Androids as an Experimental Apparatus:
Why Is There an Uncanny Valley and Can We Exploit It?

Karl F. MacDorman

Department of Adaptive Machine Systems and Frontier Research Center
Graduate School of Engineering
Osaka University, Suita, Osaka 565-0871 Japan

Abstract

Abstract. It seems natural to assume that the more closely robots come to resemble people, the more likely they are to elicit the kinds of responses people direct toward each other. However, subtle flaws in appearance and movement only seem eerie in very humanlike robots. This uncanny phenomenon may be symptomatic of entities that elicit a model of a human other but do not measure up to it. If so, a very humanlike robot may provide the best means of finding out what kinds of behavior are perceived as human, since deviations from a human other are more obvious. In pursuing this line of inquiry, it is essential to identify the mechanisms involved in evaluations of human likeness. One hypothesis is that an uncanny robot elicits an innate fear of death and culturally-supported defenses for coping with death's inevitability. An experiment, which borrows from the methods of terror management research, was performed to test this hypothesis. Across all questions subjects who were exposed to a still image of an uncanny humanlike robot had on average a heightened preference for worldview supporters and a diminished preference for worldview threats relative to the control group.

Introduction

An experimental apparatus that is indistinguishable from a human being, at least superficially, has the potential to contribute greatly to an understanding of face-to-face interaction in the social and neurosciences. Such a device could be a perfect actor in controlled experiments, permitting scientists to vary precisely the parameters under study. It could also serve as a testbed for cognitive theories, including theories about how the brain acts as a control system in mediating whole-bodied communication. The device would also have the advantage of having the physical presence that simulated characters lack. Unfortunately, no such device yet exists, nor will one any time soon; nevertheless, robots are being built that with each new generation more closely simulate human beings in appearance, facial expression, and gesture [Minato et al., 2004] [MacDorman et al., 2005] [Matsui et al., 2005]. They are capable of eliciting some of the kinds of responses that people direct toward each other but not toward mechanical-looking robots.

Humanlike robots, often referred to as *androids* in the robotics literature to distinguish them from mechanical-looking humanoid robots, may prove more capable of eliciting a subject's model of a human other than any other contrivance to date. One apparent symptom of their potential for eliciting human-directed responses is
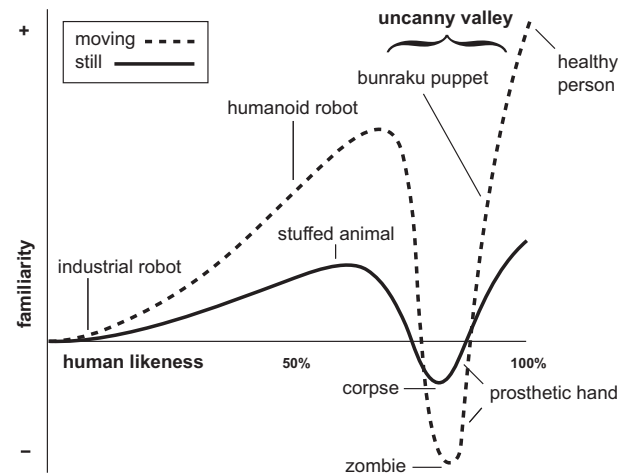

Figure 1: As a robot designer, Mori graphed what he saw as the relation between human likeness and perceived familiarity: familiarity increases with human likeness until a point is reached at which subtle differences in appearance and behavior create an unnerving effect [Mori, 1970]. This he called the *uncanny valley*. According to Mori, movement amplifies the effect.

a phenomenon Masahiro Mori identified as the *uncanny valley* [Mori, 1970].

Mori predicted that, as robots appear more human, they seem more familiar until a point is reached at which subtle imperfections create a sensation of strangeness (see Fig. 1 and Appendix B). He noted that some prosthetic hands are, at first glance, indistinguishable from real hands. They simulate muscles, tendons, veins, skin pigmentation, fingernails, and even finger prints. However, if you shook one, the lack of soft tissue and cold temperature would give you a shock. The fact that these hands can move automatically only increases the sensation of strangeness (as shown by the dashed line in Fig. 1). To build a complete android, Mori believed, would only multiply this eerie feeling many times over: Machines that appeared too lifelike would be unsettling or even frightening inasmuch as they resemble figures from nightmares or films about the living dead. Therefore, Mori cautioned robot designers not to make the second peak their goal — that is, total human likeness — but rather the first peak of humanoid appearance to avoid the risk of their robots falling into the uncanny valley.

The uncanny valley can, however, be seen in a positive light. While many nonbiological phenomena can violate our expectations, the eerie sensation associated with the uncanny valley may be particular to the violation of (largely nonconscious) human-directed expectations. If very humanlike robots are capable of eliciting human-directed expectations, then subjects can be used to evaluate the human likeness of their behavior to an extent that would be impossible if mechanical-looking robots were used instead.

Unfortunately, there has been little direct scientific investigation of Mori's uncanny valley hypothesis in the past 35 years. Clearly, there are many qualitatively different ways of deviating from human norms of appearance and movement, some of which are more uncanny than others. In addition, the relation between appearance and behavior in creating a subjective impression of familiarity or human presence has not been well explored, nor how habituation affects that impression.

This paper attempts to explore one possible explanation of the uncanny valley — that when a humanlike robot elicits an eerie sensation it is because the robot is acting as a reminder of mortality. It attempts to test this hypothesis through the experimental methods used by terror management theory (TMT). TMT studies have correlated subliminal reminders of mortality with a wide range of attitude changes. If an android affects people's attitudes without them knowing it, this raises ethical concerns that need to be addressed. If, however, an android is a conscious reminder of death, this could impede its future adoption, although people would likely habituate to the effect to some extent. In either case, the looks or movement of the device would need to be enhanced to prevent unwanted effects.

Terror Management Theory

Like other species *Homo sapiens* are highly motivated to avoid dying. Yet unlike other species they are in the potentially terrifying position of knowing that death is inevitable. Inspired by Ernest Becker's *The Denial of Death* [Becker, 1973] and other works, for more than two decades Jeff Greenberg, Tom Pyszczynski, Sheldon Solomon, and their colleagues have been developing a theory concerning how human beings manage their fear of personal extinction [Solomon et al., 1998] [Greenberg et al., 1986]. The theory has been supported by more than 200 experiments. They posit a dual-process model. Conscious thoughts of death are either suppressed (e.g., by thinking about something else) or their immediate significance is rationalized away (e.g., "My grandmother lived to be 90.") [Pyszczynski et al., 1999]. Nonconscious thoughts of death elicit defense processes that mitigate anxiety concerning the certainty of death by supporting a person's worldview and self-esteem:

> Along with the evolutionary emergence of cognitive abilities that enabled members of our species to comprehend our own mortality, our ancestors developed a solution to the problem of death in the form of a dual-component cultural anxiety buffer consisting of (a) a cultural world-
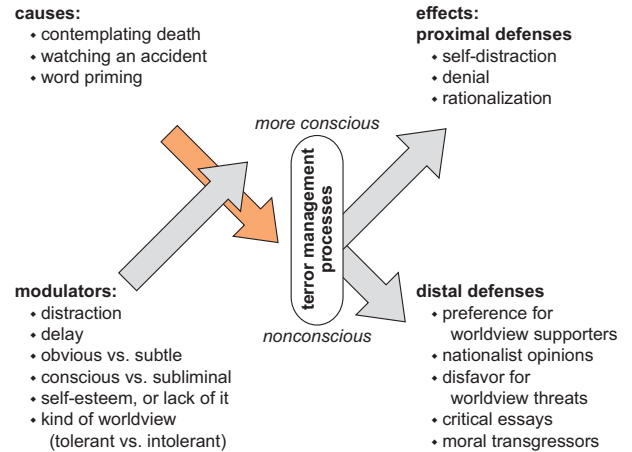

Figure 2: Terror management theory explores the relationship between reminders of death and the defense processes they elicit, including the modulating effects of intervening treatments.

view — a humanly constructed symbolic conception of reality that imbues life with order, permanence, and stability; a set of standards through which individuals can attain a sense of personal value; and some hope of either literally or symbolically transcending death for those who live up to these standards of value; and (b) self-esteem, which is acquired by believing that one is living up to the standards of value inherent in one's cultural worldview. [Pyszczynski et al., 1999]

Pyszczynski et al. (1999) contrast the *proximal* terror management defenses elicited by conscious thoughts of death with subliminally-elicited *distal* defenses. Distal defenses may address the threat at a level of abstraction different from that at which it is perceived and understood. Since distal defenses operate outside of consciousness (or at its fringes), they need not be rationally connected to the threat and may be best described as *experiential* in nature [Simon et al., 1997].

The *mortality salience hypothesis* predicts that, if having a worldview guards people from anxiety about the inevitability of death (e.g., by giving a literal or symbolic explanation of how death is transcended), those who have been subliminally reminded of death will react more favorably to information that supports their worldview and less favorably to information that undermines it. The hypothesis has been supported by numerous experiment, which have shown, for example, that mortality salience causes people to more strongly prefer essays that praise their country to those that criticize it [Greenberg et al., 1990] [Greenberg et al., 1994] [Greenberg et al., 2000], to prefer charismatic candidates over relationship-oriented candidates [Cohen et al., 2004], and to judge moral transgressors more harshly [Rosenblatt et al., 1989].

Such distal defenses as worldview protection are active when thoughts of death are not conscious but still accessible, such as when subjects have been distracted from death-related thoughts or after a period of delay

[Greenberg et al., 1994] [Greenberg et al., 2000]. They are, however, immediate for such subliminal priming as when the word *death* is flashed between the appearance of two other words for an interval too brief to result in one's conscious awareness of it [Arndt et al., 1997]. Although a fear of death can produce affective and physiological responses, evidence suggests that these responses do not mediate distal defenses; rather distal defences can occur in their absence [Pyszczynski et al., 1999].

Appraising human likeness by means of terror management defenses

A basic research question concerns whether very human-like stimuli sometimes cause an eerie sensation because they remind us of death and mortality, either consciously or subliminally. For example, an android that is not animated — or not animated like a living person — may look dead. This may remind us, if only subconsciously, of the fact that we too shall die, thus setting in motion defensive mechanisms that influence our attitudes in characteristic ways. If so, we can measure these changes of attitude to explore the terrain of uncanny valley.

More specifically, the mortality salience hypothesis predicts that nonconscious but accessible thoughts of death will result in distal defenses resulting in a heightened preference for stimuli that support a person's worldview and a decreased preference for stimuli that threaten it. If the appearance or behavior of a very humanlike robot, to the extent that it is uncanny, elicits proximal or distal terror management defenses, the effects of these defenses provide a means of quantitatively appraising the human likeness of its appearance and behavior. This then places the focus on the causes of TMT defenses (see Fig. 2). So while much research on terror management explores the range of manifestations of terror management defenses (e.g., "Will people who have been reminded of their mortality be more likely to judge moral transgressors harshly?"), the current research assumes the manifestations past studies have correlated with mortality salience are valid indicators of worldview defense, and then considers the range of stimuli that elicit them.

Experiment: Does an Uncanny Appearance Elicit Distal Defenses?

The experiment was designed to test the hypothesis that an android with uncanny appearance elicits the same distal defenses that reminders of death do. The evaluation criteria are derived from known mortality salience effects in the terror management theory literature: a heightened preference for charismatic politicians relative to relationships-oriented ones [Cohen et al., 2004] and a heightened preference for foreign students who praise a participant's country relative to those who criticize it [Greenberg et al., 1990] [Greenberg et al., 1994] [Greenberg et al., 2000]. In addition, mortality salience is gauged by word completion puzzles that are expected to show a participant's preference for death-related word completions indicative of the nonconscious activation of death-related associations.

Experimental image   Control image

Figure 3: The image on the left is the visual stimulus used for the experimental group. It is the head, neck, and upper torso of an android robot. The eyes are turned up and there is a gap between the eyes and eye lids because this part of the android has been powered off and disconnected from the rest of its body. The image on the right is the visual stimulus used for the control group. It depicts an Asian female in her early 20s.
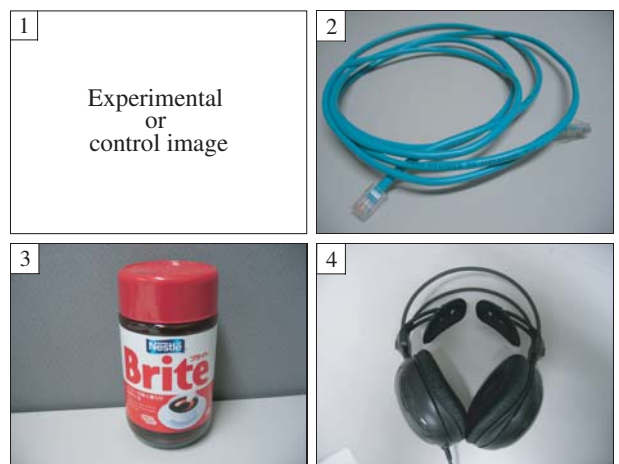

Figure 4: Participants sequentially viewed either the experimental image or the control image shown in Fig. 3 and then three other images (2–4). The images were 460 pixels in height and were captioned female figure, ethernet cable, nondairy creamer, and headphones.

Method

Participants. There were 63 participants, 25 male and 38 female, of whom 17 were 16 to 20 years old, 18 were 21 to 25, 9 were 26 to 30, 11 were 31 to 40, and 8 were over 40. All participants could communicate in English with 48 growing up in English speaking countries (5 Australia, 3 Canada, 1 Ireland, 1 New Zealand, 7 United Kingdom, and 32 United States) and 14 growing up in non-English speaking countries (1 Austria, 4 Israel, 1 Italy, 2 Korea, 1 the Netherlands, 1 Portugal, 2 Spain, 1 Turkey, and 1 Yugoslavia). Participants were recruited from Zone.com, an online gaming site. Four participants did not submit results for the word completion puzzles only. The participants were all volunteers and none received remuneration.

Procedure

Instructions. The solicitation for the experiment explained that (1) it involves filling out an online questionnaire; (2) it is for research on a cognitive mechanism that

is common to all people; (3) the participant's abilities would not be evaluated; and (4) further details concerning its purpose will be revealed only after the questionnaire has been completed. Potential participants were also told (5) it takes about 10 minutes to complete the questionnaire; (6) it must be completed in order and in one uninterrupted sitting; and (7) they should relax and just give the first answer they think of. Those who agreed to participate were given a link to the questionnaire website.

The website reiterated points 5 to 7 above and summarized the contents of the questionnaire: "You'll be shown some pictures, and you'll be asked some questions to see what you remember about them. Then you will be asked about a couple of excepts from political speeches and comments made by foreign students. Then you will solve some word puzzles." The wording of the questionnaire was intentionally informal because past studies have found that an informal experimental setting is more conducive to mortality salience effects [Simon et al., 1997], perhaps because participants tend to follow base their judgments on gut feelings rather than rational arguments. The computer, session, and starting time were uniquely identified to ensure that the same individual had not filled out the questionnaire more than once, and indeed nobody had.

Participants were asked their gender, approximate age, the country where they grew up, and then had to consent to the experiment: "This questionnaire is voluntary, so you may quit at any time. Data collected may be used in future studies, but it will be stripped of personal information. Clicking *I consent* means you want to go ahead."

Group assignment and stimuli. Participants were then randomly assigned with equal probability to either an experimental group or a control group. There were 31 participants in the experimental group and 32 in the control group. Those in the experimental group viewed the uncanny image of an android, while those in the control group viewed the image of a young Asian female (see Fig. 3). (Perhaps the main reason the android looked uncanny was because the eyes were looking up and sunken because the android was powered off and separated from its base.) In all other respects, the questionnaire was identical for both the experimental and control group. The subjects then viewed in sequence three "neutral" images. For each of the four images, they were instructed to view the image for a couple of seconds and then press a button labeled *view next image.*

Delay. The participants were then asked eight questions about the images. The answers to these questions were not kept. The questions served two functions: they distracted participants about the true purpose of the questionnaire since they concerned recall, and they added a delay before the questions relevant to terror management theory. Past TMT research has found that mortality salience effects appear immediately after subliminal priming on death but only after a delay when death is perceived consciously. Without knowing in advance whether the android would serve as a reminder of death and, if so, whether participants would be conscious of it as such, it was though prudent to insert a delay.

Worldview-related questions. Participants were next asked to read campaign speeches from two political candidates and to rate on a nine-point scale how well they liked each candidate and how insightful they thought each candidate was. They were then asked which candidate they would vote for. The first speech was *charismatic* and the second was *relationship-oriented* (see Appendix A). The speeches were loosely paraphrased from a previous study that indicated subjects in whom a subconscious fear of death has been elicited are more likely to prefer charismatic leaders [Cohen et al., 2004].

The same five questions were repeated for two foreign students who commented on their experience living in the participant's home country: Participants had to rate on a nine-point scale how well they liked each student, how insightful each student was, and which student they would support if both were running for president of the student government. The first student praised the participants' country, while the second student criticized the participants' country. These questions were motivated by a previous study that indicated subjects in whom a subconscious fear of death has been elicited are more likely to prefer people who support their worldview [Greenberg et al., 1990] [Greenberg et al., 1994] [Greenberg et al., 2000].

Word completion puzzles. Participants were next given 35 word completion puzzles of the following form:

$$\text{RELA}\underset{?}{\text{\_\_}}\text{G}$$

A button under each puzzle reads, "Give yourself three seconds to think of the missing letter with a ? under it, and then click here." After clicking the button, the puzzle vanishes and several choices would appear among which the participant may select only one. In the above puzzle, for example, a participant might select X to signify *relaxing*:

◯ T ⦿ X ◯ Y ◯ other / don't know

The participants would then be taken to the next question.

Following the TMT literature, dispersed among this set of 35 puzzles are 7 that allow participants to choose among word completions, one of which is related to death. These puzzles are intended to detect a subconscious activation of death-related concepts. The puzzles in the questionnaire are listed below with italics denoting the death-related option: COFF--: *coffin,* coffee; SK--L: skill, *skull*; MUR--R: murmur, *murder*; GRA--: grace, grade, grate, *grave*, graze; BUR-E-: burden, burger, *buried*, burned/burner, burped, burred; -EAD: bead, *dead*, head, lead, mead, read; STI--: stick, *stiff*, still/stile/stilt, stink/stint/sting.

A further 7 questions are intended to detect a subconscious activation of concepts that are roughly synonymous with the uncanny. The puzzles are listed below with with italics denoting the option related to the uncanny: WEI--: weigh, *weird*; --EEPY: *creepy,* sleepy; -DD: add, *odd*; UN-A--Y: *uncanny*, unhappy, unmanly; ----LIAR: familiar, *peculiar*; ST--GE: storage, *strange*; OM--US: *ominous*, omnibus.
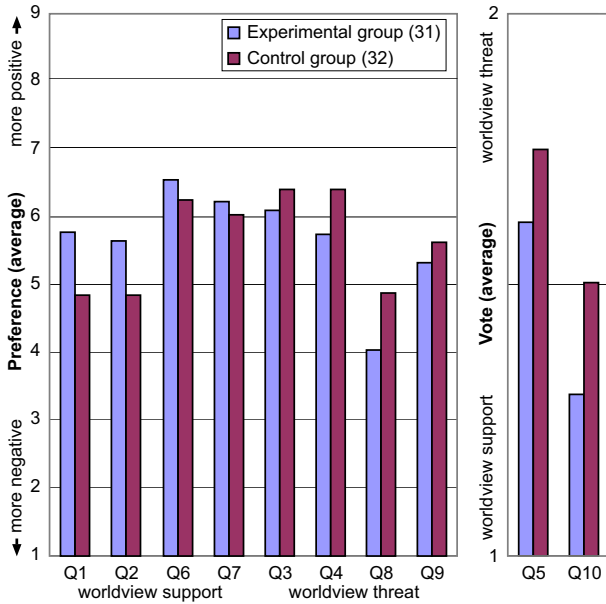
Figure 5: The experimental group shows more of a preference for the charismatic candidate (questions 1 and 2) and less of a preference for the relationship-oriented candidate (question 3 and 4) relative to the control group. They were also less likely to vote for the relationship-oriented candidate (question 5). This reproduces part of the results of [Cohen et al., 2004] but using an android as the experimental stimulus. The experimental group likewise shows more of a preference for the foreign student who praises their home country (questions 6 and 7) than the one who criticizes it (questions 8 and 9) relative to the control group, and they would vote for the praising student by a wide margin (question 10). This reproduces part of the results of [Greenberg et al., 1990] [Greenberg et al., 1994] [Greenberg et al., 2000] but using an android as the experimental stimulus.

The remaining 23 puzzles are unrelated to death and the uncanny. They are intended to disguise the true purpose of the experiment. Among 63 participants, 59 submitted results for the word completion section.

Suspicion and qualitative remarks. Finally, the participants were asked whether they had any difficulty completing the questionnaire; whether they had any suspicion concerning what the questionnaire was about; and what their impression was of the four images shown at the beginning. Some participants in the experimental group were selected for further questions concerning their impression of the uncanny image of the android. The participants were finally debriefed concerning the purpose of the experiment.

Results

Worldview-related questions. The results show on average a consistent preference for worldview supporters and against worldview threats in the experimental group (see Table 1 and Fig. 5). The experimental group rated the charismatic political candidate nearly a point higher for likeability (+0.93) and insight (+0.80) and rated the relationships-oriented candidate lower on likeability

Table 1: Worldview: Average (median) values

Question	Experimental	Control
1	5.77 (5)	4.84 (5)
2	5.65 (6)	4.84 (5)
3	6.10 (6)	6.41 (7)
4	5.74 (6)	6.41 (7)
5	1.61 (2)	1.75 (2)
6	6.55 (7)	6.25 (6)
7	6.23 (6)	6.03 (6)
8	4.03 (4)	4.88 (5)
9	5.32 (5)	5.63 (6)
10	1.29 (1)	1.5 (1.5)

(−0.31) and insight (−0.66). The experimental group rated the foreign student who praised the partipants' country higher for likeability (+0.30) and insight (+0.19) and rated the one who criticized it lower on likeability (−0.84) and insight (−0.30). (For preference questions, 1 = strongly negative, 5 = neutral, and 9 = strongly positive.)

The charismatic candidate lost by 7 votes in the experimental group (12 to 19) but by 16 votes in the control group (8 to 24), more than double the margin. The praising and critical foreign students tied in the control group (16 to 16), while the praising foreign student won by a 13 vote margin in the experimental group (22 to 9). (For voting questions, 1 = worldview-supportive candidate and 2 = worldview-critical candidate.)

It is difficult to show strong statistical significance for any isolated question in the worldview section owing to a high variance in the data. The average standard deviation for questions 1 to 4 and 6 to 9 was 2.00 and for questions 5 and 10 it was 0.48. However, when the values were summed together with sign changes on questions that the mortality salience hypothesis predicts will be disfavored by the experimental group, Student's t-test (two tails, heteroscedastic) showed statistical significance among the candidate-related questions ($t = 0.0307$) and overall ($t = 0.0348$).

Word completion puzzles. Among the 28 participants in the experimental group, there were 49 death-related word completions as compared to 38 among the 31 participants in the control group, 85 uncanny-related word completions as opposed to 66 in the control group, and 134 combined death and uncanny-related word completions as opposed to 104 in the control group (see Table 2).

On average the experimental group had 0.524 more death-related word completions, 0.907 more uncanny-related word completions, and 1.430 more death and uncanny-related word completions than the control group (see Table 3).

Student's t-test (two tails, heteroscedastic) showed strong statistical significance uncanny-related questions ($t = 0.0132$), but not for death ($t = 0.0963$) related questions. For combined death and uncanny-related questions, the statistical significance was the highest

$(t = 0.00542)$.

The median time spent viewing the first image was 12 seconds, and the average time was 21.9 seconds. (The first image is the android head in the experimental group and the Asian woman in the control group.) The minimum time was 2 seconds, and the three longest times were 262, 127, and 67 seconds.

The median time spent in the worldview section of the questionnaire by the 63 participants was 336 seconds, and the average time was 350.9 seconds. The minimum time was 162 seconds, and the three longest times were 735, 604, and 593 seconds. No figures were recorded for the word completion section.

When asked at the end of the questionnaire, "What was your feeling about or impression of the four images shown at the beginning?" 16 participants did not comment on the android or other figures or said that they had no particular feeling. The remaining 12 appeared to be commenting on the android. An exhaustive list of comments about the android head are as follows:

1. Scary female image.
2. I thought the first wax head was really interesting.
3. The women figure kind of freaked me out a bit.
4. Strange.
5. The first two images were disturbing.
6. I thought it was kind of bizarre. The Japanese girl's eyes were messed-up looking.
7. The first image seemed to come across as very sickly. It made me feel sick just looking at it.
8. Weird lady.
9. The woman seemed to be dead. It was a shocking image and given the chance I would have preferred not to see it.
10. The woman was frightening.
11. Weird.
12. I thought it humorous that two of them were related to Asian things.

No one made explicit reference to the content of an image other than the image of the android. The participants were not shown the images again at the end of the questionnaire, so all these comments were based on memory.

Table 2: Word completion: Totals

Type	Experimental	Control
Subjects	28	31
Death	49	38
Uncanny	85	66
Combined	134	104

Table 3: Word completion: Average (median) values

Type	Experimental	Control
Death	1.75 (2)	1.23 (1)
Uncanny	3.04 (3)	2.13 (2)
Combined	4.79 (5)	3.35 (3)

Interviews

Several participants were interviewed through instant messaging. They were all asked, **"What do you see when you look at this image?"** Margot[1] and her children thought it looked dead. Craig and Judith agreed after being prompted. Judith at first thought it was a real person who had fainted. Penelope thought it could not look dead because it had never been alive. Derek and Loretta both thought it looked human at first; however, Derek did hot have an emotional response, whereas Loretta had a strong response. Clearly, there is considerable diversity of results in this small sample.

Penelope: A scary robot.

Do you see something death related?

No, I see something not natural, but nothing to do with death. It is something that never has been alive, so it cannot be dead. The skin, the opened mouth, the right eye make it seems like it had an overload of electricity or a short circuit.

But in the experiment it is registering death saliency.

Well no, the thing that makes it scary is that it is a copy of a person. That's the truth. But not a good copy. You know, it is like the old dummies in empty shops.

Is it still as scary as it was before?

Yes.

Judith: Oh, my God, it's scary. She look like she is fainting.

You said "she." What do you mean? What is "it"? What makes it scary?

It looks like a female. The look on the face looks like she is fainting.

But it isn't a real female, right?

It is a real female, isn't it?!

No. It is a robot.

What?

That is the head of a robot.

So that's why the eyes look like that. The eyes make it look like its fainting, Yes, it looks dead.

So looking dead makes it scary?

A bit, yes, for a first impression, of course. It makes a big difference. It look so real though.

Craig: It looks like a Korean or Chinese fake woman. It gives an ugly impression.

Does it seem strange, scary, or dead?

I dislike it. Scary and dead, yes.

Derek: Being an android myself, I can't say that I really have any feelings about it. But seriously, I don't know.

Does it seem uncanny or eerie? Alive or dead? In pain? Do you have any strange feeling?

I didn't think of it in those terms.

In what terms did you think of it?

---

[1]The names have been changed. Grammatical errors have been fixed. Participants had to message the following consent in addition to the consent of the experiment: "I consent that my answers may be used in a publication or presentation once stripped of identifying information."

I noticed features, but didn't form an impression of the significance of the features.

Do you mean you looked at it more scientifically or analytically? Not emotionally?

E.g., the eyebrows appear to be shaved in a "nonstandard" way. The eyes appear to be "askew." The skin tone is slightly mottled or uneven.

So you are not grossed out? You don't feel any emotional response?

No, of course not. It wasn't really an emotional reaction, or even a judgment of artificiality.

Does it look alive or dead?

I can't really tell.

Well, if it looks like an android, then I suppose neither?

It could go either way. I only said android because that's the name in the URL.

Oh. Well, what if that were not the name. Then what would you say?

I guess the default response would be to anthropomorphize it (or her, rather).

Human?

Yeah, probably. The default would probably start with human, pending further data. The teeth are imperfect, which might argue for human.

Loretta: I see a woman that does not look real, but there is pain in the face, I also see torture. It could just be inside though.

What is your impression of it?

I feel this is some sort of art work. I don't like the eyes, but the lips are beautiful.

How do you feel about it?

It makes me feel sorrow, sadness, and abuse. Around the neck and face area that can be seen. What are my real thoughts? This is not a real human!

What you say sound contradictory. Does it look human or mechanical?

Well, when I first looked at it, I saw torture and pain. That was my first impression. Then the more I looked, I could see it was not real and as you said now a machine. But all the hurt was there.

And does it seem alive, dead, or what?

Dead.

But at first did it look alive or dead?

At first I saw torture and then it looked dead.... I was wondering if I saw all that torture and pain because I'd been through it? What you think?

Margot: It reminds me of death. That's the first thing I thought of when I saw it, the first time I took the survey.

Oh, I see. And how do you feel about it? What do you think when you look at it? What emotions do you feel?

It's unpleasant to look at. Hmmm. Nervous, a bit fearful, uncomfortable.

I see. So for you it is definitely a reminder of death.

Yeah, it is. I asked my daughter what it reminded her of, and she said death too with no prompting. And she
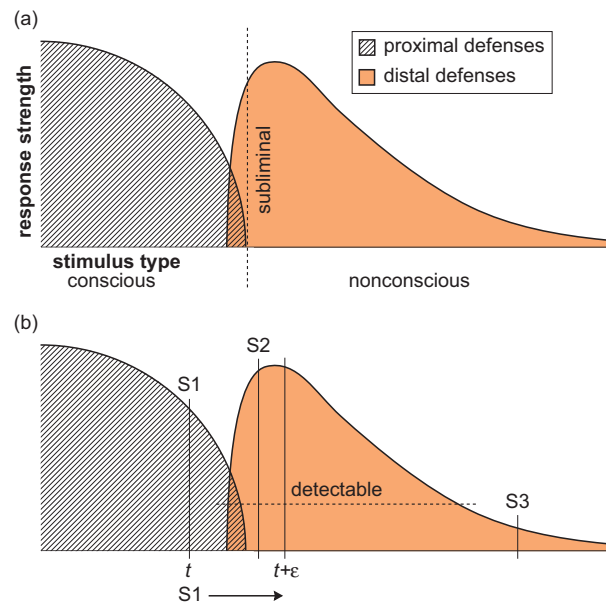

Figure 6: (a) Stimuli may be perceived in focal or fringe consciousness or subliminally. This can produce proximal or distal defenses with varying response strengths. (The curves in the figure are given only for the sake of example; their actual shapes are unknown.) (b) Different stimuli may elicit different kinds of defenses (S1 and S2) as may the same stimulus (S1) at different times $(t, t + \epsilon)$ owing to the effects of delay. In addition, for a give sample size, the effects of some stimuli may be impossible to measure owing to variance in the data (S3.)

says it reminded her of *The Grudge*, a Japanese horror film.

Oh, I saw it! Wow, you let her see that?!

No! Well, some parts like two minutes at the start. Then I sent her away. She didn't want to see anymore anyway. I'll ask my other daughter. One sec. I'll call her.

Great.

Yep, death also.

No prompting?

Nope, and this is also without hearing my other daughter's answer because she was on the phone.

Discussion

Most people had no difficulties with the questionnaire. A few people had to do extra scrolling in the word completion questions because they had missed a question. A few complained that the word completion puzzles were difficult, and a couple non-native speakers said that was owing to the language barrier. Only a few people had suspicions about the questionnaire, but their suspicions had nothing to do with what the questionnaire was actually about.

There was a minor flaw in the experimental procedure. It would have been useful to record the precise amount of time a participant spent on each part of the experiment so that the influence of delay could be considered in more detail.

The effects of distal terror management defenses may not be such a reliable indicator of the degree of mortality salience of a given stimulus. The same stimulus will affect the attitudes of individuals differently. (For example, in one study "low authoritarian individuals did not derogate attitudinally dissimilar others when mortality was made more salient," while high authoritarian individuals did ([Solomon et al., 1998], p. 40, citing an experiment in [Greenberg et al., 1992]). In addition, proximal and distal defenses have varying response strengths depending on whether the stimulus is perceived in focal or fringe consciousness or subliminally (Fig. 6(a)). The same stimulus can produce varying effects owing to delay, some of which will be too weak to detect owing to the high degree of variance in the data (Fig. 6(b)).

A more fundamental concern relates to affect. The eerie sensation identified with the uncanny valley may be characterized as affective, although it seems difficult to identify it with one or more primary emotions like disgust. Terror management studies have indicated that affect does not mediate distal defenses [Pyszczynski et al., 1999]. And it seems that subliminal priming of death-related words does not create an eerie sensation. That leaves the question of what nonconscious processes do underpin the eerie sensation. The uncanny android still seems to be a reminder of death, but perhaps a conscious reminder, in which case the distal (i.e., nonconscious) defenses that showed up in the experimental results may have occurred owing to delay.

Probably at the end of the study all participants should have been showed the uncanny android and asked to rate how eerie they felt it was and given a questionnaire to evaluate its emotional impact (e.g., PANAS), but already some were complaining about the questionnaire's length. In addition, another experiment could have measured the experimental stimulus's impact on such physiological responses as heart rate, respiration, and galvanic skin response.

The fact that androids are often not capable of satisfying the human-directed expectations they elicit may be one reason why we may perceive them to not be fully alive. However, it also suggests alternative explanations of the uncanny. If one person elicits expectations in another, that person elicits contextually-appropriate behavior (i.e., behavior that can be described in terms of norms). Given further expectations for real-time responding (regarding the responsive behavior that is appropriate in this context) androids tend to violate human expectations for how to go on. This suggests that some of the peculiarities of interacting with androids may be owing to failures to model the microbehavior central to the expectational cycle.

This opens the door to hypotheses about the uncanny valley that are unrelated to reminders of mortality. So an important question is whether there is something distinctive about expectations elicited by the human form that causes their violation to result in sensations that are qualitatively different from those elicited by other forms. Anecdotal evidence suggests that there is.

Conclusion

Many kinds of media (e.g., computers, robots, films) are capable of eliciting, to varying degrees, different kinds of social responses (e.g., verbal [Reeves and Nass, 1996], gestural, gaze [MacDorman et al., 2005]). Nevertheless, qualitative (and quantitative) differences emerge depending on the type of media and how it acts. The eerie feeling elicited by human-looking but not mechanical-looking robots is one such qualitative difference, and its significance is worth exploring.

This study has hypothesized that an uncanny-looking android may be uncanny because it elicits a fear of death, and it has attempted to verify this through questions designed to measure such distal terror management defenses as worldview protection. The results are favorable. On average the group exposed to an image of an uncanny robot consistently preferred information sources that supported their worldview relative to the control group.

The results, however, only apply to one particular stimulus, so it is important to ascertain whether they generalize across uncanny stimuli and, in particular, to uncanny movement in a robot that otherwise looks human and natural.

Appendix A:
Worldview-related Statements

Politicians. Two candidates are running for our nation's highest office. An excerpt from each candidate's campaign speech is given below:

"I will be the best leader of this great nation because I am committed to a brighter vision of our future. I set high standards for my cabinet and myself and expect them to work as hard as I do to achieve these standards. I want everyone, both public employees and private citizens, to do their best for our great nation, so that we can all achieve our full potential. My goal is to do things differently from my predecessors, and I am willing to take risks to show voters how things can be improved. You are not just ordinary citizens. You are part of a great nation, and by working together we can have a big impact." – J.N.

"I will be the best leader of this nation because I am concerned about people's welfare. I treat everyone with consideration and respect, no matter how high the political tensions may rise. I stress communication among my staff and the general public. I keep everyone informed about proposed legislation, and I am open to suggestions. I frequently have meetings with constituents to discuss policies. I encourage all citizens to take a role

in improving things because I know that each individual can make a difference. Everyone's contributions are recognized and appreciated." – B.E.

Foreign Students. Two foreign students comment on their experiences living in your country:

"I am a foreign student from Phnom Penh, which is the capital of Cambodia. I have been studying civil engineering at university in your country for three years. I love your country because there are many opportunities for me here that do not exist in my home country. Most people I have met have been kind to me, although they know I am a foreigner. I already had a positive impression of your country before coming here because your government has contributed much to the United Nations to support the clearance of landmines in Cambodia, and there are also many private citizens who either volunteer to work in Cambodia or make donations to support aid. After graduation, I would like to help develop my country too." – S.L.

"I am a foreign student from Bandung, which is the largest city in West Java, Indonesia. Last year, I started a Master's degree here. Although my course is challenging, I am surprised how ignorant other students are about other countries. That may be one reason the people have been mislead by your government, which follows short-sighted policies, which appear to be aligned with the interests of the United States, but which actually won't help anyone in the long run. If your country really wants to bring an end to terrorism, they should stop provoking Moslems with bellicose policies or propping up autocratic sheiks." – H.M.

Appendix B: *The Uncanny Valley*

Valley of familiarity

There are mathematical functions of the form $y = f(x)$ for which the value of y increases (or decreases) continuously with the value of x. For example, as the effort x increases, income y increases, or as a car's accelerator is pressed, the car moves faster. This kind of relationship is ubiquitous and easily understood. In fact, it covers most phenomena, so we might think that this function can represent all relations. That is why people are usually puzzled when faced with some phenomenon it cannot represent.

Climbing a mountain is an example of a function that does not increase continuously: a person's altitude y does not always increase as the distance from the summit decreases owning to the intervening hills and valleys. I have noticed that, as robots appear more humanlike, our sense of their familiarity increases until we come to a valley. I call this relation the *uncanny valley*.

Recently there are many industrial robots, and as we know the robots do not have a face or legs, and just rotate or extend or contract their arms, and they bear no resemblance to human beings. Certainly the policy for designing these kinds of robots is based on function-ality. From this standpoint, the robots must perform functions similar to those of human factory workers, but their appearance is not evaluated. If we plot these industrial robots on a graph of familiarity versus appearance, they lie near the origin (see Fig. 1). So they bear little resemblance to a human being, and in general people do not find them to be familiar. But if the designer of a toy robot puts importance on a robot's appearance rather than its function, the robot will have a somewhat humanlike appearance with a face, two arms, two legs, and a torso. This design lets children enjoy a sense of familiarity with the humanoid toy. So the toy robot is approaching the top of the first peak.

Of course, human beings themselves lie at the final goal of robotics, which is why we make an effort to build humanlike robots. For example, a robot's arms may be composed of a metal cylinder with many bolts, but to achieve a more humanlike appearance, we paint over the metal in skin tones. These cosmetic efforts cause a resultant increase in our sense of the robot's familiarity. Some readers may have felt sympathy for handicapped people they have seen who attach a prosthetic arm or leg to replace a missing limb. But recently prosthetic hands have improved greatly, and we cannot distinguish them from real hands at a glance. Some prosthetic hands attempt to simulate veins, muscles, tendons, finger nails, and finger prints, and their color resembles human pigmentation. So maybe the prosthetic arm has achieved a degree of human verisimilitude on par with false teeth. But this kind of prosthetic hand is too real and when we notice it is prosthetic, we have a sense of strangeness. So if we shake the hand, we are surprised by the lack of soft tissue and cold temperature. In this case, there is no longer a sense of familiarity. It is uncanny. In mathematical terms, strangeness can be represented by negative familiarity, so the prosthetic hand is at the bottom of the valley. So in this case, the appearance is quite human like, but the familiarity is negative. This is the uncanny valley.

I don't think a *bunraku* puppet is similar to human beings on close observation. Its realism in terms of size, skin, and so on, does not approach that of a prosthetic hand. But when we enjoy a puppet show in the theater, we are seated far from the puppets. Their absolute size is ignored, and their total appearance including eye and hand movements is close to that of human beings. So although the puppets' bodies are not humanlike, we can feel that they are humanlike because their bodies and movements when taken together are humanlike. And from this evidence I think their familiarity is very high.

From the above maybe readers can understand the concept of the uncanny valley. So in the following I consider the relationship between movement and the uncanny valley.

The effects of movement

For creatures, including robots, movement is generally a sign of life. Adding movement changes the shape of the uncanny valley by exaggerating the peaks and valley (see Fig. 1). For the industrial robot, the impact of

movement is relatively slight because we see it as just a machine. If it stops moving, it becomes a mere oily machine. But if programmed properly to generate humanlike movements, we can enjoy some sense of familiarity. Humanlike movement requires similarity of velocity and acceleration. Conversely, if we add movement to a prosthetic hand, which is at the bottom of the uncanny valley, our sensation of strangeness grows quite large. Some readers may know that recent technology has enabled prosthetic fingers to move automatically. A commercially available prosthetic hand made with the highest technique was developed in Vienna. To explain how it works, the intention to move the forearm, even if missing, produces current in the arm muscles that can be detected by an electromyogram. So the prosthetic hand detects the current by means of electrodes and amplifies the signal to activates a small motor in the prosthetic arm to move the fingers. This hand can move in a way that causes some ordinary people to feel uneasy. If you shook a woman's hand with this hand in a dark place, the woman must be shocked!

Since these effects are apparent for just a prosthetic arm, the strangeness will be magnified if we build an entire robot. You can imagine going to a work place where there are many mannequins: if a mannequin started to move, you might be shocked. This is a kind of horror.

In the World Expo held in Osaka this year, the robots displayed a more elaborate design. For example, one robot has 29 artificial muscles in the face to make humanlike facial expressions. According to the designer, laughing is a kind of sequence of face distortions, and the distortion speed is an important factor. If we cut the speed in half, laughing looks unnatural. This illustrates how slight variations in movement can cause a robot, puppet, or prosthetic hand to tumble down into the uncanny valley.

Escape by design

We hope to design robots or prosthetic hands that will not fall into the uncanny valley. So I recommend designers take the first peak as the goal in building robots rather than the second. Although the second peak is higher, there is a far greater risk of falling into the uncanny valley. We predict that it is possible to produce a safe familiarity by a nonhumanlike design. So designers please consider this point. A good example is glasses. Glasses do not resemble the real eyeball, but this design is adequate and can make the eyes more charming. So we should follow this principle when we design prosthetic eyes. We can create an elegant prosthetic hand - one that must be fashionable. Artist who makes statues of Buddhas created a model of a human hand that is made from wood. The fingers bend at their joints. The hand has no finger print, and it assumes the natural color of wood. But we feel it is beautiful and there is no sense of the uncanny. Maybe wooden hand can serve as a reference for future design.

The significance of the uncanny

In Fig. 1, a healthy person is at the top of the second peak. And when we die, we fall into the trough of the uncanny valley. Our body becomes cold, our color changes, and movement ceases. Therefore, our impression of death can be explained by the movement from the second peak to the uncanny valley as shown by the dashed line in the figure. We might be happy this line is into the still valley of a corpse and that of not the living dead! I think this explains the mystery of the uncanny valley: Why do we humans have such a feeling of strangeness? Is this necessary? I have not yet considered it deeply, but it may be important to our self-preservation.

We must complete the map of the uncanny valley to know what is human or to establish the design methodology for creating familiar devices through robotics research.

References

[Arndt et al., 1997] Arndt, J., Greenberg, J., Pyszczynski, T., and Solomon, S. (1997). Subliminal presentation of death reminders leads to increased defense of the cultural worldview. *Psychological Science*, 8:379–385.

[Becker, 1973] Becker, E. (1973). *The denial of death*. Free Press, New York.

[Cohen et al., 2004] Cohen, F., Solomon, S., Maxfield, M., Pyszczynski, T., and Greenberg, J. (2004). Fatal attraction: The effects of mortality salience on evaluations of charismatic, task-oriented, and relationship-oriented leaders. *Psychological Science*, 15:846–851.

[Greenberg et al., 1986] Greenberg, J., Pyszczynski, T., and Solomon, S. (1986). The causes and consequences of the need for self-esteem: A terror management theory. In Baumeister, R. F., editor, *Public self and private self*, pages 189–212, New York. Springer-Verlag.

[Greenberg et al., 1990] Greenberg, J., Pyszczynski, T., Solomon, S., Rosenblatt, A., Veeder, M., Kirkland, S., and Lyon, D. (1990). Evidence for terror management theory ii: The effects of mortality salience reactions to those who threaten or bolster the cultural worldview. *Journal of Personality and Social Psychology*, 58:308–318.

[Greenberg et al., 1994] Greenberg, J., Pyszczynski, T., Solomon, S., Simon, L., and Breus, M. (1994). Role of consciousness and accessibility of death-related thoughts in mortality salience effects. *Journal of Personality and Social Psychology*, 67:627–637.

[Greenberg et al., 2000] Greenberg, J., Simon, L., Arndt, J., Pyszczynski, T., and Solomon, S. (2000). Proximal and distal defenses in response to reminders of ones mortality: Evidence of a temporal sequence. *Personality and Social Psychology Bulletin*, 26:91–99.

[Greenberg et al., 1992] Greenberg, J., Simon, L., Pyszczynski, T., Solomon, S., and Chatel, D. (1992). Terror management and toleraance: Does mortality salience always intensify negative reactions to others

who threaten one's worldview? *Journal of Personality and Social Psychology*, 62:212–220.

[MacDorman et al., 2005] MacDorman, K., Minato, T., Shimada, M., Itakura, S., Cowley, S., and Ishiguro, H. (2005). Assessing human likeness by eye contact in an android testbed. In *Proceedings of the XXVII Annual Meeting of the Cognitive Science Society*, Stresa, Italy.

[Matsui et al., 2005] Matsui, D., Minato, T., MacDorman, K., and Ishiguro, H. (2005). Generating natural motion in an android by mapping human motion. In *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems*, Edmonton, Canada.

[Minato et al., 2004] Minato, T., Shimada, M., Ishiguro, H., and Itakura, S. (2004). Development of an android robot for studying human-robot interaction. In Orchard, R., Yang, C., and Ali, M., editors, *Innovations in Applied Artificial Intelligence*, pages 424–434, Berlin.

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

[Pyszczynski et al., 1999] Pyszczynski, T., Greenberg, J., and Solomon, S. (1999). A dual-process model of defense against conscious and unconscious death-related thoughts: An extension of terror management theory. *Psychological Review*, 106(4):835–845.

[Reeves and Nass, 1996] Reeves, B. and Nass, C. (1996). *The media equation: How people treat computers, television, and new media like real people and places.* Cambridge University Press (CSLI), New York.

[Rosenblatt et al., 1989] Rosenblatt, A., Greenberg, J., Solomon, S., Pyszczynski, T., and Lyon, D. (1989). Evidence for terror management theory: I. the effects of mortality salience on reactions to those who violate or uphold cultural values. *Journal of Personality and Social Psychology*, 57:681–690.

[Simon et al., 1997] Simon, L., Greenberg, J., Harmon-Jones, E., Solomon, S., Pyszczynski, T., Arndt, J., and Abend, T. (1997). Terror management and cognitive-experiential self-theory: Evidence that terror management occurs in the experiential system. *Journal of Personality and Social Psychology*, 72(5):1132–1146.

[Solomon et al., 1998] Solomon, S., Greenberg, J., and Pyszczynski, T. (1998). Tales from the crypt: The role of death in life. *Zygon: Journal of Religion and Science*, 33:9–43.