Toward Natural Human-Robot Interaction:
Exploring Facial Expression Synthesis on an Android Robot

Laurel D. Riek [*]
Computer Laboratory, University of Cambridge
15 JJ Thompson Avenue, Cambridge, CB3 0FD, UK
`Laurel.Riek@cl.cam.ac.uk`

Abstract

Robots are entering domestic environments in increasing number. However, the present means for interacting with them is quite limited and difficult for people who are not technically inclined or able-bodied. Thus, many in the field of robotics are moving toward natural human robot interaction, which means allowing people the ability to communicate with robots in ways similar to how they communicate with other people, i.e., via verbal and nonverbal channels. My work focuses on a part of this problem: how to accurately model and synthesize natural human facial expressions on a realistic, human-like robot head. The main goal of this research is to see if by providing such expressions on a robot lead to people feeling more at ease while interacting with it, and thus be more likely to accept such technology in domestic settings.

Keywords: Affective Computing, Emotion Synthesis, Empathy, Human-Robot Interaction

1 INTRODUCTION

Each year, robots are entering domestic environments in greater and greater numbers. According to a 2007 report by the International Federation of Robotics, 3.4 million personal service robots are in use worldwide. The report forecasts that this number is expected to increase by 4.6 million robots by 2012 (IFR, 2008). These domestic robots are being used to serve as health aids and companions, help with household chores, and provide education and entertainment to their users.

The domestic robot user presents a unique challenge to robot designers. Elderly users are likely to be uncomfortable with domestic robots due to a lack of exposure to technology, disabled users might have difficulty using robots that do not provide interaction modalities that accommodate their needs, and people using robots for household chore assistance are unlikely to have much time to devote to learning to use complexly designed systems. One way to address some of these problems is to design robots that allow people the ability to interact with robots naturally.

Natural interaction means allowing people the ability to communicate with robots in ways similar to how they communicate with other people. This includes both verbal communication (speech and non-speech vocalization) and nonverbal communication (body gesture, gaze, movement, and facial expression). Most people are able to express themselves in this way and easily interpret such expressions in others. While people generally do not expect such ease of interaction with machines, evidence suggests having it would help improve user engagement with the robot (Sidner et al., 2005). Indeed, by taking advantage of these interactive modalities, robot designers can go a long way toward ensuring their robots are accepted.

The task of enabling natural human-robot interaction is by no means trivial; it is a complex problem that requires drawing from many areas of computer science (computer vision, human-computer interaction, natural language processing, machine learning, robotics), cognitive science, and the social sciences. It requires being able to accurately characterize how humans interact

---

[*] Parts of this paper have previously appeared in the following publications: (Riek and Robinson, 2008), (Riek and Robinson, 2009), (Riek et al., 2009a), (Riek et al., 2009b), (Riek et al., 2009c)

with one another so that robots can be programmed to appropriately recognize or synthesize such behaviors themselves. It also requires robots to have at least a rudimentary understanding of the social context in which they are placed.

My work tackles a small piece of this larger problem; namely, how to accurately model and synthesize natural human facial expressions on a realistic, human-like android robot head. This is an unexplored area of research; to date most emotion synthesis work on facially-expressive robots has used zoomorphic or mechanical-looking robots with limited expressivity. Furthermore, most of these robots display exaggerated, non-naturalistic, repetitive facial expressions that are generated by an actor or animator. My approach differs in that it is based on a large collection of human facial expressions made in naturalistic conversational settings, which are then normalized to create a great repertoire that the robot can use to generate novel expressions.

The main goal of this research is to see if by providing such natural, human-like expressions on a robot people feel more at ease while interacting with it. And, further, that people may be more willing to accept the presence of such a robot in domestic settings.

2 Background

The problem of natural interaction with robots is something that spans across a variety of different disciplines. It is not practical to give a detailed survey of each one; instead I will briefly introduce the two most relevant to my work: Affective Computing and Human-Robot Interaction.

2.1 Affective Computing

Affective computing is a discipline dedicated to the idea of giving machines the ability to recognize and generate affect (Picard, 1997). In some ways, the field exists to address the failings of traditional human-computer interactive systems, which typically neglect affective state changes in users. In fact, some argue that such neglect is a reason many users view interactions with computers as “cold, incompetent and socially inept.” To address this, several leaders in the field have said that it is critical that interfaces of the future are able to “detect subtleties of and changes in the user’s behavior, especially his/her affective behavior, and to initiate interactions based on this information rather than simply responding to the user’s commands” (Zeng et al., 2009).

Until recently, most of the approaches to affect recognition centered around posed data with exaggerated affective expressions, were limited to a small set of emotions (such as anger, fear, and happiness), and were restricted to single modes of expression (just face or just speech). However, the field is now shifting toward looking at recognizing multi-modal, less-constrained naturalistic expressions (Zeng et al., 2009). For example, el Kaliouby (2005) worked on the generalization of facial affect inference for complex mental states while Sobol Shikler (2007) worked on inferring affect from naturally-evoked speech. Bernhardt and Robinson (2007) worked on inferring affect from body posture and gesture.

2.2 Human-Robot Interaction

In the field of human-robot interaction (HRI), quite a number of interactive robots have been designed to try to facilitate natural interaction by recognizing and generating affect. Breazeal et al. (2008) and Fong et al. (2003) present thorough surveys of many such robots and their theoretical emotional underpinings. I will present a subset of these robots and also introduce a few others using role categories commonly used in the literature. For each category I will list the names of some representative robots, and highlight one robot in particular as an example.

2.2.1 Epigenetic Robots (Cog, HOAP-3, iCub, *Kismet*, Leonardo)

A number of interactive robots have been created with some degree of affective understanding and generation capability using an epigenetic approach. This approach uses ideas from developmental psychology to help robots learn sophisticated social behaviors (Scassellati, 1998). Many of these developmentally-based robots inherently take social context into account in order to learn to adapt to the humans interacting with them. One of the first of these robots is Kismet, an

Figure 1: A few exemplar robots that recognize and generate affect. From left to right: KeepOn (*Photo: Janne Moren*), PARO (*Photo: Shoko*), Kismet (*Photo: Carol Nichols*).

anthropomorphic, expressive robot designed entirely for emotional interaction with humans. By understanding the social cues of humans in the environment, Kismet is able to respond in an emotionally appropriate way to people (Breazeal, 2002). Its thoughtful design has lead to it being a very well accepted and regarded robot.

2.2.2 ENTERTAINMENT ROBOTS (AIBO, ASIMO, AUR, *Keepon*)

Kozima and Michalowski were interested in building a robot that could interact with children in a pleasant and natural way. Their first attempt was the Infanoid robot, which was a highly mechanical-looking, very expressive robot. From observational studies the researchers found that the appearance and behavior of this robot was overwhelming children. This insight led them to the successful design of the robot Keepon, which is a minimally-designed interactive dancing robot. The robot only has 4 degrees of freedom, but is easily able to express attention via head direction and emotion via rocking motions. Its design was well informed by observing hundreds of children interacting with the robot for over 400 hours in total (Kozima et al., 2008).

2.2.3 THERAPEUTIC ROBOTS (HUGGABLE, ICAT, KASPAR, *PARO*, SHYBOT)

Shibata et al. (1997) describe their desire to build an affect robotic pet that was capable of sensing the emotions of the people it was interacting with and alter its affect accordingly. From the outset they concerned themselves with how their robot would interact emotionally with users, and tailored the robot's design accordingly. This mindset led the researchers to later create the very successful implementation of PARO the robotic seal, which has been used effectively to reduce stress and depression among the elderly (Wada and Shibata, 2007).

2.2.4 PEER ROBOTS (VIKIA, ROBONAUT, *Valerie*, GRACE, MEL)

Kirby et al. designed Valerie, a robot receptionist designed to facilitate long-term social interaction with people. The robot was thoughtfully designed to facilitate natural interaction - the robot's physical appearance, its station, and its behaviors were carefully considered to create an engaging experience with users (Kirby and et al., 2005).

2.2.5 MENTOR ROBOTS (*Basketball Coach*, CHIPS, ROCO)

Liu et al. (2006) describe a robotic basketball coach that monitored the physiological signals (heartrate and galvanic skin response) of people while they shot baskets. Depending on how anxious people seemed to be, the robot altered the game's level of difficulty. The researchers found through this style of interactive teaching people's performance improved.

3 WORK TO DATE

3.1 AFFECTIVE CENTERED DESIGN PARADIGM

As I began studying previous work in HRI, HCI, and Affective Computing, I came to realize that the fields had very little overlap when it came to robot design. In particular, one of the primary design methodologies espoused by interactive robot engineers, human-centered design, completely lacked an explicit affective element. And, unsurprisingly, most commercially available robots lack any sort of ability to recognize or generate affect.

Thus, I attempted to address this problem by introducing a new paradigm for the design of interactive robots called affective-centered design (Riek and Robinson, 2009). The idea behind this contribution was to provide practical guidelines to interactive robot designers to help them improve the affect aspects of their robots' designs. Furthermore, for my own research, it was helpful to think about ways in which one can measure successful interactions.

3.2 Empathizing with Robots that Mimic Expressions in Real-Time

When evaluating the affective quality of natural human-robot interaction, it is very important to consider the expression of empathy. Empathy expression is a key aspect of human-human social communication that allows people to experience and understand what others are emotionally conveying (Ross et al., 2008). One of the most basic forms of expressive empathy is known as emotional contagion, where an observer mimics the behavior of a target, and by virtue of that mimicry, comes to experience an emotional state similar to that of the target (Davis, 2006).

Facial expression and head gesture mirroring are common forms of empathic conveyance that typically include head nodding, laughing, smiling, etc. This mirroring is so vital to emotional communication that if an individual's ability to mirror others is physically blocked, that individual will actually be impaired in their ability to identify emotions (Oberman et al., 2007).

Given how important facial mimicry is in human-human communication, I wondered if it might also be important in human-machine communication. In particular, might a conversational robot that mimics a few low-fidelity expressions and head gestures in real-time create a more satisfying interactive experience for people? To address this question, we built a real-time, autonomous, head gesture mimicking robot and performed two pilot studies (Riek and Robinson, 2008; Riek et al., 2009a). In the studies, subjects sat in front of the robot and were asked to perform two conversational tasks. The first task served as an acclimation task, "Please tell the robot the route that you took to the lab today", and the second was intended to be emotionally salient, "Please tell the robot your first memories of arriving in Cambridge." Following the study subjects completed a 7-point (strongly agree - disagree), 13-question survey intended to probe emotional interaction satisfaction.

The quantitative results from these two studies were not statistically significant, however, our qualitative analyses revealed a number of valuable insights. The first surprising insight is that one ought to be cautious when making assumptions about where the robot should "look" when mimicking. We assumed subjects would primarily move their head while talking, and positioned our camera accordingly. However, some subjects never moved their head at all, but moved their head, neck, shoulders, hands, or feet quite a lot. As far as our software was concerned, those subjects were sitting still, and thus they were not mimicked.

The second insight was with regards to response appropriateness. One subject really helped to illustrate this problem by saying that she only ever told the robot positive things; thus, it was less critical for the robot to make the appropriate responses back to her. Whereas if she was saying something negative, sympathetic expressions would be far more important. This was a profound insight, because not only does an empathetic robot need to be able to make appropriate expressions, it also needs to make them at the appropriate time.

The third surprising insight is that people often behave in completely unexpected (albeit interesting) ways during an affective interaction. One subject spoke deliberately slowly to the robot. Some subjects "leaned in" to talk to the robot, or otherwise adopted odd postures. And by far the most interesting unexpected behavior was co-nodding. Two participants co-nodded with the robot; meaning, the participant nodded, the robot nodded in response, and then the participant nodded to acknowledge the robot's nod. This is something worthy of further exploration, as it seems to indicate a high level of engagement with the robot.

3.3 Empathizing with Robots of Varying Degrees of Human-likeness

In the previous studies, the kind of empathy we tested might be thought of as a first-order level of empathy - how does a subject feel after directly interacting with a robot. However, another aspect of empathy is the ability to put one's self into the shoes of another, in other words, 'simulating'

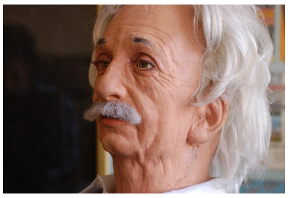

Figure 2: The Einstein robot, developed by Hanson Robotics.

their situation (Goldman, 2006). In the psychology literature, this is a well-established theory called “Simulation Theory”, and might be thought of as a kind of second-order empathy.

In robotics, a debate has been ongoing for many years regarding the degree of human-likeness a robot ought to have. Some researchers argue that robots “should look like robots”, and we ought not build androids. Others argue that we should push the envelope and try to make robots as human-like as possible. The debate is usually framed within the context of the Uncanny Valley, which is a theory proposed by Mori (1970) that posits as robots become more humanlike they become more familiar (and thus more likeable) until the mismatch between their form, interactivity, and motion quality makes people feel uncomfortable. Despite this theory remaining unproven, it is still widely cited and used as justification for robot design decisions (Draude, 2009).

Nonetheless, if Simulation Theory is correct and humans have an innate simulative system, it follows that it should be easier to empathize with the emotions and mental states of a robot that appears similar to us than with one that does not. Certainly this is true for how we interact with other humans - in-group bias and consequent referential treatment can be triggered by markers of physical similarity (e.g. skin color) (Turner, 1978); therefore, it would be unsurprising to find similar effects for how humans interact with robots.

Thus, to test second-order empathy with robots, we conducted a web-based experiment that measured how people empathized with four robots shown to be experiencing mistreatment by humans. The robots varied in appearance from not anthropomorphic (i.e., machine-like) to highly anthropomorphic (i.e., human-like). Subjects viewed both emotionally evocative and emotionally neutral film clips for each protagonist (10 in total). Following each clip, subjects were asked to rate on a scale from 1-6 how sorry they felt for the protagonist. Our quantitative results showed that indeed people empathize more strongly with human-like robots and less-strongly with mechanical-looking ones. Our detailed experimental methodology is described in Riek et al. (2009b).

4 Future Work

Thus far, I have explored the affective aspects of natural interaction with robots, from both an appearance and an empathetic interaction perspective. Next I will focus specifically on synthesizing natural, human-like expressions on an android robot. The main goal of this research is to see if by providing such expressions people feel more at ease during their interactions with the robot, and, further, if they are more willing to accept the presence of such technology in domestic settings.

The robot I will use for my work will look similar to the one depicted in Figure 2. The robot is developed by Hanson Robotics. It has 19 degrees-of-freedom in its face, and each degree of freedom is intended to represent the human facial musculature. Its movements are incredibly natural and life-like; in fact this robot is undoubtedly the most realistic facial humanoid in the world.

The robot resembles an older man, and our choice to acquire such a robot was deliberate for a few reasons. First, an older person has many more wrinkles on their face; thus on a robot it is far easier to create believable facial expressions with little effort. Second, should our future research include elder care, it may be useful to have a robot that looks elderly. Third, while we were ambivalent toward the robot’s apparent gender, it turned out that most of Hanson Robotics’ baseline robots are either male or androgynous, and we wanted to purchase one that had a well-established, believable appearance. Thus, we chose a less iconic version of the Einstein robot.

4.1 NATURAL DATA COLLECTION FROM HUMAN CONVERSATIONAL DYADS

As mentioned previously, posted data is viewed by many people in the affective computing community as less useful than natural data, because it is rarely generalizable to the real world, and in daily life most people make far more subtle expressions. However, to date, most facial robots make exaggerated expressions programmed by an animator, or are otherwise based on acted data. My work aims to fill this gap, by using natural data as the basis for expression synthesis.

I will collect data in a way similar to Morency et al. (2008), in which subjects will be asked to participate in face-to-face, quasi-monologic storytelling dyads. One of the pair will be assigned to be a speaker, and the other a listener. The faces of both pairs of the dyad will be videotaped, though I will primarily be interested in data from subjects who are listeners. This is because I will be programming our robot to behave as an active listener (Bavelas et al., 2000), primarily to avoid issues with speech synchronization, well-known to cause cognitive dissonance in subjects.

4.2 DATA TRANSFORMATION AND NORMALIZATION

Once the video data is collected, I will run each subject's video through the Neven Vision facial feature tracker. The tracker produces 31 facial feature points per video frame. Thus, for a 10-minute long video, one would expect to see about 18,000 sets of 31 points.

After the data has been transformed into feature points, I will attempt to normalize it by performing some rudimentary multivariate visualizations on it to determine what sorts of similarities exist between subjects. Once the data is adequately normalized, I will use it to train an HMM.

4.3 NOVEL FACIAL EXPRESSION SYNTHESIS

In order to generate novel, natural-looking expressions on the robot, appropriate "paths" need to be chosen from the HMM. In computer animation, some researchers have had success with novel facial generation using Active Appearance Models (Bettinger and Cootes, 2004). It is presently unclear whether this approach will work for physical robots, since subtle oddities in behavior can be extremely noticeable, simply due to the physicality of the robot. But this approach will be a very reasonable starting point.

4.4 EVALUATION

As stated previously, the main goal of this research is to see if by providing a robot with natural, human-like expressions people are more at ease during their interaction with it, and are thus more likely to accept such technology in domestic settings. This evaluation will center around the robot's expression generation capability, measured quantitatively and qualitatively using the techniques described below.

4.4.1 ROUND-TRIP SYNTHESIS

In the field of machine translation (MT) there is a technique known as "Round-trip translation". The idea is that one takes a sentence, uses an MT engine to translate it into a foreign language, and then uses the same system to translate the sentence back into the original language. For experts in the MT community this is not a well-regarded technique, but for non-experts it is a useful way to get a rough idea of what an MT engine can do.

For robot expressions, "round-trip synthesis" may actually prove to be a fruitful first evaluative step. For robotic expressions, the requirements for precision are less stringent than in the case of MT. And further, the circumstances are different: the robot is trained on human expressions, but is then making its own kind of expressions on a non-human face.

The evaluation would work as follows: after the robot has been trained to generate its expressions, it will be filmed making them. Then, the films of the robot will be run through the face tracker. Next, the HMM used to train the robot for synthesis will be instead used for recognition. Finally, the recognition rates would be examined for significance.

4.4.2 Human Labeling

In affective computing, a common evaluative technique involves humans labeling emotional corpora, and then inter-rater reliability is assessed as the primary metric for label accuracy. Afzal and Robinson (2008) describe an interface and methods to perform such annotations on natural data sets; I will likely use this interface to perform human evaluations of the robot's expressions.

4.4.3 HRI Approaches

Another possible avenue of evaluation is to employ evaluative techniques commonly used in HRI, such as common ground analysis, embodiment analysis, questionnaires, interviews, and other techniques. These techniques can be used to both survey the affective states of users and evaluate the affect generation capabilities of the robot. This type of evaluation can be performed "live", i.e., the user is placed in front of the robot, or else via video playback (the user watches pre-recorded videos of the robot).

4.4.4 Physiological and Neurological Approaches

It may be possible to use physiological and neurological approaches to assess how a user responds to (and possibly mimics) the robot's facial displays. There is a precedent for this approach in the cognitive psychology and affective neuroscience literature for measuring how humans perceive and respond to the emotional expressions of others. Researchers have successfully measured people's cardiovascular activity, skin conductance, fMRI, and electromyography in response to faces.

5 Conclusion

To date I have explored the affective aspects of natural interaction with robots, from both an appearance and an empathetic interaction perspective. Soon I will be working on the synthesis of natural, human-like expressions on an android robot. The goal is to see if by providing such naturalistic expressions people feel more at ease during their interactions with the robot, and, further, if they are more willing to accept the presence of similar technology in domestic settings. The results of this work will hopefully be useful to the affective computing community.

6 Acknowledgments

I would like to thank my supervisor, Professor Peter Robinson, for his support, and the Qualcomm Research Studentship for generously funding my research.

References

Afzal, S. and Robinson, P. (2008). An interface to simplify annotation of emotional behaviour. In *Proceedings of HUMAINE Workshop on Corpora in Emotion Research (LREC)*.

Bavelas, J. B., Coates, L., and Johnson, T. (2000). Listeners as co-narrators. *J Pers Soc Psychol*, 79(6):941–952.

Bernhardt, D. and Robinson, P. (2007). Detecting affect from non-stylised body motions. In *Proc. of 2nd Int'l Conf. on Affective Computing and Intelligent Interaction (ACII)*. Springer-Verlag.

Bettinger, F. and Cootes, T. F. (2004). A model of facial behaviour. In *Proc. of IEEE Int'l Conf. on Automatic Face and Gesture Recognition*.

Breazeal, C. (2002). *Designing Sociable Robots*. MIT Press, Cambridge, MA, USA.

Breazeal, C., Takanishi, A., and Kobayashi, T. (2008). Social robots that interact with people. In Siciliano, B. and Khatib, O., editors, *Springer Handbook of Robotics*, pages 1349–1369. Springer.

Davis, M. H. (2006). Empathy. In Stets, J. E. and Turner, J. H., editors, *Handbook of the Sociology of Emotions*. Springer Press, New York.

Draude, C. (2009). Who's afraid of virtual humans? In *Proc. of the AISB Symposium on The Social Understanding of Artificial Intelligence (SSoAI)*, Edinburgh.

el Kaliouby, R. (2005). *Mind-reading Machines: the automated inference of complex mental states from video*. PhD thesis, Computer Laboratory, University of Cambridge.

Fong, T., Nourbakhsh, I., and Dautenhahn, K. (2003). A survey of socially interactive robots. *Robotics and Autonomous Systems*, 42:143–166.

Goldman, A. I. (2006). *Simulating Minds: The philosophy, psychology, and neuroscience of mindreading.* Oxford University Press, New York.

IFR (2008). World robotics 2008: Statistics, market analysis, forecasts, case studies, and profitability of robot investment. Technical report, International Federation of Robotics.

Kirby, R. and et al. (2005). Designing robots for long-term social interaction. In *Proc. of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS).*

Kozima, H., Michalowski, M., and Nakagawa, C. (2008). Keepon: A playful robot for research, therapy, and entertainment. *International Journal of Social Robotics.*

Liu, C., Rani, P., and Sarkar, N. (2006). Human-robot interaction using affective cues. In *Proc. of the IEEE Int'l Symposium on Robot and Human Interactive Communication (RO-MAN).*

Morency, L.-P., de Kok, I., and Gratch, J. (2008). Context-based recognition during human interactions: automatic feature selection and encoding dictionary. In *Proc. Int'l ACM Conf. on Multimodal Interfaces (ICMI).*

Mori, M. (1970). Bukimi no tani [the uncanny valley]. *Energy*, 7(4):33–35.

Oberman, L. M., Winkielman, P., and Ramachandran, V. S. (2007). Face to face: Blocking facial mimicry can selectively impair recognition of emotional expressions. *Soc Neurosci*, 2(3):167–178.

Picard, R. (1997). *Affective Computing.* MIT Press.

Riek, L., Paul, P., and Robinson, P. (2009a). When my robot smiles at me: Enabling human-robot rapport via real-time head gesture mimicry. *Under Review.*

Riek, L. D., Rabinowitch, T.-C., Chakrabarti, B., and Robinson, P. (2009b). Empathizing with robots: Fellow feeling along the anthropomorphic spectrum. In *Proc. of 3nd Int'l Conf. on Affective Computing and Intelligent Interaction (ACII).* Springer-Verlag.

Riek, L. D., Rabinowitch, T.-C., Chakrabarti, B., and Robinson, P. (2009c). How anthropomorphism affects empathy toward robots. In *Proc. of the ACM/IEEE Int'l Conf. on Human-Robot Interaction (HRI).*

Riek, L. D. and Robinson, P. (2008). Real-time empathy: Facial mimicry on a robot. In *Workshop on Affective Interaction in Natural Environments (AFFINE) at the Int'l ACM Conf. on Multimodal Interfaces (ICMI).*

Riek, L. D. and Robinson, P. (2009). Affective-centered design for interactive robots. In *Proc. of the AISB Symposium on Workshop on New Frontiers in Human-Robot Interaction*, Edinburgh.

Ross, D. M., Menzler, S., and Zimmermann, E. (2008). Rapid facial mimicry in orangutan play. *Biology Letters*, 4(1):27–30.

Scassellati, B. (1998). Building behaviors developmentally: A new formalism. In *1998 AAAI Spring Symposium "Integrating Robotics Research"*. AAAI.

Shibata, T., Yoshida, M., and Yamato, J. (1997). Artificial emotional creature for human-machine interaction. In *IEEE Conference on Systems, Man, and Cybernetics.*

Sidner, C. L., Lee, C., Kidd, C. D., Lesh, N., and Rich, C. (2005). Explorations in engagement for humans and robots. *Artif. Intell.*, 166(1-2).

Sobol Shikler, T. (2007). *Le ton fait la musique - Analysis of expressions in speech.* PhD thesis, University of Cambridge.

Turner, J. C. (1978). Social comparison, similarity and ingroup favouritism. In Tajfel, H., editor, *Differentiation between social groups: Studies in the social psychology of intergroup relations*, pages 235–250. Academic Press, New York.

Wada, K. and Shibata, T. (2007). Living with seal robotsits sociopsychological and physiological influences on the elderly at a care house. *IEEE T Robot*, 23(5):972–980.

Zeng, Z., Pantic, M., Roisman, G., and Huang, T. (2009). A survey of affect recognition methods: Audio, visual, and spontaneous expressions. *IEEE T Pattern Anal*, 31(1):39–58.