Exploring the Aesthetic Range for Humanoid Robots

David Hanson (david.hanson@utdallas.edu)
The University of Texas at Dallas PO Box 830688
Richardson TX 75083-0688

Abstract

Although the uncanny exists, the inherent, unavoidable dip (or valley) may be an illusion. Extremely abstract robots can be uncanny if the aesthetic is off, as can cosmetically atypical humans. Thus, the uncanny occupies a continuum ranging from the abstract to the real, although norms of acceptability may narrow as one approaches human likeness. However, if the aesthetic is right, any level of realism or abstraction can be appealing. If so, then avoiding or creating an uncanny effect just depends on the quality of the aesthetic design, regardless of the level of realism. The author's preliminary experiments on human reaction to near-realistic androids appear to support this hypothesis.

Introduction

While Masahiro Mori's (1970) uncanny valley paradigm has dominated robotics design for over 30 years, there has been little formal challenge to the paradigm's premises. Is there really a strong, immutable relationship between the human realism and acceptability of robots? Is there an unavoidable discontinuity of acceptability (i.e., a valley) between abstract and highly realistic anthropomorphic depictions?

To answer this question, *human realism* needs to be clearly defined. We define realism as "being within the possible, naturally-occurring appearance of real human beings." Realism then can be considered across several dimensions including static and dynamic appearance and contextual responsiveness (i.e., contingent interaction). Within each dimension, there are many sub-characteristics of realism, such as physical geometry, texture, and coloration, which will be constrained by human biology.

A humanoid figure may exhibit extreme realism in some characteristics while deviating from realism in others (e.g., a realistic face with a cartoon body). Alternately, the characteristics of a figure may evenly deviate from realism (e.g., a face and a body that are both slightly cartoonish). With so many ways to deviate from realism, and so many ways to modulate the aesthetic, it would seem plausible that human reaction could vary at any given level of realism.

If human reaction is indeed variable at any given level realism, this implies that the aesthetic space is more densely populated, more like a cloud of aesthetic possibilities rather than the definite curve drawn in Mori's uncanny valley graph (Mori, 1970). In (Hanson et al., 2005), anecdotal examples indicated that there can be indeterminately many possibilities for aesthetic humanlike depictions that lie outside the curve of Mori's valley. This implies that human reactions to an anthropomorphic depiction are more strongly related to good or bad design than to its level of human realism.

This paper describes a series of preliminary tests that attempt to map out human reaction to robots that are nearly human-looking in appearance. The results of these tests appear to contravene the uncanny valley hypothesis. An alternative to the uncanny valley paradigm is then proposed.

Background

In recent years, neuroscientists and evolutionary psychologists have found abundant evidence that our tastes of beauty and ugliness are stamped into our nervous system (Rhodes and Zebrowitz, 2002), shaped by evolutionary pressures into universal, neural-templates that filter distinctly for beauty (Etcoff, 2000; Cunningham et al., 2002), for ill health and danger (Darwin and Ekman, 1872/1998; La Bar et al., 2003; Etcoff, 2000; Kesler-West et al., 2001), and for "things we are or are not accustomed to" (Dion, 2002). These neural-templates represent a primary obstacle course for social robot designers. Any "uncanny" perceptual phenomenon depends on these neural systems.

While studies indicate that we are much more sensitive to real human faces (Gauthier et al., 2000), the specific forms of beauty and ugliness inspire remarkably consistent human responses, regardless of their level of realism (Etcoff, 2000; Zaidel, 1997; Thomas and Johnston, 1995). The scientific literature on facial attractiveness shows that even among real humans, minor deviations in appearance can change a face from beautiful to ugly or disturbing (Etcoff, 2001; Cunningham et al., 2002).

Universally, clear skin, well-groomed hair and large expressive features are considered attractive (Etcoff, 2000; Cunningham et al., 2002). Likewise, the large eyes and forehead, and small nose and jaw associated with neoteny (the "baby scheme") are universally considered endearing and inspiring of protection (Eibl-Eibesfeldt, 1970; Etcoff, 2000; Cunningham et al., 2002; Breazeal, 2002). In general, averaged faces are more attractive than the median (presumably by canceling unhealthy deviations from the norm) (Rubenstein et al., 2002; Rhodes et al., 2002). However, average faces are not the most attractive. The most attractive faces deviate from the average, but only in very specific ways, usually in features associated with neoteny, sexual maturity, or senescence (Cunningham et al., 2002; Etcoff, 2000). Each of these exaggerated feature-sets inspires different behavior in humans. Neoteny features inspire nurturing, sexual maturity features inspire both sexual attraction and friendship, while senescence features inspire mentoring relationships (Cunningham et al., 2002; Zebrowitz and Rhodes, 2002).

It is well demonstrated that human aesthetic preferences transfer to nonhuman objects and beings (Norman, 1992; Kanwisher, 1997; Breazeal, 2002; Fong et al., 2003).

Conversely, other aesthetic patterns are universally regarded as ugly, disturbing, or eerie. Sickly eyes, bad skin, extreme asymmetry, and poor grooming are all repulsive to people (Etcoff, 2000). Generally, signs of illness or injury are found to be disturbing (Darwin and Ekman, 1872/1998; Etcoff, 2000). Facial forms akin to expressions of terror, psychosis, and subterfuge are also universally found to be alarming (Darwin and Ekman, 1872/1998; Ekman, 1970; Adolphs et al., 2001). These kinds of eerie signifiers are used in cartoons and art to depict villains or monsters. Such negative features would certainly be associated with a "walking corpse"—Mori's example at the bottom of the purported uncanny valley. But as discussed, these features are not attached to a given level of realism, any more than a big smile is, or large cute eyes are. Avoiding perceptual templates that trigger fear may help avoid the uncanny reaction, regardless of the level of realism.

But what about sensitivity to realism—is there any evidence that realism does make a difference? Studies do show that people are especially sensitive to the real human face (Tzourio-Mazoyer, 2002; Kanwisher, 1997; Kanwisher, 2000). We are much more sensitive to familiar faces and objects (Gauthier, 1998) and can more easily recognize such faces (Golby, 2001). People appear to find more familiar types of faces to be more attractive (Reiman et al., 2000; Etcoff, 2000; Cunningham et al., 2002). People are especially sensitive to subtleties of real human faces—moving one facial feature by just 1mm will change a real face from attractive to unattractive (Etcoff, 2000, p. 134). These sensitivities imply that more realistic faces trigger more demanding expectations for anthropomorphic depictions (Hanson et al., 2005).

Sending Robots in to Explore the Valley

Social robotics research has reapplied techniques of animatronics (entertainment robotics) in AI-driven robots, with notable examples including Cynthia Breazeal's collaboration with Stan Winston on the Leonardo robot, and Hiroshi Ishiguro's work with Kokoro Co., Ltd. on robots including the Repliee Q1. The author's robots continue this trend, being realistic in expression and interactive, but differ in that always some features put them in the region of the uncanny valley—for example, the back of the head is missing from the Philip K. Dick robot, exposing wires and mechanisms (see Fig. 1). These robots are intended to plumb the uncanny valley and challenge the premises of the paradigm.

Fig. 1. Philip K. Dick Android.

In addition to animated appearance, the author's robots engage with conversational speech, via AI-driven intelligent software using face tracking, face recognition, automatic speech recognition, and speech synthesis.

In November 2004, the author led two informal web-surveys that showed videos of two Hanson robots, animated with humanlike facial expressions. Reactions to each robot were similar: more than 80% of respondents found the stimuli "entertaining," 73% found the stimuli "appealing," and over 85% found the robots to look "lively" and "not dead."

From June 2005 to November 2005, the Philip K Dick Android (with the back of the head missing) was shown in 3 public exhibitions, where people's behavior was observed and noted by curators. Following their interactions with the robot, people were given exit interviews. According to the observers, people who interacted with the robot appeared entertained, not disturbed or afraid. The robot held peoples' attention in conversation for many minutes and even hours. People held the android's hand while talking with it, and even spontaneously hugged the android at the end of the conversation. In the exit interviews, 71% said the robot was "not eerie," and 89% "enjoyed" interacting with the robot. These results seem to merit more formal experiments.

Experiments and Results

To further test the uncanny valley theory, in October of 2005 we administered a new series of assays wherein we showed human participants series of images of the Philip K. Dick android (PKD-A), Qrio, and humanlike images, with varying levels of realism, and varying aesthetic qualities.

The test consisted of a series of images that morph from abstract robots, to our realistic robots, to images of the human models on which the robots were based. The human participants were asked to rank the images from 1 to 10, on several metrics: realism, appeal, eeriness, and familiarity.

Method

Participants. There were 25 participants, ranging in age from 18 to 77. The national origin of the participants was diverse, with 12 U.S. nationals, 8 south Asian, and 6 other. 12 Participants were male and 13 female. Participants were recruited in public thoroughfares on two college campus of the University of Texas at Dallas. All participants were volunteers and none received remuneration.

Procedure. In these experiments our control morph (see Fig. 2) was inspired by a morph used in the experimental work of Karl MacDorman, using the robots of Hiroshi Ishiguro, which contains a continuum of morphed images that elicit reactions from participants which follow the pattern predicted by the uncanny valley.

In our experimental morph, meanwhile, the morph images were designed with the intention of making them appealing and not eerie (see Fig. 3). If human participants reacted with consistently low-eeriness ratings, this would imply that the uncanny valley is avoidable, at least in the static domain.

Reaction to the control figures followed the pattern predicted by the uncanny valley theory (see Fig. 2). Reactions to the tuned morph, however, were striking in that the attractively-tuned figures were found to be consistently low in eeriness and high in appeal. This strongly implies that reaction is at least partially decoupled from realism. These results also imply that, with well-tuned faces, there can exist a continuum of appealing anthropomorphism across the range of realism, thus supporting the hypothesis of no inherent uncanny valley.

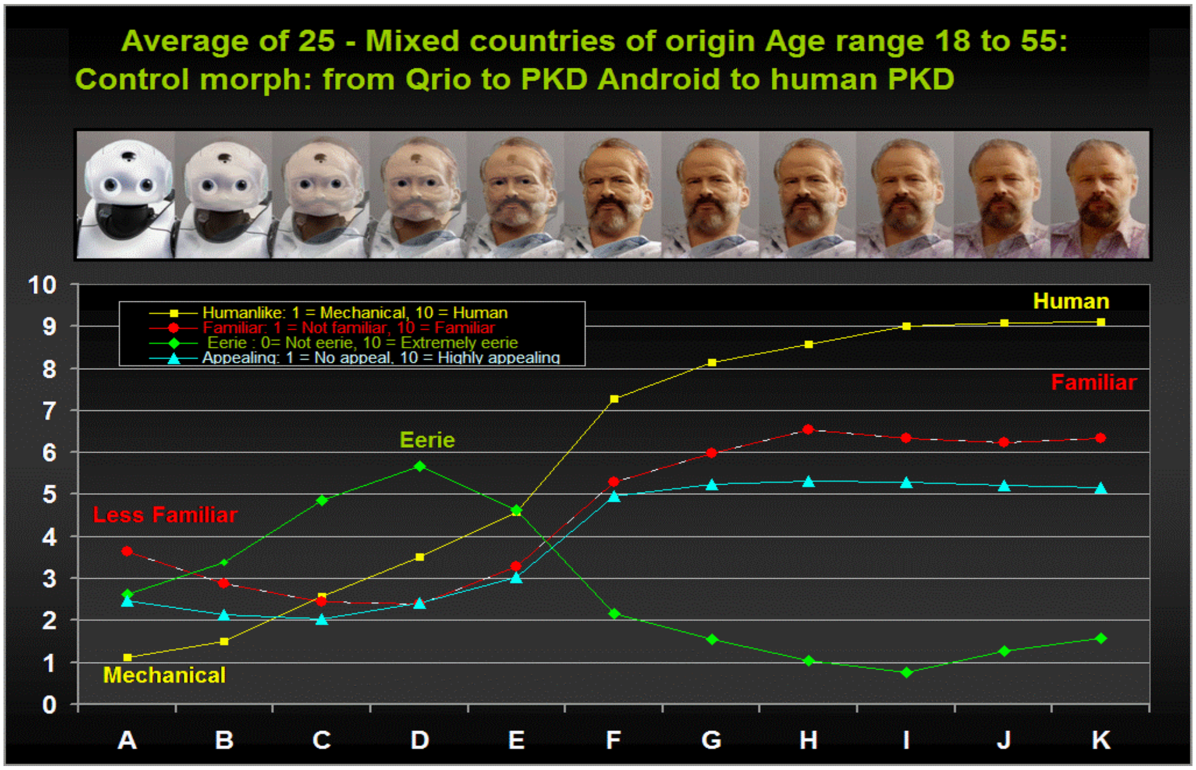
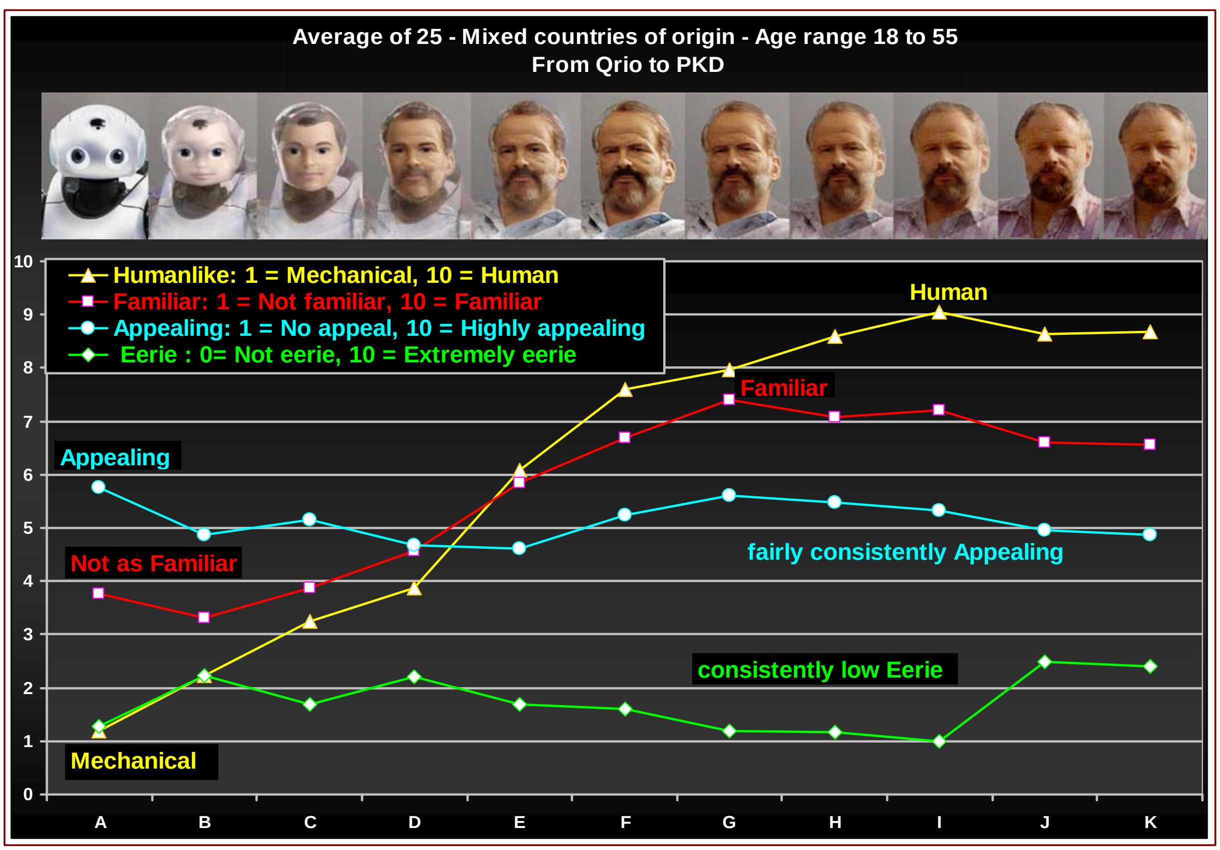


Fig. 2 (above). Uncanny morph: Qrio—android—PKD. Fig. 3 (below). Not Uncanny morph: Qrio—android—PKD.



Thus the data results are not consistent with the uncanny valley hypothesis. Further experiments are merited. Future experiments should be animated, interactive, and with participants in the presence of actual robots. A range of realism should be produced in the robots.

Pursuing New Theory for Robot Design

We propose a preliminary a replacement paradigm for the Uncanny Valley. If the illusion of life can be created and maintained, the uncanny effects may be mitigated. It may be that any level of realism can be socially engaging if one designs the aesthetic well. This, in effect, would represent a bridge of good aesthetic, which inspires us to name the revised theory the *path of engagement* (POE).

Conclusions and Future Work

As robots proliferate, they will more frequently engage people in face-to-face interactions. The success of such encounters will depend substantially upon the aesthetics of a given robot. Identification of fundamental principles of robot aesthetics can greatly accelerate the successful deployment of robots.

Presently even the most realistic robots may seem partly-dead, because in many ways they are. They are only partly aware. They shut down instead of going to sleep, and then they sit there frozen. They break. These flaws in a humanlike appearance, can remind us of our own mortality. They also may imply dead matter impersonating humans, conveying the threat of an imposter. But, if we remove these flaws to make them friendly, attractive, and seemingly alive, then the level of realism may not matter.

Ultimately, good design can help to make robots lovable and part of the human family. More freely exploring the full range of robot aesthetics will certainly accelerate the evolution of humanoid robot design. Moreover, the expanded exploration promises to help us better understand human social perception, interaction, and cognition.

Acknowledgments

The author would like to acknowledge the help and guidance of Karl MacDorman, Alice Otoole, Thomas Linehan, and Dennis Kratz. The author especially thanks Amanda and Elaine Hanson for their support and assistance.

References

Adolphs, R. (2001). The neurobiology of social cognition. *Current Opinion in Neurobiology, l1*, 231-239.

Breazeal C. (2002). *Designing Sociable Robot.*, Cambridge, Mass.: MIT Press.

Cunningham, M.R., Barbee A.P., & Philhower C. (2002). Dimensions of facial physical attractiveness: The intersection of biology and culture, in *Facial Attractiveness: Evolutionary, Cognitive, and Social Perspectives*. Westport, Conn.: Ablex Publishing.

Darwin, C. & Ekman, P. (Ed.). (1998/1872). *The expression of the emotions in man and animals.* New York: Oxford University Press.

Dion K.K. (2002). Cultural perspectives on facial attractiveness, in *Facial attractiveness: evolutionary, cognitive, and social perspectives.* Ablex Publishing.

Eibl-Eibesfeldt, I. (1970). *Ethology: The biology of behavior.* New York: Holt, Rinehart and Winston.

Ekman, P. & Friesen, W. V. (1971). Constants across cultures in the face and emotion. *Journal of Personality and Social Psychology, 17,* 124-129.

Etcoff, N. (2000). *Survival of the prettiest*. New York: Anchor.

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

Gauthier, I., Skudlarski, P., Gore, J.C., & Anderson, A.W. (2000). Expertise for cars and birds recruits brain areas involved in face recognition. *Nature Neuroscience, 3*(2), 191-197.

Gauthier, I., Williams, P., Tarr, M. J., & Tanaka, J. (1998). Training "Greeble" experts: A framework for studying expert object recognition processes. *Vision Research,* special issue on Models of Recognition, *38,* 2401-2428.

Golby, A. J., Gabrieli, J. D. E., Chiao, J. Y. & Eberhardt, J. L. (2001). Differential responses in the fusiform region to same-race and other-race faces. *Nature Neuroscience, 4,* 845-850.

Hanson, D., Olney, A., Zielke, M., Pereira, A. (2005). Upending the uncanny valley, in *AAAI conference proceedings.*

Kanwisher, N., McDermott, J., & Chun, M. M. (1997). The fusiform face area: A module in human extrastriate cortex specialized for face perception. *Journal of Neuroscience, 17,* 4302-4311.

Kesler-West, M.L., Andersen, A.H., Smith C.D., Avison, M.J., Davis, C.E., Kryscio, R.J., Blonder, L.X. (2001). Neural substrates of facial emotion processing using fMRI. *Cognitive Brain Research, 11,* 213-226.

Mori, Masahiro (1970). Bukimi no tani (the uncanny valley). *Energy, 7,* 33–35. (In Japanese).

Norman, D. (1992). *Turn signals are the facial expressions of automobiles*. Cambridge, Mass.: Perseus Publishing,..

Rhodes, G., & Zebrowitz, L.A. (2002). *Facial attractiveness: Evolutionary, cognitive, and social perspectives.* Westport, Conn.: Ablex Publishing.

Rubenstein A.J., Langlois J.H., Roggman L.A. (2002). What makes a face attractive and why: The role of averageness in defining facial beauty, in *Facial attractiveness: Evolutionary, cognitive, and social perspectives.* Westport, Conn.: Ablex Publishing.

Thomas, F. & Johnston, O. (1995). *The illusion of life: Disney animation* (Rev. ed.). Hyperion.

Tzourio-Mazoyer, N., De Schonen, S., Crivello, F., Reutter, B., Aujard, Y., Mazoyer, B. (2002). Neural correlates of woman face processing by 2-month-old infants. *NeuroImage, 15,* 454-461

Zaidel, D. W. (1997).Art and science. *Nature, 390,* 330.