My Robotic Doppelgänger – A Critical Look at the Uncanny Valley

Christoph Bartneck, Takayuki Kanda, *Member, IEEE*, Hiroshi Ishiguro, *Member, IEEE*, and
Norihiro Hagita, *Member, IEEE*

Abstract— The Uncanny Valley hypothesis has been widely used in the areas of computer graphics and Human-Robot Interaction to motivate research and to explain the negative impressions that participants report after exposure to highly realistic characters or robots. Despite its frequent use, empirical proof for the hypothesis remains scarce. This study empirically tested two predictions of the hypothesis: a) highly realistic robots are liked less than real humans and b) the highly realistic robot's movement decreases its likeability. The results do not support these hypotheses and hence expose a considerable weakness in the Uncanny Valley hypothesis. Anthropomorphism and likeability may be multi-dimensional constructs that cannot be projected into a two-dimensional space. We speculate that the hypothesis' popularity may stem from the explanatory escape route it offers to the developers of characters and robots. In any case, the Uncanny Valley hypothesis should no longer be used to hold back the development of highly realistic androids.

I. INTRODUCTION

THE Uncanny Valley hypothesis 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 toward them become. This trend continues until a certain point is reached, beyond which the emotional response quickly becomes repulsion. As the appearance and motion become indistinguishable from humans, the emotional reactions also become similar to those toward real humans. When the emotional reaction is plotted against the robot's level of anthropomorphism, a negative valley becomes visible (Figure 1), and this is commonly referred to as the Uncanny Valley. Moreover, movement of the robot amplifies the emotional response in comparison to static robots.

With the arrival of highly realistic androids and computer-generated movies, such as "Final Fantasy" and "The Polar Express" or "Beowulf", the topic has grabbed much public attention. The computer animation company Pixar developed a clever strategy by focusing on non-human characters in its initial movie offerings, e.g. toys in "Toy Story" and insects in "It's a Bug's Life". In contrast, the annual "Miss Digital

World" beauty competition [3, 4] has failed to attract the same level of interest.

The Uncanny Valley has been a hot topic in the research fields of computer graphics (CG) and human-robot interaction (HRI). The ACM Digital library lists 36 entries for an exact search for the phrase "uncanny valley" and Google Scholar lists an amazing 364 entries as of August 15th, 2008. We will review a few papers to provide a short overview of the current research.

Even though Mori's hypothesis was proposed in the framework of robotics, it appears to have been discussed in the field of CG before it became a topic in HRI. The development of computer software and hardware allowed the creation of increasingly realistic renderings of humans before highly realistic androids became available. In the field of CG, Mori's hypothesis is often used to explain why artificial characters are perceived as being eerie [5]. Even Gergle, Rosé & Kraut [6] used the Uncanny Valley hypothesis to explain their results during the presentation of their award-winning paper. The CG community has also specifically investigated the Uncanny Valley [7] and proposed solutions to overcome the Valley [8, 9]. A considerable number of studies in this area have been published, which led to the need for a first review article [10].

But also in the field of HRI, the hypothesis has been used to explain results [11, 12] and to motivate research [13-16]. In addition, Mori's hypothesis is often used as an engineering challenge [17-20]. The first research projects in HRI were conducted to investigate the hypothesis empirically, including studies on the participants' gaze [21, 22], androids as telecommunication devices [23], the perception of morphed robot pictures [24, 25], a sample of robot pictures [26], and the relationship of the hypothesis to the fear of death [27]. A basic limitation that these studies share is that they either focus on a single robot or use pictures and videos instead of real robots. Most research labs, including our own, simply cannot afford multiple sophisticated robots to perform comparative studies.

Mori's hypothesis has also treaded on theoretical grounds [28-30], and even legal considerations have been discussed [31]. These discussions also resulted in proposals of alternative views on the uncanny phenomena [32-34].

This short survey demonstrates that Mori's hypothesis has been widely used to motivate research & development in the fields of CG and HRI. It has also been used to explain the phenomenon of users perceiving highly realistic characters and robots as disturbing. However, the amount of empirical

proof for Mori's hypothesis has been rather scarce, as Blow, Dautenhahn, Appleby, Nehaniv, & Lee [32] observed.
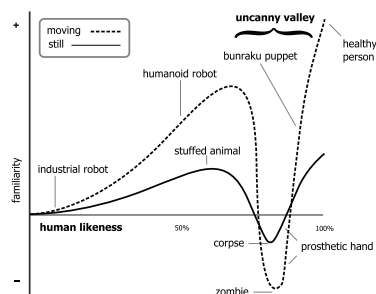

Figure 1: The uncanny valley (source: based on image by Masahiro Mori and Karl MacDorman).

This study attempts to empirically test two aspects of Mori's hypothesis. First, we are interested in the degree to which highly realistic androids are perceived differently from a human. The uncanny valley hypothesis predicts that androids would be perceived as less human-like and less likeable compared to humans. To test this hypothesis, we use Hiroshi Ishiguro and his robotic copy named "Geminoid HI-1" (see Figure 2). Moreover, we want to test whether a more human-like android would be perceived as more likeable compared to a less human-like robot. This hypothesis does not of course hold true when comparing a robot at the first peak in the graph to a robot that resides at the very bottom of the valley (see Figure 1). Accordingly, we made a small alteration to Geminoid HI-1 to make it appear less human-like.

Figure 2: Hiroshi Ishiguro and his robotic Doppelgänger Geminoid HI-1

Second, we are interested in the effect of the android's movement. Mori's hypothesis predicts that movement intensifies the users' perception of an android. A moving android would be perceived differently from an inert android. If the android were not yet human-like enough to fall into the Uncanny Valley, movement would make the android more likeable. When it does fall into the Valley, a moving android would be perceived as less likeable compared to an inert android. In summary, we are interested in the following three hypotheses.

1. Androids that are distinguishable from humans will be liked less than humans.

2. A fully moving android will be liked differently compared to an android that is limited in its movements.

3. Androids with different levels of anthropomorphism will be liked differently.

Before discussing the method of our experiment, we would like to discuss the relevant studies and the hypothesis itself. Several studies have begun empirical testing of the Uncanny Valley hypothesis. Both Hanson [24] and MacDorman [25] 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 no surprise that these pictures have been rated as unfamiliar. In our study, we exposed the participants to real androids and humans.

In his original paper, Mori plots human-likeness against 親和感 (shinwa-kan), which has previously been translated as "familiarity" as previously proposed in [26, 35]. 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 [32]. To that end, the Uncanny Valley hypothesis may only represent a short phase and hence might not deserve the attention it is receiving. We 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. It does not appear in any dictionaries and hence it does not have a direct equivalent in English. The best approach is to look at its components "shinwa" and "kan" separately. The Daijirin Dictionary (second edition) defines shinwa as "mutually be friendly" or "having similar mind". Kan is being translated as "the sense of". Given the different structure of Japanese and English, not perfect translation is possible, but "familiarity" appears to be the less suitable translation compared to "affinity" and in particular to "likeability". After an extensive discussion with native English and Japanese speakers we therefore decided to translate "shinwa-kan" as "likeability".

There is yet another reason that makes the concept of "likeability" salient for our interaction with robots. It is widely accepted that given a choice, people like familiar options because these options are known and thus safe, compared to an unknown and thus uncertain option. Even though people prefer the known option to the unknown option, this does not mean that they will like all of 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 of the robots they

know. If we consider a society in which humans interact with robots on a daily basis, the more important concept is likeability, not familiarity.

However, the discussion of the correct translation of the term shinwa-kan is not concluded and may never come to a full consensus. It has even been proposed to treat shinwa-kan as a technical term similar to Kensei Engineering, without trying to translate it. However, this would inhibit its operationalization and empirical studies would thereby become difficult to conduct. To be able to compare the results of this study to other studies that use the more widespread "familiarity" translation, we decided to also include the concepts of familiarity and eeriness in our measurements. These two measurements are not the prime focus of this study, but we included them to maintain the opportunity to view our results in relation to other studies. We used simple Likert-type scales to measure familiarity and eeriness.

We conducted a literature review to identify relevant measurement instruments for likeability. 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 [36] and that positive first impressions (e.g. likeability) of a person often lead to more positive evaluations of that person [37]. 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 [38]. 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 [39] complemented her liking question with 5-point semantic differential scales that rate nice/awful, friendly/unfriendly, kind/unkind, and pleasant/unpleasant, since these judgments tend to share considerable variance with liking judgments [40]. 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 take a measurement of the stimuli's anthropomorphism level. Anthropomorphism refers to the attribution of a human form, human characteristics, or human behavior to non-human things such as robots, computers and animals. An interesting behavioral measurement has been presented by Minato et al. [22] They attempted to analyze the differences in the gazes of the participants, who looked 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, assuming that they can overcome the technical difficulties. Powers and Kiesler [13], in contrast, were able to compile an anthropomorphism questionnaire that resulted in a Cronbach's alpha of .85. We decided to use this questionnaire for our study.

In addition to a measurement for likeability and anthropomorphism, one may also consider whether humans experience a general anxiety when interacting with robots, similar to computer anxiety [41, 42]. Such an underlying robot anxiety could have a considerable influence on the measurements, so we included the Robot Anxiety Scale (RAS) in our experiment [43].

II. METHOD

We conducted a 3 (anthropomorphism) x 2 (movement) experiment in which anthropomorphism was the within-participants factor and movement was the between-participants factor. The anthropomorphism factor had three conditions: masked android, android, and human. The movement factor had two conditions: full movement and limited movement. The Ethics Committee of the Osaka University approved this project and all participants signed a written informed consent.

III. PARTICIPANTS

The participants in this study were 19 men and 13 women in their early 20's attending Japanese universities in the Kansai area (Doushisha University 22, Kyoto Sangyo University 2, Ryukoku University 2, Ritumeikan University 2, Kyoto Bunkyo University 1, Osaka Keizai University 1, Kansei University 1, Doushisha Zyoshi University 1). The male and female participants were distributed approximately even across the experimental conditions. They were not exposed to any previous study in the laboratory. This study was conducted shortly before the official release of the Geminoid HI-1 and hence they were not exposed to the considerable media exposure that the Geminoid HI-1 received.

IV. PROCEDURE

The participants were welcomed and then asked to fill in a questionnaire, which took approximately 10 minutes. Next, the participants were guided to the experiment room. They were seated on a chair that was placed one meter away from the android/person. Afterwards the experimenter introduced them to each other without explicitly labeling Ishiguro as a human and the androids as robots.

Figure 4: Hiroshi Ishiguro and Geminoid HI-1 talk to a participant

In the full-movement condition, the android/person alternated its/his head direction and gaze between the experimenter and the participant. In addition, the android/person would show randomized subtle movements. In the limited-movement condition, the android/person

271

would only look straight ahead at the participant. After the introduction, the experimenter stepped away and invited the android/person to ask a question. The android/person would ask a different question for each of the three anthropomorphism conditions. The questions were: 1) What is your age? 2) Which university do you attend? 3) What is your name? (see Figure 4).

After receiving the answer, the android/person thanked the participant and the experimenter ended the session by asking the participant to follow her to the next room. This section of the experiment took approximately 5 minutes. In the next room, the participants filled in another questionnaire, which took again about 10 minutes. The procedure was repeated for all three anthropomorphism conditions.

The participants were randomly assigned to one of the two between-participants conditions, and the order of the three within-participants conditions was cross-balanced.

A. Measurements

The difference in appearance and behavior between humans and their robotic counterparts is still significant, and so far it has only been possible to trick participants into believing otherwise for a maximum of two seconds [28]. Accordingly, we did not attempt to present Ishiguro himself as being a robot. The users' perception of the android's/person's anthropomorphism, likeability and robot anxiety was measured using questionnaires. The questionnaires for the human condition asked the participant to evaluate this person, and in each of the robot conditions, the questionnaire asked the participants to evaluate this robot.

The anthropomorphism questionnaire was based on the items proposed by Powers and Kiesler [13]. They reported a Cronbach's alpha of .85, which gave us sufficient confidence in their items. For consistency, we transformed their items into 7-point semantic differential scales: Fake/Natural, Machinelike / Humanlike, Unconscious / Conscious, Artificial/Lifelike and Moving rigidly/Moving elegantly. To avoid confusion with the anthropomorphism condition we will refer to this measurement as human-likeness.

The likeability questionnaire was based on the scales proposed by Jennifer Monathan [39]. For consistency we used 7-point scales instead of the 5-point scales for her semantic differential scales: awful/nice, unfriendly/friendly, unkind/kind, and unpleasant/pleasant.

Robot anxiety was measured using the Robot Anxiety Scale (RAS) proposed by Nomura et al. [43] Their eleven 6-point Likert-type questions are clustered into three subscales: anxiety toward communication capacity of robots (S1), anxiety toward behavioral characteristics of robots (S2), and anxiety toward discourse with robots (S3). For the human condition, the RAS questionnaire was rephrased for asking the participants to evaluate the human. To disguise the intention of the questionnaire items we also included several dummy items, such as: "I may touch robots and break them".

B. Setup

For the anthropomorphism factor in this experiment, we used Hiroshi Ishiguro in the human condition and his robotic copy Geminoid HI-1 in the two android conditions. By using a human and his almost indistinguishable copy, we were able to eliminate a possible bias due to appearance. Otherwise, participants might have liked or disliked the android because it looks more or less handsome than the human.

Geminoid HI-1 is currently the only android available for such comparison studies. The previously developed Repliee R1 android, which is a copy of the famous newscaster Ayako Fuji, is rarely available for comparison studies because Ayako Fuji is intensely preoccupied with her work.

For the second, less human-like android condition, we considered exposing parts of Geminoid's mechanical interior, such as its arms or torso, but this looked too much like a severe injury. Sick or injured people are generally perceived as disturbing [44], so we decided to implement a plausible modification to the android to make it look less human-like. We built a visor, which displayed LED lights, as a replacement for Geminoid's original glasses (see Figure 5). The visor suggests the presence of special cameras, and thus it seemed plausible for the android to wear it.

Figure 5: Geminoid HI-1 with its glasses and its visor

For the android's voice we used audio recordings from Hiroshi Ishiguro. During the audio recording, we also captured his facial gestures, including his lip movements. The motion data were then used to animate the android. A speaker was placed behind the back of the android to play the recorded voice synchronously with its lip movement. The participants were not able to see the speaker and for the participants the android clearly appeared to be the source of the audio signal. This procedure allowed us to create realistic animations for the speech act, in particular with respect to the lip movements.

The two movement conditions were based on the captured motion data. In the full-movement condition, the captured motion data were complemented with animations for the torso, arms and legs. The overall movements were designed to reflect normal conversational behavior, which included looking at the speaker, looking away from the speaker, and subtle random movements of the body. Researchers who were very familiar with Hiroshi Ishiguro

designed the animations, and their goal was to create motion that closely resembled his normal behavior. The focus was placed on subtle behaviors, since it was not yet possible to make grand movements, such as waving an arm, look natural. The control of Geminoid's HI-1 air-pressured actuators turned out to be very difficult. The animation software did not yet include a model of the physical properties of the actuators and hence a direct mapping between the screen representation and the robot itself was problematic. However, subtle movements that are frequently observed in human-human interaction could be implemented successfully. Before the experiment, Hiroshi Ishiguro verified that the android's movement appeared familiar and natural to him. For the limited-movement condition, all animations except eye blinking and lip movements were deactivated. Hiroshi Ishiguro would not have been able to speak without moving his lips, and naturally he was not able to stop his eye blinking. A completely inert android would also have no practical use: Robots are built to move, and a comparison with an inert android would only be interesting from a theoretical point of view.

The Geminoid HI-1 was originally built for a teleconference system [23], and it does not yet include an autonomous dialogue management system. Therefore, we had to restrict the interaction to a short and controllable dialogue. The answers to the questions were easy to predict, and thus the dialogue could come to a predictable and plausible conclusion.

V. RESULTS

In the experiment, 16 participants were assigned to the full-movement condition and 16 participants to the limited-movement condition. Levene's test for quality of variance was not significant for any of the measurements, so homogeneous variance can be assumed.
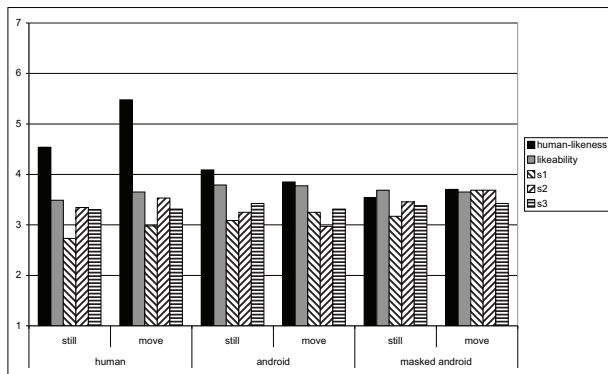

Figure 6: Means of all measurements under all conditions. Human-likeness and likeability were measured on a 7-point scale, while the RAS scales (S1, S2 and S3) were measured on a 6-point scale.

We tested the internal consistency of the questionnaire and we can we can report a Cronbach's alpha for the human-likeness questionnaire of .929 for the human condition, .923 for the android condition and .856 for the masked android condition. For the likeability questionnaire we can report a

Cronbach's alpha of .923 for the human condition, .878 for the android condition and .842 for the masked android condition. The alpha values are well above .7 and hence we can conclude that the anthropomorphism questionnaire has sufficient internal consistency reliability. The means of all measurements under all conditions are shown in Figure 6.

We conducted an analysis of variance (ANOVA) on the available data (n=32). Anthropomorphism had a significant $(F_{(2,60)}=19.864, p<.001)$ influence on human-likeness but not on likeability $(F_{(2,60)}=.614, p=.544)$. It also had a significant influence on RAS S1 $(F_{(2,60)}=3.482, p=.037)$ and RAS S2 $(F_{(2,60)}=4.195, p=.020)$ but not on RAS S3 $(F_{(2,60)}=.139, p=.870)$. Table 1 presents the F and p values of all measurements in all conditions. Post-hoc Bonferroni alpha-corrected comparisons revealed that mean for S2 in the masked android condition (3.367) was significantly lower (p=.011) than the mean for the android (3.398). The mean for S1 in the human condition (2.854) was nearly significantly lower (p=.063) than the mean for the masked android (3.427).

TABLE 1: F AND P VALUES FOR ALL CONDITIONS.

	MOVE		ANTHO		MOVE *ANTHRO	
	F	P	F	P	F	P
hum. likeness	0.523	0.475	19.864	0.000	3.404	0.040
likeability	0.035	0.853	0.614	0.544	0.161	0.851
s1	0.945	0.339	3.482	0.037	0.363	0.697
s2	0.029	0.866	4.195	0.020	1.515	0.228
s3	0.006	0.941	0.139	0.870	0.104	0.901
familiarity	0.135	0.716	1.107	0.337	2.362	0.103
eeriness	0.001	0.998	1.391	0.257	0.674	0.513

Movement had no significant influence on any measurement, but an interaction effect between anthropomorphism and movement could be observed for human-likeness $(F_{(2,60)}=3.404, p=.040)$. A human with limited movement is perceived to be less human-like than a human that fully moves. There was no difference in perceived human-likeness between a limited-movement android and a fully moving android. We performed an ANOVA to test whether the gender of the participants influenced the measurements, but the results revealed no significant difference.

We conducted a second ANOVA in which only the first session of each participant was considered to eliminate any possible influence that the repeated exposure of the stimuli to the participant might have had. It follows that Anthropomorphism is the between-subject factor. The disadvantage of this procedure is that the number of participants is spread across the three conditions of Anthropomorphism. Ten participants are in the android condition, eleven in the human condition and eleven in the masked android condition. The ANOVA revealed that Anthropomorphism had no significant effect on any of the measurements (see Table 2).

TABLE 2: F AND P VALUES FOR ALL MEASUREMENTS

MEASUREMENT	F(2,29)	P
Human-likeness	0.578	0.568
Likeability	0.962	0.394
S1	3.195	0.056
S2	1.808	0.182
S3	1.198	0.316
Familiarity	1.607	0.218
Eeriness	1.190	0.319

Next, we conducted a stepwise multiple regression analysis to determine the relation between likeability and the other measurements, in particular with the alternative translation of Shinwa-kan. Table 3 presents the Pearson Correlation Coefficient and the italic styles indicates significance levels at $p<0.05$.

TABLE 3: PEARSON CORRELATION COEFFICIENT OF THE MEASUREMENTS (THE ITALIC STYLE DENOTES CORRELATIONS THAT ARE SIGNIFICANT AT P<0.05)

	LIKEAB.	FAMIL.	EERINESS	S1	S2
FAMILIARITY	0.591				
EERINESS	*-0.159*	0.264			
S1	-0.190	*-0.316*	-0.023		
S2	*-0.304*	0.092	0.014	0.210	
S3	-0.220	-0.149	0.092	*0.482*	0.198

Only familiarity and s2 entered the final model, and the stepwise regression revealed quite a good fit (50.8% of the variance explained). The Analysis of Variance (ANOVA) revealed that the model was significant (R=0.713) (F (2,29) = 14.980, p < 0.01).

VI. CONCLUSIONS

Against Mori's prediction, androids that were distinguishable from humans were not liked less than humans in our study. The results showed that the participants were able to distinguish between the human stimulus and the android stimuli. The human was rated as being significantly more human-like compared to the two androids. However, the ratings for likeability were not significantly different. This result does not support Mori's hypothesis. Two possible interpretations came to mind. On the one hand, there really could be no difference between the likeability of humans and that of androids. On the other hand, likeability could be a more complex phenomenon. We speculate that the participants might have used different standards to evaluate the likeability of the human and the androids. As a robot, the displayed android might have been likeable to the same degree as the human was likeable as a human; however, the expectations for these two categories might have been different.

We were also unable to confirm our second hypothesis. The results show no significantly different likeability ratings for a fully moving android compared to an android that is limited in its movements. It must be acknowledged that Mori originally considered only a completely inert robot rather than a moving robot. However, such an inert robot would have only theoretical value. While we were not able to measure different likeability ratings for the three anthropomorphism conditions, we were able to detect a significant interaction effect between movement and anthropomorphism. In the human condition, the limited-movement human was rated less human-like compared to the full-movement human. Clearly, the human violated the social standard of changing his gaze. In conversations, humans alternate their gaze between the eyes of the communication partner and the environment. Staring straight ahead does not comply with this standard.

However, this misbehavior did not result in different ratings for the android. This may be explained by the participants not applying the social standard set for humans to the android. As a tangible analogy, we also do not expect our cats to follow the human standard of looking alternately at the speaker and the environment.

We propose to consider movement not as a one-dimensional value but as a factor that carries social meaning. An extensive amount of research in the area of gesture and posture is available [45-47] and should be considered. Moreover, the meaning of a certain gesture could be different for humans and robots. We may apply a different social standard and hence evaluate the gesture differently. Mori's proposed hypothesis of movement appears to be too simplistic.

We also failed to produce a conclusive result for our third hypothesis. The androids' human-likeness was not rated significantly different despite the presence of the visor. Consequently, we could not test whether androids with different levels of anthropomorphism are liked differently. The only significant effect we could measure was that the mean of the RAS S2 (anxiety toward behavioral characteristics of robots) scale was higher in the masked android condition compared to the android condition. A speculative explanation of this result is that the participants were worried about the purpose of the visor. The function of this novel device with its built-in lights might have been unclear and thus worrisome.

The absence of a significant difference for human-likeness between the two android conditions strengthens our speculation that anthropomorphism is a complex phenomenon involving multiple dimensions. Not only the appearance but also the behavior of a robot can have a considerable influence on anthropomorphism.

We also observed that in our study, the likeability of the stimuli was highly positively correlated with the rating for familiarity and significantly negative correlated with S1 rating (anxiety toward communication capacity of robots). This indicates that our interpretation of shinwa-kan may not be exactly the same as familiarity, but that it is certainly highly correlated.

VII. DISCUSSION

The results of this study cannot confirm Mori's hypothesis of the Uncanny Valley. The robots' movements and their level of anthropomorphism may be complex phenomena that cannot be reduced to two factors. Movement contains social meanings that may have direct influence on the likeability of a robot. The robot's level of anthropomorphism does not only depend on its appearance but also on its behavior. A mechanical-looking robot with appropriate social behavior can be anthropomorphized for different reasons than a highly human-like android. Again, Mori's hypothesis appears to be too simplistic.

Simple models are in general desirable, as long as they have a high explanatory power. This does not appear to be the case for Mori's hypothesis. Instead, its popularity may be based on the explanatory escape route it offers. The Uncanny Valley can be used in attributing the users' negative impressions to the users themselves instead of to the shortcomings of the agent or robot. If, for example, a highly realistic screen-based agent received negative ratings, then the developers could claim that their agent fell into the Uncanny Valley. That is, instead of attributing the users' negative impressions to the agent's possibly inappropriate social behavior, these impressions are attributed to the users. Creating highly realistic robots and agents is a very difficult task, and the negative user impressions may actually mark the frontiers of engineering. We should use them as valuable feedback to further improve the robots.

We also conclude that our translation of Shinwa-kan is different from the more popular translation of "familiarity". We dare to claim that our translation is closer to Mori's original intention and the results show that the correlation of likeability to familiarity is high enough to allow for comparison with other studies. We also feel the need to point out that just because a certain translations or interpretations are used often, does not automatically make them true. Politicians occasionally repeat their statements to make them appear to be true (e.g. "Irak has weapons of mass destruction"), but the pure repetition of a statement does not automatically make it true. Alternative, possibly better translation or interpretations, should always be considered. When we focus back onto this study, we observe that the two concepts of likeability and familiarity are related, but not completely similar. Even if we consider the more popular translation of shinwa-kan, the results of our study still do not confirm Mori's hypothesis. Anthropomorphism and Movement did neither have a significant influence on likeability nor on familiarity.

We would also like to point out that not finding significant differences between two experimental conditions is, in principle, just as valuable as finding out that two conditions are significantly different. Showing, for example, that men are not significantly more intelligent than women is an important result. No matter if significant differences were found or not, one can try to think about a third factor that explains the difference or the lack of difference thereof between two experimental conditions. This hypothesized third factor then becomes subject of a new experiment. However, no emphasis on proposing confounding factors should be given to experiments that did not find a significant difference. We will now elaborate on some limitations of our study and point towards additional factors that would be of interest of further studies.

It must be acknowledged that the interaction with robots in this work was not only short but also simple. The android or human asked a question that could be easily answered, and there were no important consequences resulting from the answer. However, it has been shown that people are able to make judgments of other humans based on a short first impression [38], so we believe that the interaction duration in this study was sufficient for the participant to form a judgment. However, in the future, robots may be involved in more complex situations and become full members in human-robot teams. We can assume that the social meaning of the robots' movements will become even more important in such a scenario. The robot would need to be able, for example, to use its body language to manage turn taking in conversations.

We are aware that our study only used one human and his robotic copy as the stimuli. One might expect that the overall ratings for likeability would be different for another human/android pair, such as for Ayako Fuji and her robotic copy Repliee Q1. The limited availability of robotic doppelgängers makes such comparison studies difficult, but we believe that pursuing such investigations could provide the needed insight to advance the research in this area. It is also clear that we did not use a perfectly inert robot, as it was suggested by Mori. In that sense, we are not able to confirm or disconfirm that specific part of his hypothesis. However, a perfectly inert robot would have very limited applications and is therefore of limited relevance. It can also be assumed that Mori did not consider movement to only consist of the two states On and Off, but that a robot may move more or less than another. We therefore feel confident that our results are still within the framework of Mori's hypothesis.

Another open question is to what degree the results presented are generalizable to users with different cultural backgrounds. While we are not able to provide a definite answer to this question, we do doubt that the results are particularly exotic. It has been demonstrated that the users' cultural background has a significant influence on their attitude towards robots, but the Japanese users were not as positive as stereotypically assumed [35]. In a previous study, the US participants had the most positive attitude, while participants from Mexico had the most negative attitude [48].

One strategy to avoid negative impressions may be to match the appearance of the agent or robot with its abilities. An excessively anthropomorphic appearance can evoke expectations that the agent/robot might not be able to fulfill. If, for example, the robot had a human-shaped face then a naïve user would expect the robot to be able to listen and to talk. However, current speech technology is not yet mature enough to enable robots to fluently converse with humans, which may lead to frustration. To prevent disappointment, it is necessary for all developers to pay close attention to the level of anthropomorphism achieved in their agents or robots.

But there is hope for the development of highly realistic androids. Our study casts considerable doubt on the Uncanny Valley hypothesis. The androids used in this study were liked just as much as the human original. Of course it is still difficult to create and animate androids, but the Uncanny Valley hypothesis can no longer be used to hold back development.

There is also hope that the recent developments in embodied intelligence may provide the key to further develop AI [49], which in turn may help to increase the abilities of robots. This may then enable us to move away from Wizard-of-Oz type experiments that are currently often necessary to enable to robot to exhibit intelligent behavior.

SUPPLEMENTARY INFORMATION

Supporting online material includes 18 movies that illustrate each of the experimental conditions. They are available at: http://www.idemployee.id.tue.nl/c.bartneck/research/doppelganger

ACKNOWLEDGMENTS

REFERENCES

[1] M. Mori, "The Uncanny Valley," *Energy,* vol. 7, pp. 33-35, 1970.
[2] M. Mori, "On the Uncanny Valley," in *Humanoids-2005 workshop: Views of the Uncanny Valley*, Tsukuba, 2005.
[3] F. Cerami, "Miss Digital World," 2006.
[4] B. Lam, "Don't Hate Me Because I'm Digital " in *Wired*. vol. 12, 2004.
[5] S. Dow, M. Mehta, A. Lausier, B. MacIntyre, and M. Mateas, "Initial lessons from AR Facade, an interactive augmented reality drama," in *ACM SIGCHI international conference on Advances in computer entertainment technology*, Hollywood, California, 2006.
[6] D. Gergle, C. P. Rosé, and R. E. Kraut "Modeling the Impact of Shared Visual Information on Collaborative Reference," in *Conference on Human Factors in Computing Systems (CHI2007)*, San Jose, 2007, pp. 1543-1552.
[7] J. i. Seyama and R. S. Nagayama, "The Uncanny Valley: Effect of Realism on the Impression of Artificial Human Faces," *Presence: Teleoperators & Virtual Environments,* vol. 16, pp. 337-351, 2007.
[8] F. Pighin and J. P. Lewis, "Introduction," in *ACM SIGGRAPH 2006 Courses*, Boston, Massachusetts, 2006.
[9] B. Windsor, "Capturing Puppets: Using the age old art of puppetry combined with motion capture to create a unique character animation approach," in *International Conference On Computer Graphics & Virtual Reality*, Las Vegas, 2006, pp. 49-55.
[10] V. Vinayagamoorthy, A. Steed, and M. Slater, "Building Characters: Lessons Drawn from Virtual Environments," in *CogSci-2005 Workshop - Toward Social Mechanisms of Android Science*, Stresa, Italy, 2005.
[11] C. F. DiSalvo, F. Gemperle, J. Forlizzi, and S. Kiesler, "All robots are not created equal: the design and perception of humanoid robot heads," in *Designing interactive systems: processes, practices, methods, and techniques*, London, England, 2002, pp. 321-326.
[12] A. Geven, J. Schrammel, and M. Tscheligi, "Interacting with embodied agents that can see: how vision-enabled agents can assist in spatial tasks," in *4th Nordic conference on Human-computer interaction: changing roles*, Oslo, Norway, 2006, pp. 135-144.
[13] A. Powers and S. Kiesler, "The advisor robot: tracing people's mental model from a robot's physical attributes," in *1st ACM SIGCHI/SIGART conference on Human-robot interaction*, Salt Lake City, Utah, USA, 2006.
[14] J. G. Trafton, A. C. Schultz, D. Perznowski, M. D. Bugajska, W. Adams, N. L. Cassimatis, and D. P. Brock, "Children and robots learning to play hide and seek," in *1st ACM SIGCHI/SIGART conference on Human-robot interaction*, Salt Lake City, Utah, USA, 2006, pp. 242 - 249.
[15] M. L. Walters, K. Dautenhahn, S. N. Woods, K. L. Koay, R. T. Boekhorst, and D. Lee, "Exploratory studies on social spaces between humans and a mechanical-looking robot," *Connection Science,* vol. 18, pp. 429-439, 2006.
[16] E. Wang, C. Lignos, A. Vatsal, and B. Scassellati, "Effects of head movement on perceptions of humanoid robot behavior," in *1st ACM SIGCHI/SIGART conference on Human-robot interaction*, Salt Lake City, Utah, USA, 2006, pp. 180-185.
[17] S. Maddock, J. Edge, and M. Sanchez, "Movement realism in computer facial animation," in *Workshop on Human-animated Characters Interaction at The 19th British HCI Group Annual Conference HCI 2005: The Bigger Picture*, Edinburgh, 2005.
[18] S. Marti and C. Schmandt, "Physical embodiments for mobile communication agents," in *18th annual ACM symposium on User interface software and technology*, Seattle, WA, USA, 2005, pp. 231-240.
[19] D. Matsui, T. Minato, K. F. MacDorman, and H. Ishiguro, "Generating natural motion in an android by mapping human motion," in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2005)*, Edmonton, 2005, pp. 3301 - 3308
[20] J. E. Young, M. Xin, and E. Sharlin, "Robot expressionism through cartooning," in *ACM/IEEE international conference on Human-robot interaction*, Arlington, Virginia, USA, 2007, pp. 309 - 316.
[21] K. F. Mac Dorman, T. Minato, M. Shimada, S. Itakura, S. Cowley, and H. Ishiguro, "Assessing Human Likeness by Eye Contact in an Android Testbed," in *CogSci-2005 Workshop - Toward Social Mechanisms of Android Science*, Stresa, Italy, 2005.
[22] T. Minato, M. Shimada, S. Itakura, K. Lee, and H. Ishiguro, "Does Gaze Reveal the Human Likeness of an Android?," in *4th IEEE International Conference on Development and Learning*, Osaka, 2005.
[23] D. Sakamoto, T. Kanda, O. Tetsuo, H. Ishiguro, and N. Hagita, "Android as a telecommunication medium with a human-like presence," in *Proceeding of the ACM/IEEE international conference on Human-robot interaction*, Arlington, Virginia, USA, 2007, pp. 193-200.
[24] D. Hanson, "Exploring the Aesthetic Range for Humanoid Robots," in *CogSci Workshop Towards social Mechanisms of android science*, Stresa, 2006.
[25] K. F. MacDorman, "Subjective ratings of robot video clips for human likeness, familiarity, and eeriness: An exploration of the uncanny valley," in *ICCS/CogSci-2006 Long Symposium: Toward Social Mechanisms of Android Science*, Vancouver, 2006.
[26] C. Bartneck, T. Kanda, H. Ishiguro, and N. Hagita, "Is the Uncanny Valley an Uncanny Cliff?," in *16th IEEE International Symposium on Robot and Human Interactive Communication, RO-MAN 2007*, Jeju, Korea, 2007, pp. 368-373.
[27] K. F. Mac Dorman, "Mortality salience and the uncanny valley," in *5th IEEE-RAS International Conference on Humanoid Robots* Tsukuba, 2005, pp. 399-405.
[28] H. Ishiguro, "Interactive Humanoids and Androids as Ideal Interfaces for Humans," in *International Conference on Intelligent User Interfaces (IUI2006)*, Syndney, 2006, pp. 2-9.
[29] H. Ishiguro, "Android science: conscious and subconscious recognition," *Connection Science,* vol. 18, pp. 319-332, 2006.
[30] F. Kaplan, "Everyday robotics: robots as everyday objects," in *Joint conference on Smart objects and ambient intelligence: innovative context-aware services: usages and technologies*, Grenoble, France, 2005, pp. 59-64.
[31] D. J. Calverley, "Android science and animal rights, does an analogy exist?," *Connection Science,* vol. 18, pp. 403 - 417, 2006.
[32] M. P. Blow, K. Dautenhahn, A. Appleby, C. Nehaniv, and D. Lee, "Perception of Robot Smiles and Dimensions for Human-Robot Interaction Design," in *15th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN06)*, Hatfield, 2006.
[33] H. Brenton, M. Gillies, D. Ballin, and D. Chatting, "The Uncanny Valley: does it exist?," in *Workshop on Human-Animated Characters Interaction at The 19th British HCI Group Annual Conference HCI 2005: The Bigger Picture*, Edinburgh, 2005.
[34] M. Shimada, T. Minato, S. Itakura, and H. lshiguro, "Uncanny Valley of Androids and Its Lateral Inhibition Hypothesis," in *Robot and Human interactive Communication, 2007. RO-MAN 2007. The 16th IEEE International Symposium on*, 2007, pp. 374-379.
[35] C. Bartneck, "Who like androids more: Japanese or US Americans?," in *17th IEEE International Symposium on Robot and Human Interactive Communication, RO-MAN 2008*, München, 2008, pp. 553-557.
[36] N. Clark and D. Rutter, "Social categorization, visual cues and social judgments," *European Journal of Social Psychology,* vol. 15, pp. 105-119, 1985.
[37] 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.
[38] 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.
[39] 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.
[40] 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.
[41] A. C. Raub, "Correlates of computer anxiety in college students," University of Pennsylvania, 1981.
[42] M. J. Brosnan, *Technophobia : the psychological impact of information technology*. New York: Routledge, 1998.
[43] T. Nomura, T. Suzuki, T. Kanda, and K. Kato, "Measurement of Anxiety toward Robots," in *The 15th IEEE International Symposium on Robot and Human Interactive Communication, ROMAN2006*, Hatfield, 2006, pp. 372-377.
[44] N. L. Etcoff, *Survival of the prettiest : the science of beauty*, 1st ed. New York: Doubleday, 1999.
[45] R. M. Krauss, P. Morrel-Samuels, and C. Colasante, "Do conversational hand gestures communicate?," *Journal of personality and social psychology,* vol. 61, pp. 743-754, 1991 1991.
[46] D. McNeill, *Hand and mind: What gestures reveal about thought*. Chicago: University of Chicago, 1992.
[47] S. Nobe, S. Hayamizu, O. Hasegawa, and H. Takahashi, "Are Listeners Paying Attention to the Hand Gestures of an Anthropomorphic Agent? An Evaluation Using a Gaze Tracking Method," in *Gesture and Sign Language in Human-Computer Interaction*, I. Wachsmuth and M. Fröhlich, Eds. Heidelberg: Springer, 1997.
[48] C. Bartneck, T. Suzuki, T. Kanda, and T. Nomura, "The Influence of People's Culture and Prior Experiences with Aibo on their Attitude Towards Robots," *AI & Society - The Journal of Human-Centred Systems,* vol. 21, pp. 217-230, 2007.
[49] R. Pfeifer and J. Bongard, *How the body shapes the way we think: a new view of intelligence*. Cambridge, USA: MIT Press, 2006.