Affection Analysis was conducted on the affection scores associated with the modalities, channels and actions shown in Fig. 2. A one-tailed t-test indicated that the total average affection score for all behavior was slightly positive (4.7 ± 1.7: $t(432) = 8.17$, $p < .001$), indicating that participants expected their behavior overall to convey some liking to a robot. A one-way ANOVA showed a statistically significant difference between the degree of affection perceived to be conveyed by modalities ($F(2, 1579) = 9.06$, $p < .001$). A Tukey post-hoc test revealed that touch was perceived to communicate significantly more affection than sound or vision (5.3 ± 1.5 vs. 4.8 ± 1.6, 4.8 ± 1.7; $p < .001$); no statistically significant differences were seen between vision and sound ($p = .93$). A second one-way ANOVA showed likewise a difference between channels: $F(7, 1574) = 6.05$, $p < .001$. A Tukey post-hoc test revealed that touch was perceived to communicate significantly more affection than proxemics ($p = .027$) and all other channels ($p < .001$), with no other significant differences.

A one-way ANOVA showed a difference between actions: $F(49, 1532) = 6.19$, $p < .001$. Hugging (average affection score: 6.5 ± 1.3), head-patting (6.4 ± .70), and head rubbing (6.4 ± .65) were perceived to communicate more affection than 14, 13, and 11 other actions respectively. Specifically, hugging was perceived as significantly more affectionate than stop, ominous, harsh, commanding, and distancing ($p < .001$); up-down, waving, neutral speech, smiling, and tilting or turning to the side ($p < .05$); with a significant trend compared with pointing, fast speech, and looking at the robot's body ($p < .1$). Thus, Prediction 2, that touches would be perceived as conveying high affection, was supported.

Monosemy Analysis was conducted in two steps: first, we identified reasons for interacting which are important for conveying affection; then we found which behavior was most associated with these reasons. A repeated measures ANOVA indicated that participants perceived some reasons for interacting to involve significantly more affection than others: $F(21, 252) = 29.14$, $p < .001$, $\eta^2 = .71$. Post hoc tests using the Bonferroni correction revealed that love and thanking (average degree of affection: 6.6 ± 0.51; 6.5 ± .66) were associated with significantly more affectionate behavior than 12 other reasons for interacting. Indicating attractiveness (6.1 ± .86), reassuring (6.0 ± .71), and accepting (6.1 ± .76) were more affectionate than six, five, and five other reasons. In contrast, Anger, (1.5 ± .52) Disgust (1.8 ± .73), Dominate (2.9 ± 1.2), and Stop (3.3 ± 1.3) were significantly less affectionate than 20, 18, 10, and 10 other reasons ($p < .05$). Thus, we can expect affectionate interactions to result from a desire to show love, thank, praise, reassure, or show acceptance toward a robot; unaffectionate interactions will result from showing anger or disgust, dominating or stopping a robot.

Next, we checked which behavior was most closely associated with these reasons. Analysis was conducted to find monosemic actions using a variant of the Term Frequency-Inverse Document Frequency (TF-IDF) heuristic[60] as (n/t-n) where n is the number of times a behavior was performed for its most frequent associated reason and t is the total number of times a behavior was performed; this variant was used because the

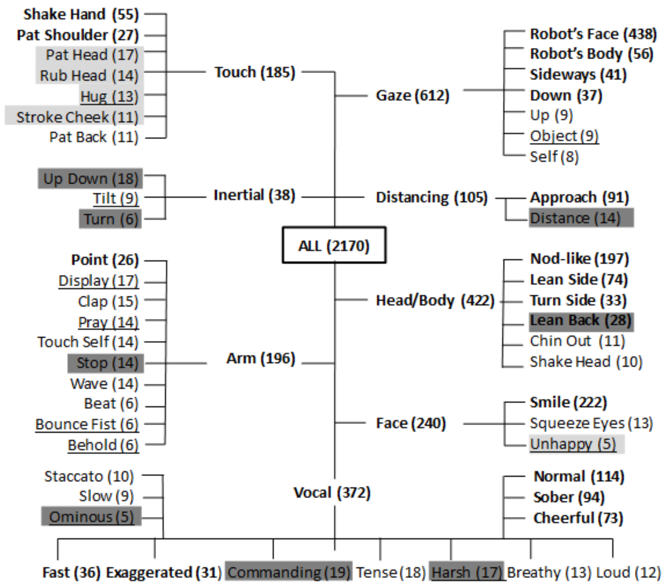

Fig. 2. Participants' typical behavior toward our robot: "Touch", "Inertial" etc. indicate channels; "Shake Hand", "Pat Shoulder", etc., indicate actions; numbers in parentheses indicate observed incidents, with actions occurring less than five times not shown and bold font indicating the most frequently occurring actions (observed more than twenty times); dark and light shading indicate the most and least affectionate actions (all actions with affections lower than four and higher than six); and underlines indicate the most monosemic actions (those strongly associated with a single designated reason, TF-IDF score ≥ 1.0).

number of reasons for interacting (documents) is very small. The TF-IDF heuristic is useful for finding which "terms" are important for characterizing a "document" or which documents are most associated with a term. In our case, we are interested in which cues are associated with affectionate reasons, with a large score indicating high association. For modalities and channels, we found that touches were performed most often to show love (23 cases), and reassure (23) with a maximum monosemy of 0.12; vision and sound channels were not most associated with the highly affectionate or unaffectionate reasons. For actions, we found that hugging and head-rubbing were most closely associated with love (monosemy scores: 1.0; 0.62), back-patting with reassuring (0.57), "Behold" and cheek-stroking with showing attractiveness (2.0, 0.83), and hand-shaking with showing acceptance (0.22). For unaffectionate reasons, harsh speech was most associated with anger (1.1), distancing and leaning back with disgust (0.40, 0.27), speaking ominously and lifting a robot with dominating (0.38, 1.5), and making a stop gesture, speaking in a commanding or staccato voice, and patting the robot's shoulder with stopping it (0.75, 0.46, 0.25, 0.17). Thus, Prediction 2c, that some touches would be highly associated with affectionate ways of interacting, was supported.

5.3. *Thin-slice video clips: conflicting signals*

A repeated measures ANOVA with a Greenhouse-Geisser correction showed that

Table 2. Descriptions of some observed actions.

Action	Description
"Nod-like"	Closely-associated head and body behavior such as forward-leaning, swaying, nodding, and bowing
"Behold"	Directing palms toward the robot sometimes accompanied by a rotation of the hands, with some resemblance to a "Kamehameha" gesture in Japanese pop culture, or to the "Orans" gesture associated with prayer in antiquity
"Bounce Fist"	Beat-like action in which one or two fists directed toward the robot were shaken up and down, with thumbs on top and arms held to the side
"Praying"	Performed with open hands, except in two cases featuring folded hands
"Displaying"	Rotating or moving an object in an arc in front of the robot's face
"Pointing"	Involved mostly the index finger (12 times) but also sometimes a thumb (three times) or all fingers (two times)
"Sober"	Gentle, quiet, serious, sad, and subdued
"Harsh"	Biting, ballistic, growly, displeased, and grumbly
"Commanding"	Prohibitive, strict, and firm
"Exaggerated"	Involved prolonged sounds and emphasis

Infrequent actions included an okay sign; arm-flapping; finger-flicking; stroking the robot's wheel, opening one's eyes wide; a Japanese gesture in which arms are crossed to express "no"; a shooing action and hands pressed into the chest to express disgust; a cyclic motion resembling eating with a fork used to indicate that the robot should continue; as well as deictic "line-tracing" used to indicate an imaginary barrier the robot should not cross, a goal for the robot to move toward, or attractive features on the robot.

participants perceived a difference in the affection communicated in the videos: $F(3.22, 61.14) = 14.59$, $p <.001$, $\eta^2 = .43$. Post hoc tests using the Bonferroni correction revealed that participants attributed significantly more affection to +Touch, -Vision than in all other cases ($p < .05$); (this was the only case which elicited average scores greater than five, meaning some affection was perceived: 5.4 ± 1.5). Also, a one-tailed t-test indicated that the average affection perceived was slightly negative (3.6 ± 1.7: $t(119) = 2.55$, $p < .05$); some participants described the behavior as domestic violence or sarcasm.

For further insight, scores were combined for each positive and negative signal; e.g., scores for +Touch, -Vision and +Touch, -Sound were grouped under +Touch. A second repeated measures ANOVA with a Greenhouse-Geisser correction confirmed that the affection communicated by each modality differed: $F(4.37, 170.24) = 18.04$, $p < .001$, $\eta^2 = .32$. Post hoc tests using the Bonferroni correction revealed that +Touch (4.5 ± 1.7) was perceived as significantly more affectionate than +Sound (3.8 ± 1.5, $p = .033$) and +Vision (2.5 ± 1.3, $p < .001$). Results are shown in Fig. 3. Thus, Prediction 2d, that touch would be important for communicating affection, was supported.

6. Discussion

In summary, the primary contribution of the current work was to confirm the importance of touch for conveying affection in a multimodal interaction with a humanoid robot. Specific results are detailed below:

- Our initial assumptions were partially supported. A new robot prototype, Penumbra, successfully conveyed an impression of being capable of recognizing various signals communicated via touch, vision and sound. All reasons for interacting except negative emotion displays were deemed likely. The fact that playing and expressing loneliness were deemed more likely than other reasons, in the sense that loneliness may arise from lack of perceived affection, supports an intuition in our ongoing work that affectionate play is a worthwhile context to study in human-robot interaction. Also, all channels which we felt might be important, but especially touch, were expected to communicate affection.

- Our results suggested the importance of touch for communicating affection. Touches were observed often, in approximately half of the cases in which participants interacted, confirming that people would touch even a robot believed capable of visual or aural communication; touches were perceived overall as conveying high affection overall, and hugging and patting or rubbing the robot's head were deemed significantly more affectionate than approximately a third of the observed typical actions; touches like hugging were closely associated with affectionate reasons such as showing love; and touch played a significantly important role in how conflicting signals were perceived, possibly because touches involve a large clear motion which could be costly if rejected.

We present some discussion based on these results revolving around three main points: comparison with previous work, generality and limitations, and future work.

6.1. *Comparison with other studies*

Comparison conducted with our previous results on touch and proprioception[11, 12] showed that most typical touches identified in our previous study were observed, with a few exceptions. Cover Face and Kiss were not observed. One form of Minimize Touch in which a participant held the robot with only their fingers was not observed, although poking touches were (Penumbra was heavier than the mock-up we used in our previous study); head-striking involved not the cheeks but the top of the robot's head twice with an open hand and once with a fist (this was facilitated by Penumbra's lower height compared to our previous mock-ups). Also, not all of the proprioceptive acts we identified in our earlier study on proprioception and play were observed. This small difference was not surprising, because in our previous study participants were restricted to using touch, whereas in the current work they often expressed themselves visually and aurally. Therefore, our prediction that we would observe all the touches observed in our previous studies was only partially supported.

Further comparison was made to a previous study which reported participants speaking to a humanoid robot as if it were an infant using four typical patterns.33 A few of our participants, but not all, spoke with our robot as with a child. The reason could be the robot's appearance; this seems intuitive based on depictions of different humanoid robots in pop culture such as Number Five, which was spoken to sometimes as a child, and Terminator, which was spoken to like an adult. Another reason was that we asked people to show anger, disgust, or dominance, which are not typically associated with child-directed speech. Additionally, we found that participants did not expect to show hate or disgust toward our robot, and expected their own actions to mostly express affection, which bears similarity to a result observed by Yohanan and MacLean[32]. We think this suggests a generally positive bent to how people expect to interact with robots.

6.2. *Generality*

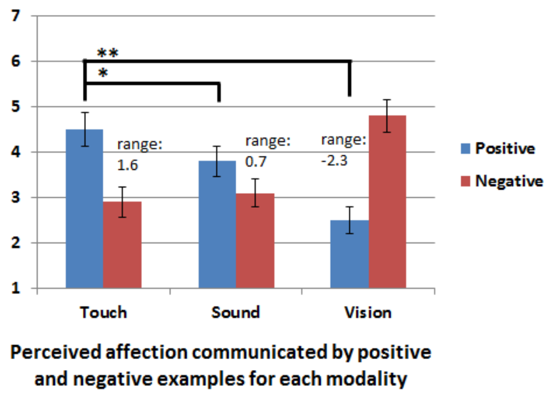

Fig. 3. Touch had a significant impact on perceived affection in mixed signals: **p < .001, * p < .05

We feel it is important to clearly state that we cannot know how a human will behave toward a robot based on this article alone. Our results are limited by the highly specific interaction scenario assumed, involving a dyadic interaction with a young Japanese adult who is not busy, far away, or has her hands full, and behaves for typical reasons, toward a small humanoid robot which can perceive multimodal signals, in a controlled environment without objects. Changing any property of this scenario will affect the behavior that will be observed. Nevertheless we feel that the selected scenario was useful toward providing some indication. For example, we expect that people interacting with a robot usually are willing and capable of interacting, and dyadic interactions are common. Also, although it is known that non-verbal behavior exhibits high cultural dependency,[61] most of the behavior we observed and analyzed such as hand-shaking, pointing, and leaning is not inherently Japanese; this could be a result of cultural hybridization,[62] and to some extent possibly also innate disposition.[63] As well, we expect robots will use touch, vision, and sound because these are familiar communication modalities for people.[10]

Another limitation regards methodology. The results presented were not obtained following a common pattern of using statistical significance tests to provide strong support for or disprove a small set of hypotheses. We also relied on subjective measurements, and thin-sliced video sequences and sounds involved the behavior of a specific confederate; behavior performed by other individuals could be interpreted differently. Motivation for our approach stems from the exploratory nature of the current work; like previous work by Yohanan et al.,[32] it is intended to act as a first step toward gaining insight into a complex phenomenon which is currently not well understood. A precedent for such an approach may be found in grounded theory[58]. As well, subjective measurements presented the only way to directly assess people's perceptions [64].

6.3. *Future work*

Our next steps will involve leveraging our current results to recognize multimodal affection and investigating other scenarios in order to improve our understanding of people's behavior toward a robot. Our basic recognition approach will involve classifying some useful multimodal actions listed in Section 5.2, and predicting overall affectionate meaning by weighting the contributions of simultaneous signals using the affection scores for modalities in Section 5.3. Many other topics will also need to be addressed, such as how to distinguish social and non-social smiling. To better model people's behavior, cues will be identified to predict a person's interactive capabilities (a person holding bags or an umbrella may not be able to touch a robot). Interactions with multiple persons and robots will be analyzed, as well as social behavior involving objects; e.g., we have observed onlookers taking photos of robots with cell-phones and cameras, and children showing off toys. Also, we will investigate how people behave toward non-humanoid robots with various or arbitrary forms. This work will contribute toward realizing rich affectionate interactions.

Acknowledgments

We'd like to thank everyone who helped. This research was supported by JST, CREST.

References

1. M. Pantic, R. Cowie, F. D'Errico, et al., Social Signal Processing: The Research Agenda, *Laser Physics Review* (2011) 511-538.
2. H. Harlow, The Nature of Love, *American Psychologist* **13**(12) (1958) 573–685.
3. J. W. Prescott, Body pleasure and the origins of violence, *Bull Atom Sci* (1975) 10-20.
4. A. H. Maslow, A theory of human motivation, *Psychological Review*, **50**(4) (1943) 370–96.
5. S. Turkle, C. Breazeal, B. Scassellati and O. Dasté, Encounters with Kismet and Cog: Children's Relationships with Humanoid Robots, in *IEEE-RAS International Conference on Humanoid Robots* (Humanoids) (Los Angeles, CA, 2004).
6. L. F. Barrett, Valence as a basic building block of emotional life, *J Res Pers* **40** (2006) 35–55.
7. J. Russell, A circumplex model of affect, *J Pers Soc Psychol* **39** (1980) 1161–1178.
8. A. Mehrabian and S. R. Ferris, Inference of Attitudes from Nonverbal Communication in Two Channels, *Journal of Consulting Psychology* **31**(3) (1967) 48-258.
9. I. Poggi, F. D'Errico and L. Vincze, Types of Nods. The Polysemy of a Social Signal, in *International Language Resources and Evaluation (LREC)*, (Valletta, Malta, 2010).
10. A. G. Brooks and R. C. Arkin, Behavioral overlays for non-verbal communication expression on a humanoid robot, *Autonomous Robots* **22**(1) (2007) 55-74.
11. M. D. Cooney, S. Nishio, and H. Ishiguro, Recognizing Affection for a Touch-based Interaction with a Humanoid Robot, in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, (IEEE Press, Algarve, Portugal, 2012), pp. 1420-1427.
12. M. D. Cooney, C. Becker-Asano, T. Kanda, A. Alissandrakis, and H. Ishiguro, Full-body Action Recognition Using Inertial Sensors for Playful Interaction with Small Humanoid Robot, in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, (IEEE Press, Taipei, Taiwan, 2010), pp. 2276-2282.
13. M. Heerink, B. Kröse, B. Wielinga and V. Evers, Enjoyment, Intention to Use And Actual Use of a Conversational Robot by Elderly People, in *ACM/IEEE International Conference on Human-Robot Interaction (HRI)* (ACM, New York, 2008), pp. 113-119.
14. M. Csíkszentmihályi, K. Rathunde K, The measurement of flow in everyday life: Towards a theory of emergent motivation, in *Developmental perspectives on motivation. Nebraska symposium on motivation*, ed. J. E. Jacobs (University of Nebraska Press, Lincoln, 1993).

15. J. Bowlby, The nature of the child's tie to his mother, *Int J Psychoanal* **39** (5) (1958) 350–373.
16. M. J. Ellis, Why people play. (Prentice-Hall, Englewood Cliffs, New Jersey, 1973).
17. V. Venkatesh, Determinants of perceived ease of use: integrating control, intrinsic motivation, and emotion into the technology acceptance model, *Inf Sys Research* **11** (2000) 342–365.
18. M. Y. Yi and Y. Hwang, Predicting the use of web-based information systems: self-efficacy, enjoyment, learning goal orientation, and the technology acceptance model, *International Journal of Human-Computer Studies* **59**(4) (2003) 431-449.
19. K. Wada and T. Shibata, Living With Seal Robots - Its Sociopsychological and Physiological Influences on the Elderly at a Care House, *IEEE Trans Robot* **23**(5) (2007) 972-980.
20. M. Kanamori et al., Pilot study on improvement of quality of life among elderly using a pet–type robot, in *IEEE international symposium on computational intelligence in robotics and automation* (CIRA), (IEEE Press, Kobe, Japan, 2003), pp. 107–112.
21. J. Cacioppo and W. Patrick, Loneliness: Human Nature and the Need for Social Connection. (W.W. Norton & Co., New York, 2008).
22. R. S. Wilson et al., Loneliness and Risk of Alzheimer Disease, *Arch Gen Psych* **64** (2007) 234-240.
23. Kohei Ogawa, Shuichi Nishio, Kensuke Koda, Giuseppe Balistreri, Tetsuya Watanabe and Hiroshi Ishiguro, Exploring the Natural Reaction of Young and Aged Person with Telenoid in a Real World, *J Adv Comp Intell Inform* **15**(5) (2011) 592-597.
24. K. Dautenhahn, C. L. Nehaniv, M. L. Walters, B. Robins, H. Kose-Bagci, N. A. Mirza, and M. Blow, KASPAR - A Minimally Expressive Humanoid Robot for Human-Robot Interaction Research, in *Special Issue on Humanoid Robots, Applied Bionics and Biomechanics* **6**(3) (2009) 369-397.
25. A. Vinciarelli, H. Salamin, H. and M. Pantic, Social Signal Processing: Understanding Social Actions through nonverbal behaviour analysis, in IEEE International Conference on Computer Vision and Pattern Recognition (CVPR) (Miami, FL, USA 2009).
26. I. McCowan, D. Gatica-Perez, Samy Bengio, Guillaume Lathoud, Mark Barnard and Dong Zhang, Automatic Analysis of Multimodal Group Actions in Meetings, *IEEE Trans. Pat Anal Mach Int* **27**(3) (2005) 305-317.
27. N. Ambady and R. Rosenthal, Thin slices of expressive behavior as predictors of interpersonal consequences: A meta-analysis, *Psychological Bulletin* **111** (1992) 256–274.
28. T. Noda, T. Miyashita, H. Ishiguro and N. Hagita, Map acquisition and classification of haptic interaction using cross correlation between distributed tactile sensors on the whole body surface, in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, (IEEE Press, California, U.S.A., 2007), pp. 1099-1105.
29. H. Knight, R. L. Toscano, W. D. Stiehl, A. Chang, Y. Wang and C. Breazeal, Real-time social touch action recognition for sensate robots, in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, (IEEE Press, U.S.A., 2009), pp. 3715-3720.
30. W. D. Stiehl and C. Breazeal, Affective Touch for Robotic Companions, in *International Conference on Affective Computing and Intelligent Interaction (ACII) LNCS 3784*, (Beijing, China, 2005), pp. 747-754.
31. D. François, D. Polani and K. Dautenhahn, On-line behaviour classification and adaptation to human-robot interaction styles, in *ACM/IEEE International Conference on Human-Robot Interaction (HRI)* (ACM, Washington DC, USA, 2007), pp. 295-302.
32. S. Yohanan and K. E. MacLean, The Role of Affective Touch in Human-Robot Interaction: Human Intent and Expectations in Touching the Haptic Creature, *Int J Soc Rob* **4**(2) (2012) 163-180.
33. C. Breazeal and L. Aryananda, Recognition of Affective Communicative Intent in Robot-Directed Speech, *Autonomous Robots* **12**(1) (2002) 83-104.

34. T. Minato, M. Shimada, S. Itakura, K. Lee, H. Ishiguro, Evaluating the human likeness of an android by comparing gaze behaviors elicited by the android and a person, *Advanced Robotics* **20**(10) (2006) 1147-1163
35. T. V. Oosterhout and A. Visser, A Visual Method for Robot Proxemics Measurements, in *ACM/IEEE International Conference on Human-Robot Interaction (HRI) Workshop on Metrics for Human-Robot Interaction*, (ACM, Amsterdam, 2008) Technical Report 471.
36. E. Oztop, D. W. Franklin, T. Chaminade and G. Cheng, Human–humanoid interaction: Is a humanoid robot perceived as a human? *Int J Hum Robot* **2**(4) (2005) 537-559.
37. W. S. Ng and E. Sharlin, Collocated interaction with flying robots, in *IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*, (IEEE Press, Atlanta, Georgia, 2011), pp. 143-149.
38. C. Liu, C. T. Ishi, H. Ishiguro, N. Hagita, Generation of Nodding, Head tilting and Gazing for Human-Robot Speech Interaction, *Int J Hum Robot* **10**(1) (2013) 1350009-1-19.
39. J. Broekens, Emotion and Reinforcement: Affective Facial Expressions Facilitate Robot Learning, *Artifical Intelligence for Human Computing* (2007) 113-132.
40. A. Mehrabian, Orientation behaviors and nonverbal attitude communication, *Journal of Communication* **17** (1967) 324–332.
41. E. T. Hall, The Hidden Dimension. (Anchor Books, New York, 1966).
42. P. E. Bull, Posture and action. (Pergamon Press, Oxford, 1987).
43. M. L. Knapp and J. A. Hall, Nonverbal Communication in Human Interaction (6th ed.) (Wadsworth/Thomson Learning, Belmont, CA, 2006).
44. L. E. Williams and J. A. Bargh, Experiencing physical warmth promotes interpersonal warmth, *Science* **322** (2008) 606-607.
45. M. LaFrance and M. A. Hecht, Why smiles generate leniency, *Pers Soc Psychol Bull* **21** (1995) 207-214.
46. P. Ekman and W. V. Friesen, The repertoire of nonverbal behavior: Categories, origins, usage, and coding, *Semiotica* **1** (1969) 49- 98.
47. K. Grammera, K. Kruck, A. Juettea and B. Fink, Non-verbal behavior as courtship signals: the role of control and choice in selecting partners, *Evol Hum Behav* **21** (2000) 371-390.
48. B. Laeng and L. Falkenberg, Women's pupillary responses to sexually significant others during the hormonal cycle, *Horm Behav* **52**(4) (2007) 520-530.
49. T. L. Chartrand and J. A. Bargh, The chameleon effect: The perception–behavior link and social interaction, *Journal of Personality and Social Psychology*, **76**(6) (1999) 893-910.
50. S. E. Jones and A. E. Yarbrough, A naturalistic study of the meanings of touch, *Communication Monographs* **52** (1985) 19-56.
51. P. Ekman and W. B. Friesen, The repertoire of nonverbal behaviour: categories, origins, usage, and coding, in *Nonverbal communication, interaction and action*, ed. A. Kendon (Mouton Publishers, The Hague, Paris, 1981).
52. K. S. Verderber, R. F. Verderber and C. Berryman-Fink, Inter-Act: Interpersonal Communication Concepts, Skills, and Contexts, (Oxford University Press, U.S.A, 2009).
53. J. Dore, Variation in preschool children's conversational performances, in *Children's language Vol. 1*, ed. K. Nelson (Gardner Press, New York, 1978), p. 397-444.
54. A. Mehrabian and M. Wiener, Decoding of Inconsistent Communications, *J Pers Soc Psych* **6**(1) (1967) 109–114.
55. N. M. Gada, F. Bernieri, J. E. Grahe, D. Zuroff, and R. Koestner, Love: How do observers perceive it, presented at the *Midwestern Psychological Association convention*, (Chicago, IL 1997).
56. M. J. Hertenstein, D. Keltner, B. App, B. Bulleit and A. Jaskolka, Touch communicates distinct emotions, *Emotion* **6** (2006) 528-533.

57. M. A. Sayette, J. F. Cohn, J. M. Wertz, M. A. Perrott, and D. J. Parrott, A Psychometric Evaluation of the Facial Action Coding System for Assessing Spontaneous Expression, *Journal of Nonverbal Behavior* **25** (2001) 167-186.
58. B. G. Glaser and A. L. Strauss, The discovery of grounded theory: strategies for qualitative research. (Aldine, Chicago, 1967).
59. W. T. Freeman and M. Roth, Orientation histograms for hand action recognition, in *International Workshop on Automatic Face and Action Recognition Vol. 12* (IEEE Computer Society, Los Alamitos, 1995) pp. 296-301.
60. C. D. Manning, P. Raghavan, and H. Schutze, Scoring, term weighting, and the vector space model. Introduction to Information Retrieval. (Cambridge University Press, New York, 2008). doi:10.1017/CBO9780511809071.007.
61. M. S. Kirch, Non-Verbal Communication Across Cultures, *Modern Language Journal* **63**(8) (1979) 417.
62. R. Robertson, Glocalization: Time-Space and Homogeneity-Heterogeneity, in Global Modernities, eds. M. Featherstone, S. Lash and R. Robertson (Sage, London, 1995).
63. J. L. Tracy and D. Matsumoto, The spontaneous display of pride and shame: Evidence for biologically innate nonverbal displays, *Proc Nat Acad Sci* **105** (2008) 11655-11660.
64. W. H. Cushman and D. J. Rosenberg, Human factors in product design. (Elsevier, New York, 1991).

Martin D. Cooney received a D. Eng. degree in engineering science from Osaka University in 2014, and was an Intern Researcher at Advanced Telecommunications Research Institute International (ATR), Kyoto, from 2008 to 2014.

Shuichi Nishio received his M.Sc. in computer science from Kyoto University in 1994, and a D. Eng from Osaka University in 2010, and is currently a Senior Researcher at the Advanced Telecommunications Research Institute International (ATR) in Kyoto.

Hiroshi Ishiguro received in 1991 a D. Eng. degree from Osaka University and was a Research Assistant at the Department of Electrical Engineering and Computer Science, Yamanashi University. In 1992, he joined the Department of Systems Engineering, Osaka University, as a Research Assistant. In 1994, he became an Associate Professor at the Department of Information Science, Kyoto University, where he researched distributed vision using omnidirectional cameras. From 1998 to 1999, he was a Visiting Scholar at the Department of Electrical and Computer Engineering, University of California, San Diego. In 1999, he was a Visiting Researcher at Advanced Telecommunications Research Institute International (ATR) Media Information Science Laboratories. He became an Associate Professor in 2000 and Professor in 2001 at the Department of Computer and Communication Sciences, Wakayama University. He is currently a Distinguished Professor at the Department of Systems Innovation, Graduate School of Engineering Science, Osaka University, and a Group Leader of Hiroshi Ishiguro Laboratory at ATR. His research interests include sensor networks, interactive robotics, and android science.