

## Is The Uncanny Valley An Uncanny Cliff?

Christoph Bartneck<sup>1</sup>, Takayuki Kanda<sup>2</sup>, Member, IEEE, Hiroshi Ishiguro<sup>2</sup>, Member, IEEE  
and Norihiro Hagita<sup>2</sup>, Member, IEEE

<sup>1</sup> Department of Industrial Design, Eindhoven University of Technology, Den Dolech 2, 5600MB Eindhoven, The Netherlands, email: c.bartneck@tue.nl

<sup>2</sup>ATR Intelligent Robotics and Communications Labs, 2-2-2 Hikaridai Seikacho Sorakugun, Kyoto, Japan, email: [kanda,ishiguro,hagita]@atr.jp

**Abstract**— The uncanny valley theory proposed by Mori in 1970 has been a hot topic in human robot interaction research, in particular since the development of increasingly human-like androids and computer graphics. In this paper we describe an empirical study that attempts to plot Mori's hypothesized curve. In addition, the influence of framing on the users' perception of the stimuli was investigated. Framing had no significant influence on the measurements. The pictures of robots and humans were rated independently of whether the participants knew a particular picture showed a robot or human. Anthropomorphism had a significant influence on the measurements, but not even pictures of real humans were rated as likeable as the pictures of humanoid or toy robots. As a result we suggest the existence of an uncanny cliff model as an alternative to the uncanny valley model. However, this study focused on the perception of pictures of robots and the results, including the suggested model, may be different for the perception of movies of moving robots or the perception of standing right in front of a moving robot.

### I. INTRODUCTION

The uncanny valley theory was proposed originally by Masahiro Mori [1] and was discussed recently at the Humanoids-2005 Workshop [2]. It hypothesizes that the more human-like robots become in appearance and motion, the more positive the humans' emotional reactions towards them become. This trend continues until a certain point is reached beyond which the emotional responses quickly become negative. As the appearance and motion become indistinguishable from humans the emotional reactions also become similar to the ones towards real humans. When the

emotional reaction is plotted against the robots' level of anthropomorphism, a negative valley becomes visible (see Figure 1) and is commonly referred to as the uncanny valley. Movement of the robot amplifies the emotional response in comparison to static robots.

Ever since it was first proposed, the uncanny valley has been a hot topic in human robot interaction research. However, the amount of empirical proof has been rather scarce, as Blow, Dautenhahn, Appleby, Nehaniv, & Lee [3] observed. With the arrival of highly realistic androids (see Figure 2) and computer-generated movies, such as, "Final Fantasy", and "The Polar Express", the topic has grabbed much public attention. The computer animation company Pixar developed a winning/clever strategy by focusing on non-human characters in its initial movie offerings/projects, e.g., toys in, "Toy Story", and insects/bugs in, "It's a Bug's Life". In contrast, the annual "Miss Digital World" beauty competition [4, 5] has failed to attract the same level of interest.

A possible explanation of the uncanny phenomenon may be related to the framing theory [6]. When we encounter new situations or artifacts, we select from our memory a structure called a frame. Frames are data structures for representing stereotyped situations or artifacts. When we enter a restaurant, for example, we already have certain expectations. Attached to each frame are several kinds of information which help us in knowing how to use the frame, anticipating what will happen next and also knowing what to do when our expectations are not fulfilled.

When we encounter a very machine-like robot, we select a 'machine frame' and its human-like features deviate from our expectation and hence attract our attention. This deviation is usually positive since we tend to like other humans. In contrast, when we encounter an android, we select our 'human frame' and its machine-like features grab our attention. However, the machine-like features are deviations that are otherwise only found in sick or injured people, which we find to be disturbing [7]. This study attempts to identify how strong this framing effect might be.

In his original paper, Mori plots human likeness against 親和感 (shinwa-kan), which has previously been translated to "familiarity". Familiarity depends on previous experiences and is therefore likely to change over time. Once people have been exposed to robots, they become familiar with them and the robot-associated eeriness may be eliminated [3]. To that end, the uncanny valley model may only represent a short phase and hence might not deserve the at-

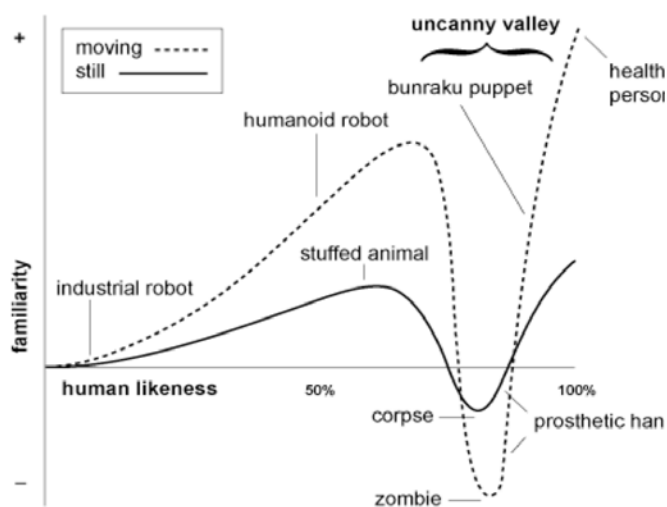


Figure 1: The uncanny valley (source: Wikipedia. Karl MacDorman translated this graph from Mori's original article).

tention it is receiving. We also questioned whether Mori's shinwa-kan concept might have been "lost in translation", and in consultation with several Japanese linguists, we discovered that shinwa-kan is not a commonly used word, nor does it have a direct equivalent in English. In fact, "familiarity" appeared to be the least suitable translation compared to "affinity" and in particular to "likeability".

It is widely accepted that given a choice, people like familiar options because these options are known and thereby safe, compared to an unknown and thereby uncertain option. Even though people prefer the known option over the unknown option, this does not mean that they will like *all* the options they know. Even though people might prefer to work with a robot they know compared with a robot they do not know, they will not automatically like all the robots they know. Therefore, the more important concept is likeability, and not familiarity.

Several studies have started empirical testing of the uncanny valley theory. Both Hanson [8] and MacDorman [9] created a series of pictures by morphing a robot to a human being. This method appears useful, since it is difficult to gather enough stimuli of highly human-like robots. However, it can be very difficult, if not impossible, for the morphing algorithm to create meaningful blends. The stimuli used in both studies contain pictures in which, for example, the shoulders of the Qrio robot simply fade out. Such beings could never be created or observed in reality and it is of no surprise that these pictures have been rated as unfamiliar. This study focuses on robots and humans that can actually be observed in reality. However, it is difficult to find examples of entities in the section of the uncanny valley between the deepest dip and the human level. We are not certain if this section actually exists thereby prompting us to suggest that the uncanny valley should be considered more of a cliff than a valley, where robots strongly resembling humans could either fall from the cliff or they could be perceived as being human. We therefore included pictures of the most human-like artificial faces (computer graphics) and pictures of slightly altered humans in our study. In addition we also included pictures from real humans in our stimuli. In essence, we are approaching the uncanny valley from its far end, backwards.

Unfortunately, MacDorman's study does not offer any statistical tests to analyze whether the observed differences were actually significant. MacDorman and Hanson, both, apply the "one clock approach" that is based on the proverb, "a person with one clock always knows the exact time while a person with two clocks can never be certain". Both authors assume that one question will measure a certain concept perfectly. Experimental psychologists are well aware of the difficulty of creating valid and reliable measurement tools, and consequently, guidelines have been developed for this purpose [10]. Instead of only asking a single question, it is often better to use multi-indicator measurement instruments, of which reliability statistics such as Cronbach's alpha, are known.

We therefore conducted a literature review to identify relevant measurement instruments. It has been reported that the way people form positive impressions of others is to

some degree dependent on the visual and vocal behavior of the targets [11] and that positive first impressions (e.g., likeability) of a person often lead to more positive evaluations of that person [12]. Interviewers report that within the first 1 to 2 minutes they know whether a potential job applicant will be hired, and people report knowing within the first 30 seconds the likelihood that a blind date will be a success [13]. There is a growing body of research indicating that people often make important judgments within seconds of meeting a person, sometimes remaining quite unaware of both the obvious and subtle cues that may be influencing their judgments. Therefore it is very likely that humans are also able to make judgments of robots based on their first impressions.

Jennifer Monathan [14] complemented her liking question with 5-point semantic differential scales: nice/awful, friendly/unfriendly, kind/unkind, and pleasant/unpleasant, because these judgments tend to share considerable variance with liking judgments [15]. She reported a Cronbach's alpha of .68, which gives us sufficient confidence to apply her scale in our study.

To measure the uncanny valley, it is also necessary to have a good measurement of the anthropomorphism of the stimuli. Anthropomorphism refers to the attribution of a human form, human characteristics, or human behavior to non-human things such as robots, computers and animals. Hiroshi Ishiguro, for example, developed highly anthropomorphic androids, such as the Geminoid HI-1 robot (see Figure 2). Some of his androids are, for a short period, indistinguishable from human beings [16].



Figure 2: Geminoid HI-1 robot

Even if it is not the intention of a certain robot to be as human-like as possible, it still remains important to match the appearance of the robot with its abilities. An overly anthropomorphic appearance can evoke expectations that the robot might not be able to fulfill. If, for example, the robot has a human-shaped face, then the naïve user will expect that the robot is able to listen and to talk. To prevent disap-

pointment, it is necessary for all developers to pay close attention to the anthropomorphism level of their robots.

An interesting behavioral measurement for anthropomorphism has been presented by Minato et al., [17]. They attempted to analyze differences in the gazes of participants looking at either a human or an android. They have not been able to produce reliable conclusions yet, but their approach could turn out to be very useful, once technical difficulties are overcome. Powers and Kiesler [18], in comparison, used a questionnaire with six items and were able to report a Cronbach's alpha of .85 which gives us sufficient confidence to apply this questionnaire in our study. In the following text, we will refer to the measurement of anthropomorphism as "human likeness" to be able to distinguish the measurement from the independent factor anthropomorphism.

In summary, this study attempts to plot the uncanny valley with particular emphasis on the last ascending section of the hypothesized curve. This study attempts to plot the curve with examples of existing humans and robots and uses more extensive measurements compared to previous studies.

A second research question is if highly human-like androids are perceived more likeable when they are being framed as robots compared to when they are being framed as humans? This question can also be posed the other way around. Will real humans be perceived less likable when they are being framed as robots compared to when they are being framed as humans?

## II. METHOD

We conducted a 3 (framing) x 4 (anthropomorphism) within participants experiment. Framing contained three conditions: human, robot, none. Anthropomorphism consisted of four conditions: real human (rh), manipulated human (fh), computer graphic (cg) and android (an). We included two additional anthropomorphism conditions that were only presented in the robot framing condition: humanoid (hd) and pet robot (pr). The three pictures within each anthropomorphism conditions were then labeled rh1, rh2, rh3, etc. We measured the human likeness (measurement of anthropomorphism) and likeability of the stimuli.

### A. Measurements

We measured the likeability of the stimuli by using Monathan's [14] liking question in addition to four of her semantic differential scales: nice/awful, friendly/unfriendly, kind/unkind, and pleasant/unpleasant. However, we deviated from her questionnaire by using a 7-point scale instead of a 5-point scale. To ensure consistency, we converted the human likeness items found in Powers and Kiesler [18] to 7-point semantic differential scales: fake/natural, machine-like/human-like, unconscious/conscious, artificial/lifelike.

### B. Stimuli

MacDorman [9] also presented movie sequences of existing robots to his participants, however as these robots were shown in different contexts and behaved differently

(only some were able to talk), MacDorman concluded that these differences create considerable noise in the measurements. To eradicate this, we picked pictures that focused on the face and did not provide any context information.

MacDorman and Hanson used sequences of morphed pictures and thereby generated entities that would be impossible to create in reality. We only used pictures of entities that either already exist or that are extremely similar to existing entities, such as computer-generated faces. It can be argued that computer-generated faces are also impossible to observe in reality. However, artists who generated these pictures focused explicitly on the creation of realistic faces, which is a great challenge in computer graphics.

Hanson pointed out that the beauty of a face already influences its likeability and that therefore great care should be taken to create beautiful androids. To avoid a possible bias we selected pictures of reasonably beautiful women. It was necessary to focus on women because only female androids were available at the start of the study. The first male highly human-like android, Geminoid HI-1 (see Figure 2), only became available shortly after the start of this study. Presenting only female androids introduces a gender bias. Female entities might in principle be preferred over male entities. To be able to at least control this bias systematically, we only presented pictures of female entities at the price of not being able to generalize to male entities.

To further prevent a possible bias by accidentally selecting an extraordinary beautiful or ugly picture of an android or human, we presented pictures of three different entities from each category, resulting in a total of 18 pictures. The pictures in the android category were of Actroid (Kokoro), EveR 1 (KITECH) and Repliee Q1 (Osaka University). The computer graphic pictures were of Kaya (Alceu Baptista), Maxim Cyberbabe (Liam Kemp) and an unnamed entity by Young Jong Cho. The Manipulated Human and Real Human images were taken from the Elle fashion magazine. The names of the models are unknown. The skin color of the faces in the Manipulated Human category was adjusted to give it a slightly green hue, producing a mildly artificial look, similar to "Data" an android character from, from the television show "Star Trek". The pictures in the Humanoid category were of Qrio (Sony), Asimo (Honda) and an unnamed humanoid from Toyota. The robot pet pictures were of Aibo (Sony), PaPeRo (NEC) and iCat (Philips Research). We are not able to present the pictures used in the experiment in this paper since some companies, such as Sony, denied us the permission.

### C. Participants

58 participants aged between 18 and 41 years (mean 21.3) filled in the survey. 28 participants were female and 30 were male. All participants were associated with a University in the Kyoto district of Japan.

### D. Procedure

The participants were guided to a computer on which the instructions and the questionnaire were presented. Before filling in the questionnaire, the participant read the instruc-

tions and had the opportunity to ask questions to the experimenter. Afterwards they had to rate each stimulus, one at a time (see Figure 3).

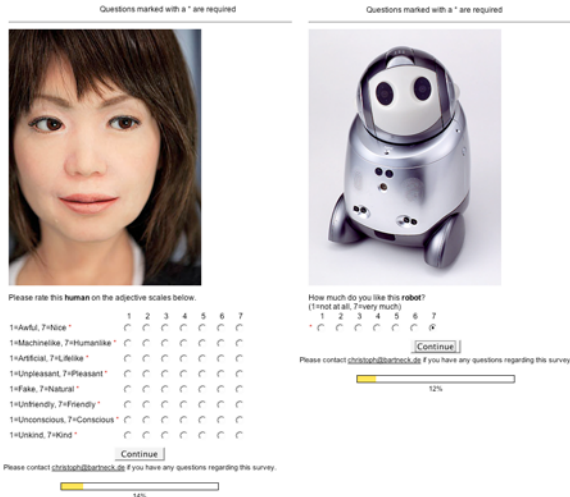


Figure 3: Two screenshots of the questionnaire.

Each of the 18 stimuli was presented twice: once with the liking question and once with the semantic differential scales. This resulted in the presentation of a total of 36 questions. The order of the 36 questions was randomized. Both times the picture was framed the same way by the question presented below the picture. The question would ask the participant to either rate this human, robot or face. The framing word was highlighted in bold typeface. Only the humanoid and toy robots were framed as “robot”, since it would be implausible to describe them as human.

The three framing conditions for this experiment were implemented by creating three different versions of the questionnaire, distinguished by different framing of the pictures. For example, in the first questionnaire version the picture of the android was framed as a human, in the second version as a robot and in the third version only as a face. The participants were randomly assigned to the three versions of questionnaire and thereby to one of the framing conditions. Each participant only filled in one of the three questionnaires.

There were three different pictures per anthropomorphism category. In each questionnaire, one of them would be framed as a human, one as a robot and one as a face. Table 1 illustrates the setup of the framing of the pictures in the three questionnaires.

TABLE 1: THE FRAMING OF THE PICTURES IN THE THREE QUESTIONNAIRES.

Questionnaire	Picture 1	Picture 2	Picture 3
1	robot	human	face
2	human	face	robot
3	face	robot	human

After the experiment the participants were debriefed and received a 1000 Yen honorarium.

### III. RESULTS

A reliability analysis across all conditions was conducted. The resulting Cronbach’s alpha for the human likeness (.878) and the likeability (.865) give us sufficient confidence in the reliability of the questionnaires.

We conducted a repeated measure analysis of variance (ANOVA) in which anthropomorphism and framing was the within participant factor. The analysis excluded data from the humanoid (hd) and toy robot (tr) conditions since they were framed only as being robots. Framing had no significant influence on the measurements. Anthropomorphism had a significant influence on human likeness ( $F(3,165)=23.451, p<.001$ ) and likeability ( $F(3,165)=9.384, p<.001$ ). The means for all anthropomorphic conditions are shown in Table 2, sorted by likeability.

TABLE 2: MEANS ACROSS THE ANTHROPOMORPHISM CONDITIONS.

	human likeness	likeability
rh	3.917	3.648
an	4.952	3.721
fh	4.840	3.882
cg	4.412	4.209
hd	2.700	5.023
tr	2.542	5.143

Since anthropomorphism had significant influences on the measurements, the analysis was extended to all anthropomorphism conditions and their individual pictures. Figure 4 and Figure 5 display the boxplots for each picture across all anthropomorphism conditions. The boxplots show that certain pictures deviate from the other pictures in the same anthropomorphism level, such as cg3 for human likeness and rh1 for likeability.

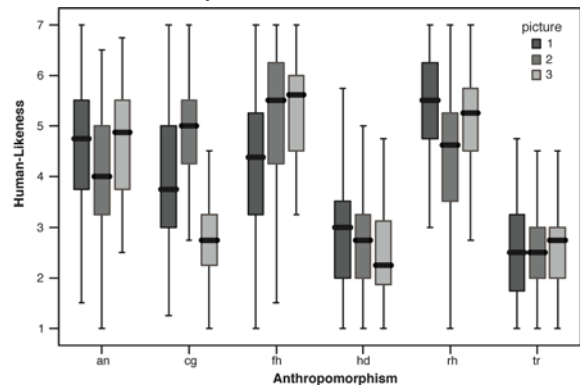


Figure 4: Boxplot of human likeness for each picture across the anthropomorphism conditions.

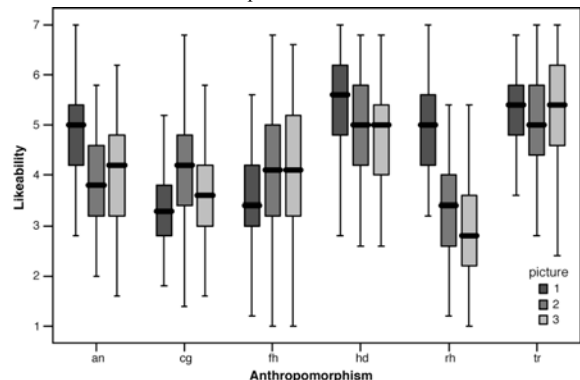


Figure 5: Boxplot of likeability for each picture across the anthropomorphism conditions.

In order to analyze each picture separately, the data was processed into a complete between participant structure. ANOVA was conducted in which anthropomorphism was the between participants factor and human likeness and likeability were the measurements. Anthropomorphism had

a significant influence on human likeness ( $F(5,1038)=128.036, p<.001$ ) and likeability ( $F(5,1038)=55.653, p<.001$ ). Post hoc t-tests with Bonferroni-corrected alpha showed that only the means for human likeness of toy robot/humanoid and real human/manipulated human were not different from each other (see Table 3). A similar pattern can be observed for the likeability means. The means for likeability were not significantly different from each other within the group of toy robot/humanoid and within the group of real human/manipulated human/computer graphics (see Table 4). The means between manipulated human and android did not differ from each other either.

TABLE 3: P VALUES OF THE COMPARISON BETWEEN THE HUMAN LIKENESS MEANS. THE BOLD TYPEFACE MARKS NON-SIGNIFICANT DIFFERENCES. A VALUE OF 0.000 INDICATES A VALUE SMALLER THAN 0.001

	an	cg	fh	hd	rh
cg	0.003				
fh	0.014	0.000			
hd	0.000	0.000	0.000		
rh	0.001	0.000	<b>1.000</b>	0.000	
tr	0.000	0.000	0.000	<b>1.000</b>	0.000

TABLE 4: P VALUES OF THE COMPARISON BETWEEN THE LIKEABILITY MEANS. THE BOLD TYPEFACE MARKS NON-SIGNIFICANT DIFFERENCES. A VALUE OF 0.000 INDICATES A VALUE SMALLER THAN 0.001

	an	cg	fh	hd	rh
cg	0.000				
fh	<b>0.193</b>	<b>0.745</b>			
hd	0.000	0.000	0.000		
rh	0.002	<b>1.000</b>	<b>1.000</b>	0.000	
tr	0.000	0.000	0.000	<b>1.000</b>	0.000

Next, the mean for each picture was plotted onto the two dimensional space of human likeness and likeability (see Figure 6). The numeric values and illustrations are available from Table 5. A Curve estimation analysis revealed that a Quadratic curve fits the data best ( $R^2 = .474$ ).

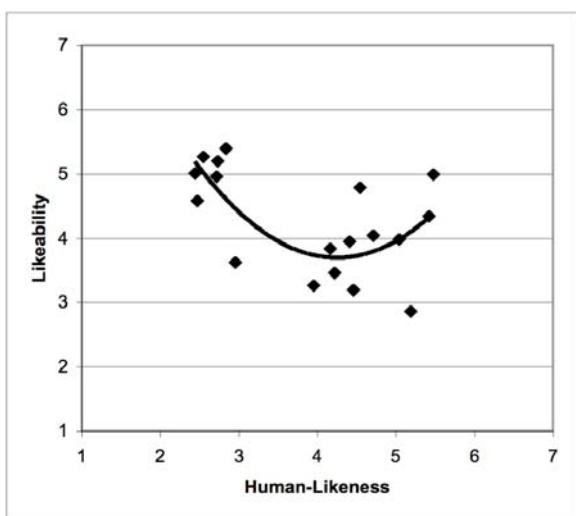


Figure 6: Means of all pictures plotted onto the human-likeness/likeability space.

#### IV. DISCUSSION AND CONCLUSION

The framing of the pictures in the questionnaire had no significant influence on the measurements. A given picture was evaluated independently from whether it was labeled human, robot or face. Knowing that a certain entity is a robot or human does in itself not constitute a positive or negative effect on its likeability or human likeness. Instead, the appearance of the entity is mainly responsible for its likeability. A highly human-like android is not uncanny because of the fact that it is a robot, but because of its appearance.

TABLE 5: MEAN HUMAN LIKENESS AND LIKEABILITY FOR EACH PICTURE SORTED BY LIKEABILITY. (CO=CODE, HL=HUMAN LIKENESS, LA=LIKEABILITY)

CO	HL	LA	CO	HL	LA
rh3	5.19	2.87	fh3	5.42	4.35
rh2	4.45	3.20	hd3	2.47	4.59
cg1	3.95	3.27	an1	4.54	4.79
fh1	4.22	3.47	hd2	2.72	4.96
cg3	2.96	3.63	rh1	5.48	5.00
an2	4.16	3.84	tr2	2.44	5.01
cg2	4.41	3.95	tr3	2.73	5.21
fh2	5.04	3.98	tr1	2.54	5.27
an3	4.71	4.05	hd1	2.84	5.40

The level of anthropomorphism and the pictures within each level of anthropomorphism had a significant influence on the measurements. Interestingly, the most liked anthropomorphism levels were all robots: toy robots and humanoids. They were even preferred over real humans.

The uncanny valley appears to be more of a cliff than a valley since even pictures of humans do not reach the level of pictures of toy robots. It has to be acknowledged that there is a small upwards trend again towards highly human-like entities, which results in a small valley. However, the most dominant feature in the graph is not the valley, but the cliff preceding it. We would like to hypothesize an alternative model for the uncanniness of robots. Figure 7 shows both the stylized curves of Mori's original prediction and our improved prediction, based on the results of this study. We speculate that maybe even the most human-like androids are not liked as much as toy robots or humanoids. While there may be a small valley in the graph, the more important feature is the cliff, and as a result we propose the existence of an uncanny "cliff" rather than an uncanny "valley".

With this in mind, it appears unwise to attempt to build highly human-like androids, since they would not be liked as much as more machine-like robots, such as Sony's Qrio.

These robots form a distinct cluster which is supported by the fact that there was no significant difference in the t-test (see Table 3 and Table 4). Even though these robots are rated significantly less human-like compared to real humans, they are liked much more. It cannot be excluded

that this preference may be based on the fact that all participants were Japanese, who are stereotypically thought of as being more avid robot-aficionados than other cultures. A comparison study with participants from other cultures would be necessary to exclude this possibility.

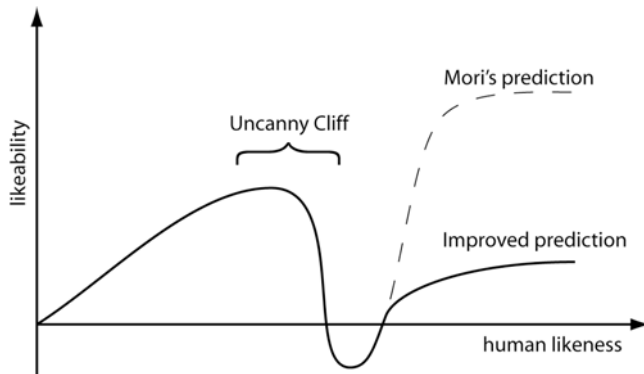


Figure 7: Hypothesized uncanny cliff

Blow et al., [3] already proposed a different description of robots' appearances. They proposed the use of the two dimensions of realism and realistic/iconic, that were originally introduced by McCloud [19], to describe the design space of robots. Hanson [8] already speculated that the mere beauty of a face might already influence the users' perception more than its level of anthropomorphism. It has also been shown that the camera angle itself influences the users' perception of a face [20]. These examples of possible influences show that human likeness is a very broad concept that includes many different aspects. So the question arises - is it possible to accurately measure and plot the uncanny cliff?

In a future study it might be beneficial to keep the framing consistent for each participant. One group might only receive pictures that are labeled "robot", and a second group might receive pictures that are all labeled "human". It might also be useful to measure the participants' frame directly by asking them to categorize each picture as being either a human or a robot. In this study, the categorization was only done indirectly through the human likeness questionnaire. It might also be worthwhile to attempt to influence the users' frame by using a visually more dominant indication, such as a larger typeface size for the framing word.

It would be interesting to compare the results of this study (using static) pictures with that of a similar study using moving pictures. The curve might have a different, more valley like, shape. In particular the comparison between moving androids and humanoids would be of interest. It would be necessary to film different robots and humans that attempt to execute the same movements. The camera position, angles and background would all need to be similar. It would even be better to test the users' perception with real robots instead of movies and pictures. The robots' social presence might have a strong effect. After all,

the goal of robotics is to bring real robots into our society and not movies and pictures of them.

## V. ACKNOWLEDGEMENTS

We thank Terumi Nakagawa for helping with the experiment and Aoife Currid for proofreading. This work was supported in part by the Ministry of Internal Affairs and Communications of Japan.

## VI. REFERENCES

- [1] M. Mori, "The Uncanny Valley," *Energy*, vol. 7, pp. 33-35, 1970.
- [2] M. Mori, "On the Uncanny Valley," presented at Humanoids-2005 workshop: Views of the Uncanny Valley, Tsukuba, 2005.
- [3] M. P. Blow, K. Dautenhahn, A. Appleby, C. Nehaniv, and D. Lee, "Perception of Robot Smiles and Dimensions for Human-Robot Interaction Design," presented at 15th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN06), Hatfield, 2006.
- [4] F. Cerami, "Miss Digital World," 2006.
- [5] B. Lam, "Don't Hate Me Because I'm Digital" in *Wired*, vol. 12, 2004.
- [6] M. Minsky, "A Framework for Representing Knowledge," in *The Psychology of Computer Vision*, P. Winston, Ed. New York: McGraw-Hill, 1975.
- [7] N. L. Etcoff, *Survival of the prettiest : the science of beauty*, 1st ed. New York: Doubleday, 1999.
- [8] D. Hanson, "Exploring the Aesthetic Range for Humanoid Robots," presented at CogSci Workshop Towards social Mechanisms of android science, Stresa, 2006.
- [9] K. F. MacDorman, "Subjective ratings of robot video clips for human likeness, familiarity, and eeriness: An exploration of the uncanny valley," presented at ICCS/CogSci-2006 Long Symposium: Toward Social Mechanisms of Android Science, Vancouver, 2006.
- [10] A. Fink, *The survey kit*, 2nd ed. Thousand Oaks, Calif.: Sage Publications, 2003.
- [11] N. Clark and D. Rutter, "Social categorization, visual cues and social judgments," *European Journal of Social Psychology*, vol. 15, pp. 105-119, 1985.
- [12] T. Robbins and A. DeNisi, "A closer look at interpersonal affect as a distinct influence on cognitive processing in performance evaluations," *Journal of Applied Psychology*, vol. 79, pp. 341-353, 1994.
- [13] J. H. Berg and K. Piner, "Social relationships and the lack of social relationship," in *Personal relationships and social support*, W. Duck and R. C. Silver, Eds. London: Sage, 1990, pp. 104-221.
- [14] J. L. Monathan, "I Don't Know It But I Like You - The Influence of Non-conscious Affect on Person Perception," *Human Communication Research*, vol. 24, pp. 480-500, 1998.
- [15] J. K. Burgoon and J. L. Hale, "Validation and measurement of the fundamental themes for relational communication," *Communication Monographs*, vol. 54, pp. 19-41, 1987.
- [16] H. Ishiguro, "Android Science - Towards a new cross-interdisciplinary framework," presented at CogSci Workshop Towards social Mechanisms of android science, Stresa, 2005.
- [17] T. Minato, M. Shimada, S. Itakura, K. Lee, and H. Ishiguro, "Does Gaze Reveal the Human Likeness of an Android?," presented at 2005 4th IEEE International Conference on Development and Learning, Osaka, 2005.
- [18] A. Powers and S. Kiesler, "The advisor robot: tracing people's mental model from a robot's physical attributes," presented at Proceeding of the 1st ACM SIGCHI/SIGART conference on Human-robot interaction, Salt Lake City, Utah, USA, 2006.
- [19] S. McCloud, *Understanding comics : the invisible art*. Northampton, MA: Kitchen Sink Press, 1993.
- [20] M. J. Lyons, R. Campbell, A. Plante, M. Coleman, M. Kamachi, and S. Akamatsu, "The Noh Mask Effect: Vertical Viewpoint Dependence of Facial Expression Perception," *Proceedings of the Royal Society of London*, vol. B 267, pp. 2239-2245, 2000.