Perception of Robot Smiles and Dimensions for Human-Robot Interaction Design

Mike Blow, Kerstin Dautenhahn, Andrew Appleby, Chrystopher L. Nehaniv, David C. Lee

Abstract— As robots enter everyday life and start to interact with ordinary people the question of their appearance becomes increasingly important. Our perception of a robot can be strongly influenced by its facial appearance. Synthesizing relevant ideas from narrative art design, the psychology of face recognition, and recent HRI studies into robot faces, we discuss effects of the uncanny valley and the use of *iconicity* and its relationship to the self-other perceptive divide, as well as *abstractness* and *realism*, classifying existing designs along these dimensions. A new expressive HRI research robot called KASPAR is introduced and the results of a preliminary study on human perceptions of robot expressions are discussed.

I. Motivations

It is an exciting time in robotics. Personal service robots, so long the science fiction dream, are becoming reality and are for sale to general consumers. Currently their uses are limited, but capabilities are improving, costs are coming down and sales are growing. In addition robots are finding a new place in society as toys, artificial pets [20], security guards, teachers [10], tour guides [24] and in search and rescue. They are finding use in areas as diverse as autism therapy [22], space exploration and research into cognition and biological systems [23].

A. RobotCub

One such research project that we are involved in at Hertfordshire is RobotCub, a multinational European project to build a humanoid child-size robot for use in embodied cognitive development research [23]. The RobotCub consortium consists of 11 core partners from Europe with collaborators in America and Japan, and the institutions involved are each working on specific areas of the robot design, engineering, developmental psychology and human-robot interaction. The software APIs and hardware plans will be published under open-source licenses, with the aim of creating a community using a common platform for robotic and cognitive research.

B. Designing Robots for Users

A previous study of people's expectations of a robot companion indicated that a large proportion of the participants in the test were in favour of a robot companion, especially one that could communicate in a human-like way [6]. Human-like behaviour and appearance were also considered important, but less so than human-like communication. In terms of role

robots were seen by the majority as suitable for personal assistant duties carrying out household tasks. Child care or friendship roles were seen as less suitable.

Existing human-human interaction studies are a good starting point for HRI research, but can only be treated as such. Robots are not people, and not all insights and results will remain valid for HRI scenarios. So given that the nature of the interaction between humans and robots is likely to be different from that between two humans, or between humans and most current consumer technology, there are many open questions. Most importantly for the general acceptance of robots, what appearance and modalities of communication are optimal for the majority of non-technical users? Will people find a machine with a human appearance or that interacts in a human-like manner engaging or frightening? If a face is humanoid, what level of realism is optimal? What role could timing in communication [25] and the movement and timing of interactive behaviour (*kinesics* [21], [1]) play?

II. Considering Design

A. The Extended Uncanny Valley

The effect of the aesthetic design of a robot is an area that has often been neglected, and only in visual science fiction media or recently with the advent of commercial household robots has it been paid much attention. A notable exception is the 'uncanny valley' proposed by Masahiro Mori in the late 1970's [17], [5]. Mori proposed that the acceptance of a humanoid robot increases as realism increases. However there comes a point where, as the robot approaches perfect realism, the effect becomes instead very disturbing and acceptance plunges, because the robot starts to look not quite human or at worst like a moving corpse (Fig. 1). In theory the realism of both appearance and movement can give rise to this effect, with movement evoking the stronger response. It is possible that there may also be 'behavioural uncanniness' affecting perception of a robot during social interaction and governed by (among other things) the appropriateness and timing of its reponses to social cues. However little empirical data exists to support Mori's theory and opinions vary as to the strength of the effect and its longevity - our initial observations with KASPAR (section III) indicated that people soon became habituated to the robot and that feelings of uncanniness decreased rapidly with time and experience. See [15], [14] for recent work on the uncanny valley by MacDorman.
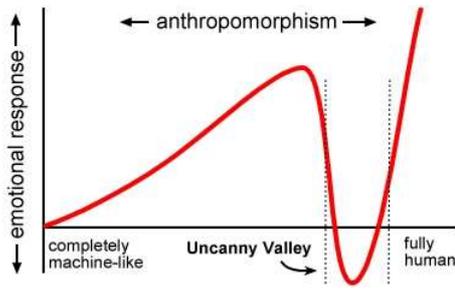
Fig. 1.   Mori's uncanny valley hypothesis.

Fig. 2.   The design space of faces in comics and narrative art (modified from [16]).

B. Managing Perceptions

DiSalvo *et al.* performed a study into how facial features and dimensions affect the perception of robot heads as human-like [7]. Factors that increased the perceived human-ness of a robot head were a 'portrait' aspect ratio (i.e. the head is taller than it is wide), the presence of multiple facial features and specifically the presence of nose, mouth and eyelids. Heads with a landscape aspect ratio and minimal features were seen as robotic. They suggest that robot head design should balance three considerations: 'human-ness' (for intuitive social interaction), 'robot-ness' (to manage expectations of the robot's cognitive abilities) and 'product-ness' (so the human sees the robot as an appliance). The idea of designing a robot to be perceived as a consumer item is noteworthy for the fact that people's *a priori* knowledge of electronic appliances can be utilised in avoiding the uncanny valley; the implication is that the robot is non-threatening and under the user's control. To fulfill their design criteria they present six suggestions: a robot should have a wide head, features that dominate the face, detailed eyes, four or more features, skin or some kind of covering and an organic, curved form.

C. Faces

Faces help humans to communicate, regulate interaction, display (or betray) our emotions, elicit protective instincts, attract others, and give clues about our health. Several studies have been carried out into the attractiveness of human faces, suggesting that symmetry, youthfulness and skin condition [9] are all factors. Famously Langlois and Roggman [12] proposed that an average face - that is, a composite face made up of the arithmetic mean of several individuals' features - is fundamentally amd maximally attractive (although there are claims to the contrary, see [19]), and that attractiveness has a social effect on the way we judge and treat others [11].

Human infants seem to have a preference for faces, and it appears that even newborns possess an 'innate' ability to spot basic facial features, such as a pair of round blobs situated over a horizontal line which is characteristic of two eyes located above a mouth. It has been debated whether this is due to special face recognition capability or sensory-based preference based on preferences for general perceptual features and broad visual cues and properties of figures such as symmetry, rounded contours etc. which form the basis
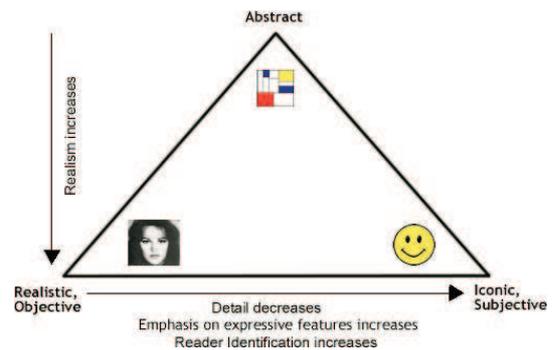
for learning to recognize faces [8]. The nature and devel-opment of face recognition in humans is still controversial. Interestingly, while the baby develops, its preference for certain perceptual features changes until a system develops that allows it to rapidly recognize familar human faces. Evidence suggests that exposure to faces in the first few years of life provides the necessary input to the developing face recognition system, e.g. [18]. The specific nature of the face stimuli during the first year of life appears to impact the development of the face processing system. While young infants (up to about 6 months of age) can discriminate among a variety of faces belonging to different species or races, children at around 9 months (and likewise adults) demonstrate a face-representation system that has become more restricted to familiar faces. The social environment, i.e. the 'kinds of faces' an infant is exposed to influences the child's preferences for certain faces and abilities to discriminate among them. Not only time of exposure, but also other factors, including emotional saliency, are likely to influence the tuning of the face recognition systems towards more precision [18].

D. The Design Space of Faces

In his book *Understanding Comics* [16], Scott McCloud introduces a triangular design space for cartoon faces (Fig. 2). The left apex is *realistic*, i.e. a perfect representation of reality, for example a photograph, or realistic art such as that by Ingres. Travelling to the right faces become more *iconic*, that is, the details of the face are stripped away to emphasise the expressive features; emoticons such as **:)** are a perfect example in the 21st century zeitgeist. The simplification has two effects. Firstly it allows us to amplify the meaning of the face, and to concentrate on the message rather than the medium. Secondly the more *iconic* a face appears the more people it can represent. Dautenhahn points out that iconography can aid the believability of a cartoon character [4]. We are more likely to identify with Charlie Brown than we are with Marilyn Monroe, as a realistic or known face can only represent a limited set of people whereas the iconic representation has a much broader range - to the extent of allowing us to project some of ourselves onto the character. Towards the top apex representations
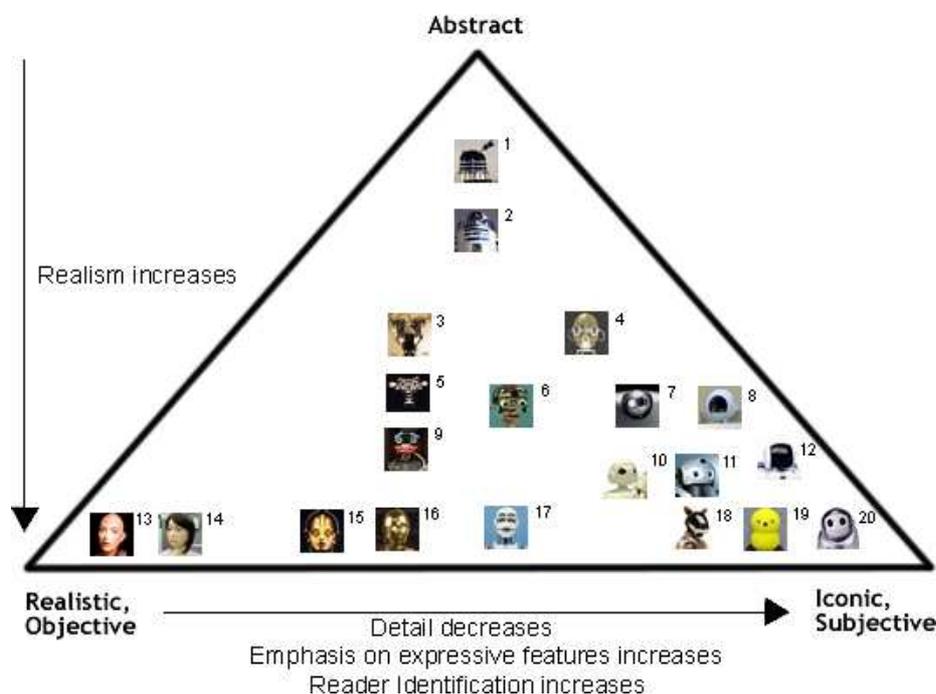
Fig. 3. Robot faces mapped into McCloud's design space. 1. Dalek ((©the British Broadcasting Corporation/Terry Nation), 2. R2D2, fictional robot from *Star Wars* (©Lucas Film Ltd.), 3. DB (©ATR Institute Kyoto), 4. MIT Humanoid Face Project (©MIT), 5. Kismet (©MIT/Cynthia Breazeal), 6. Infanoid (©Hideki Kozima), 7. Nuvo companion robot (©ZMP Inc.), 8. HOAP-2 (©Fujitsu Automation), 9. Minerva tour-guide robot (©Carnegie Mellon University), 10. Toshiba partner robot (©Toshiba), 11. QRIO (©Sony), 12. ASIMO (©Honda), 13. K-Bot, extremely realistic 24 DOF head built by David Hanson (©Human Emulation Robotics), 14. Repliee-Q1 (©Osaka University/Kokoro Inc.), 15. False Maria, fictional robot from Fritz Lang's 1927 film *Metropolis*, 16. C3PO, fictional robot from *Star Wars* (©Lucas Film Ltd.), 17. WE-4R robot (©WASEDA University), 18. AIBO robotic dog (©Sony), 19. Keepon, minimal DOF HRI robot (©Hideki Kozima), 20. Papero household robot (©NEC)

become *abstract*, where the focus of attention moves from the meaning of the representation to the representation itself. Examples in art would be (to a degree) Picasso's cubist portraits or the art of Mondrian.

E. Robot Faces in the Design Space

We can use this design space, and the accumulated knowledge of comics artists, to inform the appearance of our robots. Fig. 3 shows some robot faces and their (subjective) places on the design triangle. Most are 'real-life' robots although several fictional robots have been included, as functionality has no bearing on our classification in this context. At the three extremes are NEC's Papero (iconic), a small companion robot which is relatively simple and cheap to make and allows easy user-identification; Hanson's K-bot (realistic), complex and theoretically deep in the uncanny valley but allowing a large amount of expressive feedback, and a Dalek (abstract), potentially difficult to identify with but not as susceptible to the uncanny valley due to its non-human appearance.

Of course the design space only addresses the static appearance of the robot. The nature of most robot faces is that they encompass a set of temporal behaviours which greatly affect our perception of them. An extension of McCloud's design space to investigate behavioural aspects would be a worthwhile study, specifically how a robot's behaviour affects its perception as iconic, realistic or abstract, and the effect of social behaviour on the uncanny valley and user identification with the robot.

F. The Robot as an Extension of Self?

As one moves in the design space of the faces from realism towards iconicity, a human is more likely to identify themselves with the face due to the decrease in specific features, and the distinction between *other* and *self* becomes less and less pronounced. Could this idea be useful in robot design? If a robot is to be designed to extend the human's abilities or carry out tasks on their behalf, iconic features may possibly allow the user to more easily project their own identity onto the robot. In contrast, realistic face designs will be seen objectively as 'someone else', and *abstract* designs often as 'some*thing* else'. In this case the interaction partner's identification with the robot will be discouraged by the non-iconic nature of the design. Some robot roles (such as security guards) might benefit from reinforcing this perception. While the idea of the robot as an extension of self remains speculative at this point, future work in this area needs to shed more light on these issues.

III. KASPAR

Fig. 4 shows KASPAR (**K**inesics **A**nd **S**ynchronisation in **P**ersonal **A**ssistant **R**obotics). KASPAR is a child-sized robot which acts as a platform for HRI studies, using mainly expressions and gestures to interact with a human. The robot is a work-in-progress but when finished will comprise a static body with an 8 DOF head and two 6 DOF arms. Important features of KASPAR are minimal design, the inclusion of eyelids, and aesthetic consistency of the face (which is why eyebrows were not implemented; any mechanism to actuate them would have protruded through the skin).
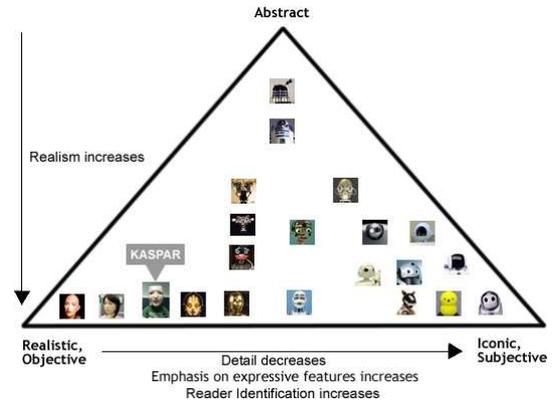
Fig. 4. KASPAR, HRI research robot.

Fig. 5. KASPAR on the design space of robots.

A. Design Motivations and Rationale

Part of Hertfordshire's input in the early stages of the RobotCub project was to suggest design motivations that would help produce a useful platform for HRI studies, and which also formed the basis of the design rationale for KASPAR. These were that there should be consistency of appearance and complexity between the head, body and hands to aid natural interaction, and also between the appearance and the capabilities of the robot to govern the human's expectations. It was also suggested that minimal expressive features should be included and that they should be used to create the impression of autonomy by (for example) allowing joint attention or expressing emotional state.

The overall hardware costs of KASPAR are in the range of a desktop PC, and by keeping the complexity and DOFs to a minimum we aim to reduce building and maintainance costs while still creating a robot capable of a wide range of behaviours. The goal in this case is not perfect realism, but optimal realism for rich interaction.

B. Face Design

The face design echoes the overall rationale, in that it aims to approximate the appearance and movements of the human face without venturing into ultra-realism. Fig. 5 shows the approximate position of KASPAR on the design space of robot faces. The decision to position the face somewhat in the iconic direction was made with a two-fold purpose. We have seen that emphasis on the features used for communication allows the robot to present facial feedback clearly, by allowing the interaction partner to focus on the message more than the medium. Furthermore a reduction in detail de-personalises the face and allows us to project our own ideas on it and make it, at least partially, what we want it to be. These are both potentially desirable features for a robot in HRI scenarios. Note, however, that the emphasis on the communicative features is achieved not by using discrete, exaggerated versions (which is the case with robots such as Feelix [3] and Kismet [2]), but by reducing the distracting effect of other details of the face. KASPAR's expressions are not as unambiguously defined as those of Kismet or Feelix, but initial observations indicate

that surprisingly subtle changes in expression can be effective (see experimmental results, section IV).

KASPAR's skin (a resuscitation doll mask) is only fixed at the ears and nose, and allows the face to be pulled into some fairly natural-looking expressions as the actuation of the mask in one place tends to slightly deform other areas; for instance, a smile also pushes up the cheeks and narrows the eyes. In humans this is typically considered an 'honest' smile compared to one which moves only the mouth [1].

C. Design Specifics

Requirements and Strategy. KASPAR's design was informed by initial studies of existing robot heads and by the application of ideas from McCloud's design space. The design requirements were: (1) Minimal design, yet expressive enough for HRI, (2) capacity to display *autonomy*, (3) capacity to display *undirected* and *directed attention*. (4) iconicity, (5) capacity to accept "projected" expressions with change of view angle (a requirement that was inspired by this ability in traditional Japanese noh masks [13]), and (6) human-like appearance.

Metal rods are used to transmit servo movement to the required part of the face or head. In addition to CMOS cameras in the eyes, micro-switches will be incorporated in the hands to provide simple tactile feedback and microphones added to the head.

D. Potential Uses

KASPAR can be used to study a variety of research issues relevant to HRI such as interaction dynamics, gesture creation and recognition, joint attention, communication through imitation and the use of expressions. The addition of arms will allow a range of interaction games to be played.

IV. SMILE EXPERIMENT

The first study to be undertaken with KASPAR investigated people's perception of the robot's expression. Such an experiment was considered necessary in order to provide baseline results that will inform future experiments where KASPAR's expressions will be used in regulating interaction dynamics with people. For this purpose, a simple experiment

was created to investigate what bearing the speed and continuity of a transition from one expression to another might have on the perception of a robot. Our expectations were that:

(1) Static expressions of a smile will be judged less appealing by subjects than expressions with dynamic transitions from a neutral expression.

(2) Dynamic expressions with transitions at natural speed will be judged more appealing than those with abrupt transitions.

(3) The larger the smile, the better will subjects recognize the expression of 'happiness'.

As this experiment investigates the use of movement in robot perception it can only be partially related to the idea of the design space which only concerns static images.

A. Methodology

Four degrees of smile were programmed into KASPAR and recorded on video with a plain static background. These were neutral (i.e. no smile, and the 'default' starting condition for all other expressions), and small, medium and large smiles (Fig. 6). Ten videos were created of 6 seconds duration each, showing:

1) The neutral face with no transition (static) as a control condition.

2) Small medium and large smiles with no transition (static).

3) Small, medium and large smiles with a natural transition (one that takes up to 2 seconds from neutral to smile).

4) Small, medium and large smiles with a sudden transition, created by editing the video to cut abruptly from neutral to finished smile with no intermediate stage.

It is important to note that the three sizes of the smiles remained consistent across all videos, and that only the transitions varied. A website was created which, after gathering consent and some minimal demographic data, presented all the videos twice in a random order. For each video the subject was asked to rate how happy, and how appealing, the robot's smile looked on a scale of 1-5, where 5 is maximal. Ratings of 'happiness' were expected to reflect how successful the robot's design conveyed this expression. As perceived 'happiness' could simply be interpreted as 'the amount of smile', we were also interested in how the robot would be regarded by subjects both visually and behaviourally and thus chose the term 'appealing' in an attempt to communicate the idea of this subjective judgement. All results were stored in a database for later analysis.

B. Results

Results from 51 subjects were obtained, from the UK, Norway, Sweden, Netherlands, Germany, Austria, Poland, Spain, Portugal and Italy. The subjects ranged in age between 23 and 58, 21% were female and almost all worked in a variety of academic and administration roles in universities.

1) 'Happiness' Rating: Fig. 7 shows the mean responses (average standard deviation = 0.94) to the question 'On a scale of 1-5, how HAPPY does this robot's smile look?' for each video. For the small and medium smiles, those with
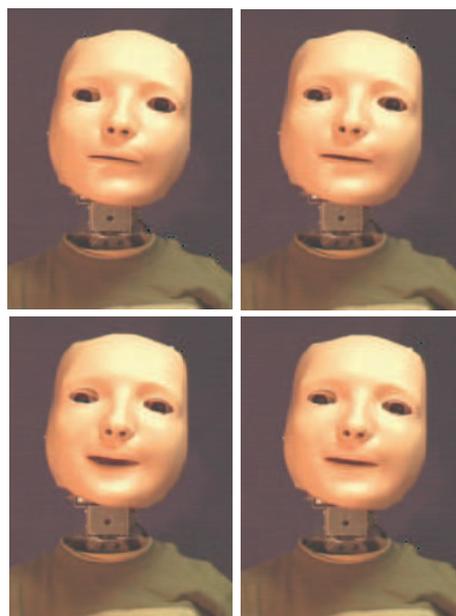

Fig. 6. The four experimental expressions, clockwise from top left: neutral, small, medium and large smiles.

transitions (4-9) are perceived as marginally happier than the corresponding static smiles (1-3). For the more obvious large smile, the static version is seen as happiest followed by the natural and sudden transition versions. It is interesting that there is such a distinct classification, especially between the small and medium smiles, as at first glance the difference between them is quite subtle.

Fig. 7. Perceived 'happiness' responses. 0 = neutral, 1 = small static, 2 = medium static, 3 = large static, 4 = small NT, 5 = medium NT, 6 = large NT, 7 = small ST, 8 = medium ST, 9 = large ST. NT = natural transition, ST = sudden transition. Average standard deviation = 0.94.

2) 'Appeal' Rating: Fig. 8 shows the mean responses (average standard deviation = 1) to the question 'On a scale of 1-5, how APPEALING does this robot's smile look?' for each video. Here the clear winners are the natural transitions (4-6). In each of the small, medium and large cases the natural transition smile is rated higher than either the corresponding static or sudden transition options. Interestingly the large smile with a natural transition (6) is the most appealing of all the large smiles (in fact the most appealing of all the expressions), and yet the large smile with a sudden transition

(9) is the least. This suggests that realism or time taken to attain an expression might be a crucial factor in how the robot is perceived by human subjects.
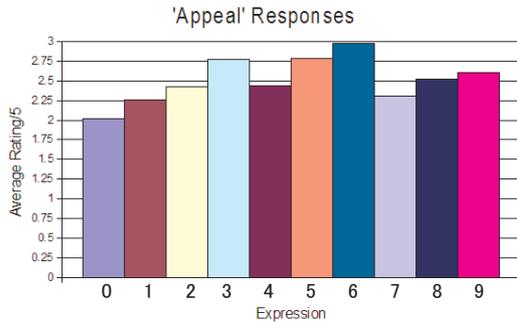

Fig. 8. Perceived 'appeal' responses. 0 = neutral, 1 = small static, 2 = medium static, 3 = large static, 4 = small NT, 5 = medium NT, 6 = large NT, 7 = small ST, 8 = medium ST, 9 = large ST. NT = natural transition, ST = sudden transition. Average standard deviation = 1.

Two of our hypotheses are supported by the results - natural transitions are seen as more appealing than sudden ones (hypothesis 2), and the larger the smile the greater the judgement of 'happiness' (hypothesis 3). However hypothesis 1 is only partially supported; smiles with a natural transition are seen as more appealing than static ones, but those with a sudden transition are not. We would suggest that in the latter case the inconsistency between appearance (fairly natural) and behaviour (unnatural) causes a negative response.

V. CONCLUSIONS

In this paper we focussed on design issues of robot faces integrating findings from psychological studies, work on narrative art design, and recent HRI studies. Consideration of these design issues strongly influenced our creation of a minimally expressive humanoid face, part of the robot KASPAR. Dimensions of face design were discussed with aims to help researchers and designers understand and exploit some ideas synthesizing those of artists, roboticists, and psychologists that pertain to human perception of robot faces in HRI. Expressions with a natural transition time are experimentally shown to be seen as more appealing than static ones or those with a sudden transition. Whether the preferred style of expression is one which has natural timing in any context or is merely one consistent with the overall aesthetic of the robot is an open question. Although these results are specific to KASPAR it is clear that robot design affects peoples' perceptions in significant ways and these results suggest that aesthetic/behavioural consistency and the temporal element in HRI are worthy of further investigation.

REFERENCES

[1] R. L. Birdwhistell. *Kinesics and Context*. University of Pennsylvania Press, Philadelphia, 1970.
[2] C. L. Breazeal. *Designing Sociable Robots*. MIT Press, 2002.
[3] L. Cañamero and J. Fredslund. I show you how I like you - Can you read it in my face? *IEEE Trans. Systems, Man & Cybernetics, Part A*, 31(5):454–459, 2001.
[4] K. Dautenhahn. Design spaces and niche spaces of believable social robots. In *Proc. IEEE Intl. Workshop Robot and Human Interactive Communication*, pages 192–197, 2002.
[5] K. Dautenhahn. Socially intelligent agents in human primate culture. In S. Payr and R. Trappl, editors, *Agent Culture: Human-Agent Interaction in a Multicultural World*, pages 45–71. Lawrence Erlbaum Associates, 2004.
[6] K. Dautenhahn, S. Woods, C. Kaouri, M. Walters, K. L. Koay, and I. Werry. What is a robot companion - friend, assistant or butler? In *Proc. IEEE IROS*, pages 1488–1493, 2005.
[7] C. DiSalvo, F. Gemperle, J. Forlizzi, and S. Kiesler. All robots are not created equal: The design and perception of humanoid robot heads. In *Proc. Designing Interactive Systems*, pages 321–326, 2002.
[8] M. J. Johnson and J. Morton. *Biology and Cognitive Development: The Case of Face Recognition*. Blackwell, 1991.
[9] B. C. Jones, A. C. Little, D. M. Burt, and D. I. Perrett. When facial attractiveness is only skin deep. *Perception*, 33(5):569 – 576, 2004.
[10] T. Kanda and H. Ishiguro. Communication robots for elementary schools. In *Proc. AISB'05 Symposium Robot Companions: Hard Problems and Open Challenges in Robot-Human Interaction*, pages 54–63, April 2005.
[11] J. Langlois, L. Kalakanis, A. Rubenstein, A. Larson, M. Hallam, and M. Smoot. Maxims or myths of beauty? A meta-analytic and theoretical review. *Psychological Bulletin*, 126:390–423, 2000.
[12] J. Langlois and L. Roggman. Attractive faces are only average. *Psychological Science*, 1:115–121, 1990.
[13] M. Lyons, R. Campbell, A. Plante, M. Coleman, M. Kamachi, and S. Akamatsu. The noh mask effect: Vertical viewpoint dependence of facial expression perception. *Proc. Royal Soc. London*, 267:2239–2245, 2000.
[14] K. F. MacDorman. Androids as an experimental apparatus: Why is there an uncanny valley and can we exploit it? In *CogSci-2005 Workshop: Toward Social Mechanisms of Android Science*, pages 106–118, 2005.
[15] K. F. MacDorman, T. Minato, M. Shimada, S. Itakura, S. Cowley, and H. Ishiguro. Assessing human likeness by eye contact in an android testbed. In *Proc. XXVII Ann. Meeting of the Cognitive Science Society*, 2005.
[16] S. McCloud. *Understanding Comics: The Invisible Art*. Harper Collins Publishers, Inc., 1993.
[17] M. Mori. Bukimi no tani [the uncanny valley]. *Energy*, 7:33–35, 1970.
[18] O. Pascalis, L. S. Scott, D. J. Kelly, R. W. Shannon, E. Nicholson, M. Coleman, and C. A. Nelson. Plasticity of face processing in infancy. *PNAS*, 102(14):5297–5300, 2005.
[19] D. Perrett, K. May, and S. Yoshikawa. Attractive characteristics of female faces: preference for non-average shape. *Nature*, 368:239–242, 1994.
[20] J. Pransky. AIBO - the No. 1 selling service robot. *Industrial Robot*, 28(1):24–26, 2001.
[21] B. Robins, K. Dautenhahn, C. L. Nehaniv, N. A. Mirza, D. François, and L. Olsson. Sustaining interaction dynamics and engagement in dyadic child-robot interaction kinesics: Lessons learnt from an exploratory study. In *Proc. 14th IEEE Ro-Man*, 2005.
[22] B. Robins, K. Dautenhahn, R. te Boekhorst, and A. Billard. Effects of repeated exposure to a humanoid robot on children with autism. In *Proc. Universal Access and Assistive Technology (CWUAAT)*, pages 225–236, 2004.
[23] G. Sandini, G. Metta, and D. Vernon. Robotcub: An open framework for research in embodied cognition. In *Proc. IEEE-RAS/RSJ Intl. Conf. Humanoid Robots*, 2004.
[24] S. Thrun, M. Bennewitz, W. Burgard, A. Cremers, F. Dellaert, D. Fox, D. Haehnel, C. Rosenberg, N. Roy, J. Schulte, and D. Schulz. Minerva: A second generation mobile tour-guide robot. In *Proc. IEEE Intl. Conf. Robotics and Automation (ICRA'99)*, 1999.
[25] M. Yamamoto and T. Watanabe. Time lag effects of utterance to communicative actions on robot-human greeting interaction. In *Proc. IEEE Intl. Workshop Robot and Human Interactive Communication*, 2003.