Review of General Psychology
2015, Vol. 19, No. 4, 393–407

1089-2680/15/$12.00 http://dx.doi.org/10.1037/gpr0000056

The Uncanny Valley: Existence and Explanations

Shensheng Wang, Scott O. Lilienfeld, and Philippe Rochat
Emory University

More than 40 years ago, Japanese roboticist Masahiro Mori (1970/2005) proposed the "uncanny valley" hypothesis, which predicted a nonlinear relation between robots' perceived human likeness and their likability. Although some studies have corroborated this hypothesis and proposed explanations for its existence, the evidence on both fronts has been mixed and open to debate. We first review the literature to ascertain whether the uncanny valley exists. We then try to explain the uncanny phenomenon by reviewing hypotheses derived from diverse theoretical and methodological perspectives within psychology and allied fields, including evolutionary, social, cognitive, and psychodynamic approaches. Next, we provide an evaluation and critique of these studies by focusing on their methodological limitations, leading us to question the accepted definition of the uncanny valley. We examine the definitions of human likeness and likability, and propose a statistical test to preliminarily quantify their nonlinear relation. We argue that the uncanny valley hypothesis is ultimately an engineering problem that bears on the possibility of building androids that may some day become indistinguishable from humans. In closing, we propose a dehumanization hypothesis to explain the uncanny phenomenon.

Keywords: animacy, dehumanization, humanness, statistical test, uncanny valley

Over the centuries, machines have increasingly assumed the heavy duties of dangerous and mundane work from humans and improved humans' living conditions. Nevertheless, humans have often described machines, including robots, as "cold" and difficult to interact with. To improve human–robot interaction (HRI), engineers have designed androids, which resemble humans highly and may eventually even become indistinguishable from them (MacDorman & Ishiguro, 2006). Because such androids are more humanlike than other machines, we may be able to relate to, and even feel an emotional connection, to them. As androids increasingly resemble humans physically, however, some of them may fall into an "uncanny valley" (Mori, 1970/2005), in which they elicit repulsive affects.

Masahiro Mori (1970/2005), a Japanese roboticist, coined the term "uncanny valley" in an essay *Bukimi No Tani*. There, he wrote that:

> 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 owing 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." (p. 33)

Although Mori proposed the uncanny valley hypothesis 45 years ago, the concept did not draw researchers' attention for nearly 30 years. Nevertheless, in popular culture, the uncanny valley hypothesis gained traction as androids and realistic computer-generated characters (avatars), such as those in the film *The Polar Express* (Zemeckis, 2004), reportedly elicited unease in some human observers (Geller, 2008; Kaba, 2013). In 2011, the hamburger franchise *Burger King* retired its huge "King" humanoid mascot from TV advertising, apparently in part because many viewers regarded it as "creepy" (Daily Mail Reporter, 2011). The uncanny valley may also account partly for the widespread clinical phenomena of coulrophobia (fear of clowns; see Moore, 2012) and pediophobia (fear of dolls; Macy & Schrader, 2008) given that clowns and dolls often appear quasihuman. Perhaps not coincidentally, dolls, mannequins, wooden dummies, facially distorted human imposters, and human-like robots featured prominently in the enormously popular 1960s TV series "The Twilight Zone," which was renowned for eliciting powerful feelings of eeriness in viewers (Wolfe, 1997). Still, in many cases such feelings may be inadvertent or unintended. Engineers and filmmakers are especially concerned that their designs may fall into the uncanny valley and have made concerted efforts to avoid it (Geller, 2008).

Nevertheless, researchers have struggled with validating Mori's uncanny valley hypothesis. This hypothesis posits that as a robot increasingly resembles a person, its familiarity increases until a point at which it abruptly drops to a negative value and elicits strong repulsion; then, as the robot's resemblance to a person continues to increase, its familiarity increases again and eventually reaches the level of a person. Although this notion has been widely discussed and has considerable intuitive appeal, it is unclear whether the uncanny valley is a real phenomenon or merely a plausible conjecture.

The current literature review focuses on two overarching questions: First, is the uncanny valley a real phenomenon? Second, how can we explain it? In this study, we first try to determine the existence of this valley by reviewing the relevant studies. We then try to explain it by reviewing hypotheses derived from diverse theoretical and methodological perspectives within psychology

Shensheng Wang, Scott O. Lilienfeld, and Philippe Rochat, Department of Psychology, Emory University.

Correspondence concerning this article should be addressed to Shensheng Wang, Department of Psychology, Emory University, 36 Eagle Row, Atlanta, GA 30322. E-mail: shensheng.wang@emory.edu

and allied fields, including evolutionary, social, cognitive, and psychodynamic approaches.

The Existence of the Uncanny Valley

Past research has tested the uncanny valley hypothesis primarily by detecting the hypothetical curve that would be anticipated if this valley exists (see Figure 1 in Mori, 1970/2005). By plotting participants' ratings on familiarity or eeriness against the ratings on the human likeness of human replicas (e.g., androids and computer-generated characters), researchers initially appeared to corroborate the uncanny valley hypothesis (MacDorman & Ishiguro, 2006; Seyama & Nagayama, 2007). Using morphed images, which ranged from a humanoid robot to an android and to the android's human inspiration, MacDorman and Ishiguro (2006) detected nonlinear relations between images' human likeness and familiarity or eeriness, which resembled the uncanny valley. Employing an asynchronous morphing technique (in which images were morphed in two steps, with the eyes morphed first, and the head morphed next, or in the reversed order) to morph a doll face into a human face, Seyama and Nagayama (2007) also detected a curve that resembled the uncanny valley. In this study, participants rated their impression of a series of morphed images on a 5-point scale of pleasantness ranging from −2 (*extremely unpleasant*) to + 2 (*extremely pleasant*). The instructions defined pleasantness in terms of "representing any emotionally positive adjective such as attractive, pretty, natural, healthy, intimate, or elegant" (Seyama & Nagayama, 2007, p. 340). Similarly, they defined unpleasantness in terms of "representing any emotionally negative adjective such as unattractive, fearful, ugly, abnormal, sick, or inelegant" (p. 340). The researchers plotted the ratings on pleasantness against the human likeness of the morphed images, which was operationalized in terms of the proportion of human face in the morphed images (i.e., morphing ratio). A recent study (Mathur & Reichling, 2016) examined participants' ratings of 80 real-world android faces. By plotting the ratings on likability against human likeness, the researchers detected a curve resembling the uncanny valley.

At the same time, evidence in support of the uncanny valley has been mixed and at times contradictory. Although some studies have shown that perceived human likeness and likability of human replicas follow a nonlinear relation, others have failed to detect a curve that resembles the uncanny valley (Bartneck, Kanda, Ishiguro, & Hagita, 2007; Hanson, 2005; MacDorman, 2006; Poliakoff, Beach, Best, Howard, & Gowen, 2013; Seyama & Nagayama, 2007). Evidence against the existence of the uncanny valley typically derives from studies that rated images or videos of existing human replicas (Bartneck et al., 2007; MacDorman, 2006; Poliakoff et al., 2013, for an exception, see Mathur & Reichling, 2016), whereas support derives from some studies that instead used morphed images. Nevertheless, this methodological difference does not appear to explain why some studies detected the hypothetical curve whereas others did not. Using morphed images, for example, Seyama and Nagayama (2007, Study 1) and MacDorman and Ishiguro (2006) obtained opposing findings. These inconsistent findings may in part be attributable to the variability in the human replicas these studies selected, which ranged from mechanistic-looking robots, to humanoids, dolls, and computer-generated characters (avatars).

In addition to the previously mentioned studies, which detected the hypothetical curve, researchers have shown that combinations of mismatched features in robots such as voice (Mitchell et al., 2011) and facial expressions (Tinwell, Grimshaw, Nabi, & Williams, 2011) with appearance can elicit the uncanny feeling. A study also revealed that macaque monkeys are more attentive to unrealistic synchronized and real monkey faces than to realistic synchronized monkey faces, raising the possibility of an uncanny valley in nonhuman primates (Steckenfinger & Ghazanfar, 2009). Finally, an infant study showed that by 12 months, infants looked longer at human or avatar faces than at the same avatar faces that were combined with uncanny features, such as disproportionately large eyes (Lewkowicz & Ghazanfar, 2012). This finding suggests that infants' experience with real human faces may play a critical role in the emergence of the uncanny phenomenon.

In summary, although the notion of the uncanny valley is plausible and is supported by plentiful anecdotal evidence, rigorous controlled studies have yielded mixed support for its existence. Whether androids' uncanniness is a function of their human likeness as the uncanny valley hypothesis predicts is debatable.

Explanations of the Uncanny Phenomenon

Researchers have proposed a variety of explanations to account for the uncanny phenomenon (i.e., that some human replicas reportedly elicit the uncanny feeling), including evolutionary, social, cognitive, and psychodynamic approaches. These hypotheses fall mainly into two broad categories: One category views the uncanny phenomenon as an "automatic, stimulus-driven, specialized processing that occurs early in perception," whereas the other views the phenomenon as "a broader and more general range of cognitive processing that occurs later" (MacDorman, Green, Ho, & Koch, 2009, p. 696).

The Uncanny Phenomenon as Perceptual Processing

Mori (1970/2005) suggested that the uncanny feeling elicited by the prosthetic hand was attributable to the contrast between its visual realism and its tactile imperfections. Researchers have shown that human observers are more sensitive to facial proportions in more realistic faces than in less realistic faces (Green, MacDorman, Ho, & Vasudevan, 2008; MacDorman et al., 2009). Nevertheless, researchers have disagreed regarding the mechanisms whereby this enhanced sensitivity to facial proportions in realistic faces could elicit the uncanny feeling. The *Pathogen Avoidance* hypothesis (MacDorman et al., 2009; MacDorman & Ishiguro, 2006) proposes that uncanny faces are indicative of a heightened risk for transmissible diseases, which trigger an evolved mechanism for pathogen avoidance, thereby eliciting disgust. The *Mortality Salience* hypothesis (Ho, MacDorman, & Pramono, 2008; MacDorman & Ishiguro, 2006) suggests that uncanny faces remind human observers of their own inevitable mortality, thereby provoking fears of death. The *Evolutionary Aesthetics* hypothesis posits that humans are highly sensitive to facial aesthetics evolved under selective pressures and that uncanny faces fail to meet those aesthetic standards, thereby eliciting the uncanny feeling (Hanson, 2005; MacDorman & Ishiguro, 2006). We review each hypothesis in turn.

Pathogen Avoidance hypothesis. Why do humans have the uncanny feeling? Mori (1970/2005) believed that the uncanny feeling may play a crucial role in the self-preservation of humans.

Based on evolutionary psychology, and in particular Rozin and colleagues' classic research on disgust (Rozin & Fallon, 1987), Christian Keysers linked the uncanny phenomenon to an evolved mechanism for pathogen avoidance (MacDorman et al., 2009; MacDorman & Ishiguro, 2006). He argued that the uncanny feeling is rooted in the basic emotion of disgust. Specifically, the perceived imperfections in a human replica elicit disgust because the pathogen avoidance mechanism "interprets" these defects as harbingers of transmissable diseases. Because more humanlike entities are perceived to be more closely related to humans genetically, their imperfections elicit stronger aversion than do mechanical-looking entities. Nevertheless, MacDorman and colleagues (2009) also noted that the uncanny feeling might be driven by fear rather than disgust. To our knowledge, the Pathogen Avoidance hypothesis has not been tested directly. Nevertheless, MacDorman and Entezari (2015) showed that individual differences in disgust sensitivity predict differential sensitivity to the uncanny valley, an important finding that offers indirect support for this hypothesis.

One study (Tinwell, Nabi, & Charlton, 2013) alluded to this hypothesis by showing that adult participants associated uncanniness with psychopathic personality traits, which include callousness, guiltlessness, and dishonesty. In the experiment, participants viewed virtual characters either exhibiting startle responses to a scream or failing to do so and rated them as either low or high in psychopathic personality traits, respectively. Researchers found that characters associated with pronounced psychopathic personality traits elicited the uncanny feeling, whereas others did not. Nevertheless, the study did not test whether the uncanny characters also elicited disgust. Furthermore, startle responses to a scream are not a common feature in the existing human replicas. Therefore, failing to do so is unlikely to explain the uncanniness in the human replicas that are incapable of facial expressions or body movements. The findings at best indirectly corroborate the Pathogen Avoidance hypothesis.

Mortality Salience hypothesis. Other researchers have argued that human replicas may remind humans of death, thereby eliciting the uncanny feeling driven by a fear of death. Indeed, many or most of humanoid replicas, such as dolls, clowns, mannequins, wax figures, zombies, and humanoid robots resemble dead individuals who have seemingly come alive. They can be threatening to some people because they elicit the fear of being replaced by an android Doppelganger, losing bodily control, or being deprived of a soul (Freud, 1919/1964; Ho et al., 2008). Inspired by Terror Management Theory in social psychology (Pyszczynski, Greenberg, & Solomon, 1999), Sara Kiesler proposed the Mortality Salience hypothesis (Ho et al., 2008; MacDorman & Ishiguro, 2006), which posits that some human replicas are uncanny because they remind people of death and trigger defense systems that cope with the deeply rooted anxiety for mortality.

According to Terror Management Theory, conscious and unconscious thoughts of death elicit two distinct defense systems—proximal and distal terror management defenses (Pyszczynski et al., 1999). Distal terror management defenses involve maintaining one's self-esteem and cultural worldviews, namely, a symbolic, culturally constructed conception of reality that imbues life with a sense of "order, permanence, and stability" (Pyszczynski et al., 1999, p. 836). To test whether some androids are uncanny because they unconsciously remind people of death, MacDorman and Ishiguro (2006) showed the images of an uncanny android and a person to experimental ($n = 31$) and control groups ($n = 32$), respectively. Participants then completed a questionnaire assessing their worldview protection as a measure of distal terror management defenses. Participants read campaign speeches from a (a) charismatic candidate and (b) relationship-oriented candidate and then rated on a 9-point scale on how much they liked each candidate, how insightful each candidate was, and which one they would support. It was hypothesized that if the uncanny android reminded participants of death, they would demonstrate enhanced worldview protection and prefer the charismatic candidate, who could presumably protect them against threat to their cultural worldviews. Following the same procedure, participants then rated two foreign students, of whom one praised the participants' home country whereas the other criticized it. Again, if the uncanny android served as a reminder of death, it would lead the participants to favor the student who favored their home country. The results supported these predictions, suggesting that the uncanny android provoked distal terror management defenses.

Nevertheless, these findings do not adequately corroborate the Mortality Salience hypothesis, which posits that fear is elicited. Research has shown that affects do not mediate distal terror management defenses (Pyszczynski et al., 1999). Therefore, according to Terror Management Theory, although the uncanny android still served as a reminder of death and triggered distal terror management defenses, it did not elicit fear or any other affects. To demonstrate that the uncanny feeling was a fear of death, future researchers should measure participants' affective responses to the android or adopt a different theory, using implicit measures, such as implicit association tests, to assess fear. One advantage of such implicit measures is that they would minimize demand characteristics arising from cues regarding the nature of the study. Further research on the Mortality Salience hypothesis is clearly warranted.

Evolutionary Aesthetics hypothesis. Hanson (2005) proposed the Evolutionary Aesthetics hypothesis, which suggests that selection pressures have shaped human preference for certain physical attributes signaling fitness, fertility, and health, and that the uncanniness of a human replica is due to its low attractiveness rather than its lack of realism. To test this hypothesis, Hanson (2005) first detected an uncanny valley using morphed images. He then enhanced the aesthetic properties of the uncanny face and showed that the valley disappeared. This finding seems to suggest that human replicas' limitation in aesthetic value is responsible for its uncanniness. Nevertheless, the removal of the uncanny valley after tuning the aesthetic properties of the uncanny face does not demonstrate that aesthetic properties caused the uncanniness in the first place. Furthermore, by simultaneously manipulating realism and attractiveness, Hanson introduced a potential confound and was unable to dissociate the effects of the two variables in determining human replicas' uncanniness.

In sum, the perceptual hypotheses, including the Pathogen Avoidance, the Mortality Salience, and the Evolutionary Aesthetics, have yet to be adequately tested for their ability to account for the uncanny phenomenon in human replicas. Their limitations are in part attributable to the methodologies used to test these hypotheses, which have not successfully drawn causal relations between the uncanny phenomenon and the proposed factors, including transmissable diseases, reminders of death, and low attractiveness. In addition, these hypotheses assume that human observers perceive the human replicas as real people, which would be indispensable for the aforementioned factors to trigger the pathogen avoidance mechanism, to remind people of their inevitable death, or to activate the evolved mechanisms for facial attractiveness. Nevertheless, this assumption has not been tested or acknowledged in any of these hypotheses. Finally, researchers have created human replicas that do not exist, such as morphed images or computer-generated characters, to test these perceptual hypotheses. By doing so, they have deviated from the existing human replicas that elicit the uncanny feeling and therefore forfeit a certain degree of ecological validity. Because not all uncanny human replicas possess features that remind people of transmissable diseases, death, or are less attractive, these hypotheses will not be able to explain the uncanniness in the human replicas that do not fit into those categories.

The Uncanny Phenomenon as Cognitive Processing

In addition to perceptual theories, researchers have proposed cognitive theories to explain the uncanny phenomenon, including the *Violation of Expectation* hypothesis (Mori, 1970/2005), the *Categorical Uncertainty* hypothesis (Jentsch, 1906/1997), and the *Mind Perception* hypothesis (Gray & Wegner, 2012). We review each hypothesis in turn.

Violation of Expectation hypothesis. When Mori (1970/2005) coined the term "uncanny valley," he used the example of a prosthetic hand, which appeared real at a first glance but elicited eerie sensations as people realized that it was artificial arising from its cold temperature and the lack of soft tissue. Mori suggested that this eerie feeling was a result of the mismatch between visual and tactical information derived from the prosthetic hands. Similarly, Brenton, Gillies, Ballin, and Chatting (2005) suggested that a perceptual mismatch between graphical and behavioral cues could become disturbing as behavioral cues fail to match graphical realism, rendering an avatar Zombielike. Furthermore, Moore (2012) proposed that perceptual tension derived from conflicting perceptual cues (i.e., perceptual mismatches) may elicit the uncanny feeling. Using a Bayesian model, he applied this hypothesis to reproducing the uncanny valley statistically.

In line with this view, the Violation of Expectation hypothesis posits that human replicas elicit the uncanny feeling by creating expectations for a human but failing to match them (Mitchell et al., 2011; Saygin, Chaminade, Ishiguro, Driver, & Frith, 2012). A large body of research has tested the Violation of Expectation hypothesis and found that a variety of cross-modal mismatches, such as appearance—motion mismatch (Saygin et al., 2012) and face—voice mismatch (Mitchell et al., 2011), elicit the uncanny feeling. Using the repetition suppression technique, Saygin and colleagues (2012) found distinctive responses in the action perception systems (APS) in human brains when the participants viewed androids in comparison with humans or robots. This finding suggests that the uncanny phenomenon could be attributable to violating the expectation that human appearance predicts biological movements. The characters with incongruent features, such as those with human face—synthetic voice or robot face—human voice mismatches, elicited significantly eerier sensations than those with congruent face and voice (Mitchell et al., 2011). Mismatch induced by disproportionate facial features had similar effects (MacDorman et al., 2009; Seyama & Nagayama, 2007). Seyama and Nagayama (2007) showed that faces with mismatched facial features (e.g., eyes and head) elicited the strongest unpleasant feeling compared with those without a mismatch.

Although converging evidence seems to support a link between the violation of expectation and the uncanny feeling, the hypothesis is at best partially explanatory of the uncanny phenomenon. First and foremost, perceptual mismatch in androids does not necessitate the violation of expectations,[1] and previous studies have neglected to assess participants' expectations for the androids and the violations of those expectations. Saygin and colleagues (2012) noted that by the time scanning took place, participants had experienced a brief exposure to the android and were certain about its identity. It is therefore unlikely that participants would experience a violation of expectation due to unanticipated perceptual mismatches. Because fMRI data do not permit a fine-grained temporal analysis, it remains unclear whether the distinctive neural responses to androids reflect a violation of expectation. In addition, the Violation of Expectation hypothesis neglects to consider the effects of positive violations of expectations, for instance, in humor comprehension (Veale, 2004). Surprising violations that are positive in valence rarely, if ever, elicit an uncanny feeling. For the hypothesis to account for the uncanny phenomenon, researchers should explore the expectations people have for the androids and the cognitive underpinnings of the violations of these expectations, which were proposed to elicit the uncanny feeling. As MacDorman and Ishiguro (2006) wrote,

> While many nonbiological phenomena can violate our expectations, the eerie sensation associated with the uncanny valley may be peculiar to the violation of human-directed expectations, which are largely subconscious. If androids are more likely to fall into the uncanny valley than mechanical-looking robots, the reason may be that our brains are processing androids as human. (p. 301)

Categorical Uncertainty hypothesis. The concept of categorical uncertainty was first proposed by Ernst Jentsch (1906/1997), who argued that uncanniness is associated with "a lack of orientation," which arises when individuals incorporate new information with "mistrust, unease and even hostility (misoneism)" (p. 8). Similarly, Ramey (2006) suggested that the uncanny phenomenon concerns the process whereby uncertainty emerges at any category boundary, and is not specific to androids.

In three experiments, Yamada, Kawabe, and Ihaya (2013) tested whether categorization uncertainty might be responsible for the uncanny phenomenon. In a first experiment, they morphed each pair of a real, a stuffed, and a cartoon human face, creating three continua of 11 equally stepped morphed images. Participants performed a categorization task and an evaluation task in counterbalanced order. In the categorization task, participants categorized morphed images as either category (e.g., real human or stuffed human) as quickly as possible and their response latencies were

[1] For example, after repeated interactions with a human replica, observers are habituated to its perceptual mismatches and expect it to possess inconsistent levels of human likeness.

recorded. In the evaluation task, participants rated the likability of the 33 morphed images presented in a random order on a 7-point scale ranging from −3 (*strongly dislike*) to 3 (*strongly like*). Yamada and colleagues found that the longest response latency, the highest category ambiguity, and the lowest ratings on likability coincided at a consistent location on the morphing spectrum. They also found a strong negative correlation between response latencies and likability scores. These findings corroborated the Categorical Uncertainty hypothesis. In a second study, Yamada et al. (2013) replicated these findings with morphed images of dog faces. By showing that it is not limited to androids, this replication supports Ramey's (2006) idea that the uncanny phenomenon is linked to general categorization uncertainty arising at any category boundary. In a third experiment, Yamada et al. (2013) did not find the same effect using morphed images of real human faces (e.g., a female and a male face or two male faces).

Recently however, Kätsyri, Förger, Mäkäräinen, and Takala (2015) questioned the validity of these results. They argued that image morphing artifacts (e.g., "ghosting" and "color interpolation," for a discussion, see Kätsyri et al., 2015, p. 8) might have lower likability scores, acting as confounds to categorization uncertainty. Inconsistent with the results of Yamada et al. (2013), recent studies have refuted the Categorical Uncertainty hypothesis. Using carefully controlled continua of morphed human and avatar faces, Cheetham, Wu, Pauli, and Jäncke (2015) measured late positive potential (LPP), facial electromyography (EMG), and a self-assessment manikin (SAM) in participants viewing these faces. Running counter to the Categorical Uncertainty hypothesis, they showed that participants' negative affective responses to categorically ambiguous images were not stronger than to unambiguous images, consistent with an earlier finding (Cheetham, Suter, & Jäncke, 2014). Finally, a recent study (Mathur & Reichling, 2016) replicated the findings of Yamada et al. (2013) by showing that participants spent longer time rating human likeness of categorically ambiguous faces and rated these faces as least likable compared with categorically unambiguous ones. Nevertheless, the authors also showed that rating time, which indicates categorization uncertainty, did not statistically mediate the relation between human likeness and likability scores.

In sum, in an attempt to account for the uncanny phenomenon, the Violation of Expectation and the Categorical Uncertainty hypotheses focus on cognitive processes in their analysis. Nevertheless, at least in their current forms, they neglect to explain what the expectations for humans and those for robots are, and why violating such expectations could elicit the uncanny feeling. To address these questions, we next review the Mind Perception hypothesis (Gray & Wegner, 2012).

Mind Perception hypothesis. Although androids and computer-generated characters have increasingly come to resemble humans, the uncanny feeling they elicit reminds people of their identity as robots as opposed to humans. The distinctions between robots and humans lie not only in their physical appearance but also in human observers' perception of them. Therefore, instead of trying to perfectly recreate the physical features of humans, researchers should focus on the question, "On what bases do people perceive each other as humans?" This question encompasses the broader perspective of social cognition, including such topics as Theory of Mind, face perception, and emotion recognition. Given the scope of this manuscript, we focus on people's perception of humanness, namely, "attributes that define what it is to be human" (Haslam & Loughnan, 2014, p. 401).

Humans and nonhuman characters differ on two major dimensions: agency (the ability to plan and do things) and experience (the ability to feel and sense things) (Gray, Gray, & Wegner, 2007). Gray and Wegner (2012) argued that androids especially lack the capacity to feel and sense, which helps define human beings. Based on this distinction, they proposed that human replicas are uncanny not because they are not realistic enough to be indistinguishable from real people. Instead, they are uncanny because they are so realistic that they trigger the attributions of human minds, particularly subjective experience, to nonhuman entities. These authors argued that the attribution of human experience to androids, which violates humans' deep-rooted expectations for a robot, is responsible for the uncanny phenomenon. To test this hypothesis, they conducted three experiments.

In the first, they showed participants either a humanlike or a mechanical-looking robot, and asked them to rate the two robots on uncanniness and capacity to feel and sense (e.g., "This robot has the capacity to feel pain") and to plan and act (e.g., "This robot has the capacity to plan actions") on a 5-point scale. The humanlike robot received higher ratings on uncanniness than did the mechanical-looking robot; importantly, the perceived uncanniness was predicted by participants' perceived experience instead of their agency. In a second experiment, the authors investigated whether a humanlike appearance is necessary for perceiving experience in a robot to elicit uncanniness. To do so, each participant received a questionnaire, which described one of three computers (one that could feel hunger and emotions, one that could execute self-control, and one that was simply powerful in function) and rated their uncanniness. The results suggested that perceived experience elicited the uncanny feeling even in the absence of a humanlike appearance. Lastly, the researchers hypothesized that perceived lack of experience, and the resulting unfeeling humans, would also elicit uncanniness. To test this hypothesis, in the last experiment, they provided descriptions of three humans: One lacked the mental capacities to experience, one lacked the mental capacities to plan, and one (a control condition) appeared normal. Participants rated their uncanny feeling and perceived agency and experience of the three humans. The results showed that lack of experience rather than of agency in humans elicited the uncanny feeling. In summary, these findings suggest that the uncanny feeling is linked to violating the expectation that robots lack subjective experience, which characterizes humans.

Nevertheless, we argue that the attribution of experience to robots by itself cannot account for the uncanny feeling. The process of mind perception (Gray & Wegner, 2012) belongs to a broader phenomenon known as anthropomorphism, which according to the *Oxford Dictionary* (Soanes & Stevenson, 2005), refers to "the attribution of human characteristics or traits to nonhuman agents" (p. 865, as cited in Epley, Waytz, & Cacioppo, 2007). Contrary to what the Mind Perception hypothesis would predict, anthropomorphism, although violating some expectations, generally facilitates rather than hinders social interactions between humans and nonhumans. Anthropomorphism is pervasive in human societies and is reflected in humans' religious beliefs. Religious believers typically view Gods as possessing human forms, and as experiencing emotions as ordinary people do, although the anthropomorphism of Gods is implicit (Barrett & Keil, 1996). The propensity of pet owners to take care of and become attached to beloved animals (Archer, 1997) and of computer users to demonstrate interpersonal responses to computers (Nass, Steuer, & Tauber, 1994) further show humans' predispositions to anthropomorphize nonhuman and nonliving things alike. Although the process of anthropomorphism does not preclude from any negative effects caused by the anthropomorphized agents, such as a malfunctioning computer or a disobedient pet, the process of anthropomorphism itself does not seem to cause those negative effects directly.

We agree with Gray and Wegner (2012) that mind perception provides new insights into the uncanny phenomenon by focusing on not merely humans' hypersensitivity to the distortion of facial features, but also humans' perception of other entities' ability to sense, plan, and act (i.e., humanness). Nevertheless, given that the process of attributing human experience to nonhuman and nonliving things does not seem to elicit negative effects in various domains, we question the Mind Perception hypothesis, which posits that the uncanny phenomenon is a result of the attribution of humans' subjective experience to robots.

To summarize the cognitive theories, the Violation of Expectation, Categorical Uncertainty, and Mind Perception hypotheses overlap in their accounts of the uncanny phenomenon. Both the Violation of Expectation and the Categorical Uncertainty hypotheses assume that perceptual mismatches intrinsic to human replicas elicit the uncanny feeling. The Mind Perception hypothesis is based on the Violation of Expectation hypothesis; particularly, it emphasizes the role of attributing humans' subjective experience to robots in eliciting the uncanny feeling. Both the Categorical Uncertainty and the Mind Perception hypotheses highlight human–nonhuman distinctions. Their interdependency entices the search for a more explanatory and parsimonious account of the uncanny phenomenon that would not rely on other hypotheses or any untested assumption.

We will return to this topic by proposing our account of the uncanny phenomenon, namely, the *Dehumanization* hypothesis. Before delineating this hypothesis, we provide a critique of studies on the uncanny valley hypothesis with respect to their methodological limitations. In trying to explain the inconsistent findings in the literature, we examine the definitions of human likeness and likability, and propose a statistical test to quantify their nonlinear relation.

Reconstructing the Uncanny Valley Hypothesis

Since Mori (1970/2005) proposed the uncanny valley hypothesis, researchers have used different methodologies to validate it. As noted earlier, the findings have been inconsistent. The reason for the discrepancy in the literature is threefold. First, the dependent variable—likability—has lacked a consistent definition and a valid measurement. Second, the independent variable—human likeness—is a multifaceted construct, which is difficult to define and manipulate systematically. Third, the uncanny valley hypothesis lacks an explicit mathematical model that specifies the shape of the hypothetical curve. Without an a priori specification of the shape of the hypothetical curve, the interpretation of findings is vulnerable to researchers' biases, especially confirmation bias and the "Texas sharpshooter fallacy," the error of drawing arbitrary boundaries around random data after peeking at it (Schneps & Colmez, 2013).[2]

Definition and Measurement of the Dependent Variable

The most commonly used measurements of the dependent variable in the uncanny valley hypothesis have been semantic differential and Likert scales (Hanson, 2005; MacDorman, 2006; MacDorman & Ishiguro, 2006; Seyama & Nagayama, 2007). In the process, researchers have translated from Japanese and interpreted the dependent variable "shinwakan" in various ways (Bartneck et al., 2007, p. 369). These interpretations vary from "familiarity" (MacDorman, 2006, p. 26; MacDorman & Ishiguro, 2006, p. 299; Mori, 1970/2005, p. 33; Tinwell et al., 2011, p. 741) to "impression of pleasantness" (Seyama & Nagayama, 2007, p. 337), "acceptability" (Green et al., 2008, p. 2459; Hanson, Olney, Pereira, & Zielke, 2005, p. 30; Piwek, McKay, & Pollick, 2014, p. 271; White, McKay, & Pollick, 2007, p. 477), eeriness (Mitchell et al., 2011, p. 10), comfort level (MacDorman et al., 2009, p. 695), "valence" (Cheetham, Suter, & Jäncke, 2011, p. 2), "empathetic emotional response" (Misselhorn, 2009, p. 346), and "likability" or "affinity" (Bartneck et al., 2007, p. 369; Bartneck, Kanda, Ishiguro, & Hagita, 2009, p. 270). Researchers have often used these terms interchangeably in previous studies, raising the question of whether they are equivalent in meaning, and if they differ, which term best captures the dependent variable in the uncanny valley hypothesis. The research on this hypothesis has undergone a 45-year trajectory, during which the interpretation of "shinwakan" has changed substantially. To illustrate this development, we review the history of the first English translation of Mori's Bukimi No Tani.

Mori (1970) published the essay "Bukimi No Tani" in an obscure Japanese journal *Energy*, in which he first proposed the uncanny valley hypothesis. In 2005, Karl MacDorman and his Japanese colleague Takashi Minato translated the essay into English, in which the term "familiarity" appeared as the English equivalent for "shinwakan."

Nevertheless, as research on the uncanny valley hypothesis developed, MacDorman and other researchers began to question this translation and suggested that a more accurate translation would be terms like "affinity," "likability," or "interpersonal warmth," to capture the intended meaning of the Japanese word shinwakan. Researchers have realized that "shinwakan" does not possess a direct equivalent in English, and must be understood in terms of its two components—"shinwa" and "kan"—which mean "mutually be friendly" (or "having similar mind") and "the sense of," respectively. Bartneck and colleagues (2007) proposed using "affinity" and "likability" to replace "familiarity" as a more suitable translation of "shinwakan." Realizing the confusion caused in part by the initial translation (Mori, 1970/2005), Mori, MacDor-

[2] This fallacy is named after the proverbial Texas sharpshooter who fires a round of bullets into a wall and then draws a circular target around them after the fact, claiming that "I hit the bullseye."

man and Kageki (2012) revised the translation of “Bukimi No Tani,” in which shinwan was translated into “affinity.”[3]

In addition to the problem of what might have been lost in translation, concerns were raised regarding the reliability and validity of the measurement of the dependent variable (Bartneck et al., 2007). For example, a single question is unlikely to measure a certain affective response properly. Thus, multiindicator measurements, such as the liking questionnaire (Monahan, 1998), which allows the calculation of reliability, were advocated. Bartneck and colleagues developed a series of scales to systematically assess variables in human–robot interactions, including anthropomorphism, animacy, likability, perceived intelligence, and perceived safety (Bartneck, Kulić, Croft, & Zoghbi, 2009). These scales allow researchers to systematically measure relevant concepts with higher reliability and validity.

A more fundamental issue associated with the definition of the dependent variable in the uncanny valley hypothesis is its conceptualization. That is, what do likability and uncanniness mean and how should we measure them? Ho and colleagues (2008) examined the emotional terms that potentially capture the uncanny phenomenon. By analyzing participants’ ratings on animate robots, which varied in appearance and behavior from mechanical-looking to humanlike, they found that the term “fear” was highly predictive of attributions of eeriness or creepiness to the robots. Disgust, shock, and nervousness were also significant predictors of robots’ eeriness or creepiness. Instead of reducing eeriness or creepiness to a single emotional term, the authors suggested that “the uncanny valley may not be a single phenomenon to be explained by a single theory but rather a nexus of phenomena with disparate causes” (Ho et al., 2008, p. 175).

Across studies testing the uncanny valley hypothesis, the dependent variable is typically assumed to vary along a single dimension. For example, Mori (1970/2005) wrote,

> 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 humanlike, but the familiarity is negative. This is the uncanny valley (p. 34).

Nevertheless, as pointed by MacDorman and Ishiguro (2006), no compelling evidence exists in support of this assumption. Bartneck and colleagues (2007) found that humanoids and toy robots received higher ratings on likability than did real humans, which contradicted the uncanny valley hypothesis. Chances are high that participants hold different criteria for what constitutes likability for toy robots and for androids. Toy robots may acquire likability simply because they possess humanlike features, whereas humans and androids may require friendly and agreeable characteristics to reach a comparable level of likability, which is more challenging. Further research should clarify what the items in the questionnaire mean to participants.

In summary, the interpretations and conceptualization of the dependent variable in the uncanny valley hypothesis have been unclear, which may have contributed to inconsistent findings. Most researchers now agree with the interpretation of shinwakan as “affinity” or “likability.” From this consensus, researchers now need to examine, in the context of the uncanny valley, what likability of human replicas means for human observers. As MacDorman (n.d., quoted in Hsu, 2012) elaborated on this concept,

> I think it is that feeling of being in the presence of another human being—the moment when you feel in synchrony with someone other than yourself and experience a ‘meeting of minds.’ Negative ‘shinwakan,’ the uncanny, is when that sense of synchrony falls apart, the moment you discover that the one you thought was your soul mate was nothing more than smoke and mirrors. (Hsu, 2012)

Definition and Measurement of Human Likeness

In addition to the dependent variable of likability, researchers have called for a rigorous definition of human likeness because of its complex nature (Ramey, 2006).

Measurement of human likeness. Similar to the measurement of likability, the conventional measurement of human likeness of a human replica includes subjective ratings. In addition, implicit measures of human likeness capitalize on participants’ largely unconscious behavior such as gaze during human–android interactions (MacDorman et al., 2005; Shimada, Minato, Itakura, & Ishiguro, 2007). Research has shown that people often break eye contact from the interlocutor while thinking. This averted gaze, according to the Social Signal Theory, serves in part as a signal to others that they are thinking (McCarthy & Muir, 2003), and is expected to be influenced by the interlocutor during interactions (MacDorman et al., 2005). Shimada and colleagues (2007) quantified the similarity between the gaze behavior (i.e., duration and direction of gaze) toward an android and a human, and showed that an android with greater physical similarity to a human elicited more human-directed gaze. These findings suggest that gaze behavior may be used to measure the human likeness of human replicas at an implicit level.

Human likeness: Subjectivity versus objectivity. Another fundamental yet understudied question is whether human likeness is a subjective or an objective construct. In most studies, human likeness has been measured by means of subjective ratings. Human likeness was therefore defined subjectively. In contrast, in other studies, human likeness was manipulated and measured largely objectively either by changing the surface characteristics of a face (e.g., degree of photorealism, proportion of facial features, and texture, MacDorman et al., 2009), or by morphing a real into an artificial face (Hanson, 2005; Seyama & Nagayama, 2007).

Subjective measurement of human likeness is limited in three respects. First, individual differences in the sensitivity to uncanniness (MacDorman & Entezari, 2015) and the evaluation of human likeness are likely to introduce psychometric noise into self-reported ratings, particularly when participants see multiple categories of stimuli (e.g., toy robots, dolls, androids, and humans). Second, subjective ratings of human likeness and likability can interfere with one another so that ratings on both variables are inherently correlated (Ho & MacDorman, 2010). Third, because the subjective ratings of human likeness reflect merely the relative realism in a given set of stimuli from one study, the extent to which the sample of human replicas are representative of the population is critical. In conclusion, it becomes crucial to develop standard criteria to measure the human likeness of human replicas objectively.

[3] Researchers thereafter have typically used affinity and likability interchangeably as the dependent variable in the uncanny valley hypothesis.

In the studies that used morphing techniques or computer-generated images, researchers quantified human likeness in terms of the proportion of a human face in a morphed image (Hanson, 2005; Seyama & Nagayama, 2007), or the degree of photorealism and polygon counts of facial features (MacDorman et al., 2009). Nevertheless, these objective measurements, by introducing disproportionate facial features, effects of morphing or modifying attractiveness of the replica faces, confound human likeness with various manipulation-related artifacts. For example, Burleigh, Schoenherr, and Lacroix (2013) manipulated the level of human likeness by changing the facial proportions, thereby confounding human likeness with facial abnormality. Given that human likeness is a multifaceted construct, an objective measure of human likeness may oversimplify this concept.

Given the advantages and disadvantages of pursuing both subjective and objective measurements of human likeness, this issue remains unresolved. We suggest that to test the uncanny valley hypothesis, researchers should (a) objectively measure human likeness and (b) obtain a sizable and diverse collection of images or videos to represent a spectrum of human replicas (Mathur & Reichling, 2016).

Quantification of the Hypothetical Curve

The inconsistent findings are also attributable to the differing criteria researchers have used to ascertain whether they detected Mori's hypothetical curve. These criteria may afford researchers too many degrees of freedom for corroborating the existence of the uncanny valley, allowing confirmation bias and the Texas Sharpshooter Fallacy to creep into their judgments. MacDorman and Ishiguro (2006) claimed to find the uncanny valley by identifying a "valley" in the plotted curve. Nevertheless, the observed "valley" occurred in the middle range of human likeness, which failed to precisely match Mori's depiction of the hypothetical curve. Bartneck et al. (2007) also found a "valley," except that the second peak was lower than the first peak. The authors disputed the uncanny valley hypothesis and argued instead for an "uncanny cliff." In both studies, the authors might have interpreted the data in different ways, given that their data only partially resembled the hypothetical curve.

The divergent criteria for defining the uncanny valley in the literature stems from the longstanding absence of strict quantification criteria for the hypothetical curve. This void in the literature led not only to ambiguous interpretations of findings, but also led researchers searching for the uncanny valley to evoke extraneous variables, such as attractiveness (Hanson, 2005) and facial abnormality (Burleigh et al., 2013; Seyama & Nagayama, 2007). This approach is problematic for two reasons. First, by manipulating variables other than human likeness to map data onto an uncanny valley, researchers have deviated from the original definition of the uncanny valley, which was exclusively about the nonlinear relation between human likeness and likability. Second, across studies, researchers have created multiple uncanny valleys, which not only differ in shapes, but also lack a consistent explanation. In this case, the uncanny valley becomes too vague and open-ended to allow for rigorous investigation. Here, we propose to test the uncanny valley hypothesis against its null hypothesis—as human likeness increases, likability increases linearly.

The statistical test includes two components: a linear regression and a residual analysis. As the uncanny valley hypothesis suggests, we can regress likability against human likeness. Researchers can then check linearity to determine whether the null hypothesis—there is a linear relationship between human likeness and perceived likability—is rejected or not. Rejecting the null hypothesis would serve as a first step in corroborating the uncanny valley hypothesis. Using a residual analysis, researchers could further validate it by mapping data onto Mori's hypothetical curve. If the uncanny valley exists, as human likeness increases, the residuals of likability should scatter around zero and indicate a random pattern, until a point at which the residuals plunge abruptly into negative values; as human likeness continues to increase, the residuals become close to zero and random again. Although this method does not allow complete quantification of the uncanny valley hypothesis (in part because it does not articulate the location of the plunge, a prediction not made explicit by Mori, 1970/2005), it would provide a robust statistical test of whether a nonlinear relation between human likeness and likability exists, thereby moving a step further to validating the uncanny valley hypothesis.

Definition of the Uncanny Feeling

Until recently, the definition of the uncanny feeling has been unclear. Studies have interpreted the uncanny feeling in various ways, including the highest rating on eeriness, or the lowest rating on familiarity or acceptability (Hanson, 2005; MacDorman et al., 2009; Seyama & Nagayama, 2007). The different ways of operationalizing the uncanny feeling are partly responsible for generating inconsistent results (MacDorman & Ishiguro, 2006). These mixed results necessitate a closer look at the definition of the term "uncanny."

Historically, Freud (1919/1964) studied the uncanniness by combining the examination of its semantics in dictionaries and analyzing the situations in which it emerged. Freud regarded uncanniness as related to fear, but viewed it as unique in that it is the fear associated with novelty and unfamiliarity. He believed that what makes novelty and unfamiliarity uncanny is what Schelling defined as something "that ought to have remained . . . secret and hidden but has come to light" (Freud, 1919/1964, p. 219). Freud further interpreted the meaning of uncanny by referring to a situation that elicited an uncanny sensation, namely Hoffmann's (1817/1994) story "The Sandman." From a psychoanalytic point of view, Freud stated that the uncanny feeling stemmed from children's castration anxieties, as it was described in the story that the sandman tears out children's eyes, which ostensibly symbolize testicles. Nevertheless, given that what Freud described as the uncanny feeling in the "Sandman" does not map onto Mori's conceptualization of uncanniness during human–robot interactions, his interpretation is of questionable relevance to the uncanny valley hypothesis and the uncanny phenomenon discussed here.

In contrast, Ernst Jentsch, a German psychiatrist, contributed the first psychological literature on the psychology of the uncanny. In the essay "On the Psychology of the Uncanny," Jentsch (1906/1997) interpreted uncanniness as related to uncertainty. He also referred to Hoffmann's "The Sandman," but focused on a slightly different aspect than did Freud. Jentsch proposed that what elicited the uncanny effect was uncertainty about whether the character in the story—the automaton Olympia—was a real person. This interpretation of the "uncanny" sensation was situated in human replicas, as Mori did in his uncanny valley hypothesis. From the discussion of the Categorical Uncertainty hypothesis, we see that Jentsch's idea about the cognitive underpinnings of the uncanny phenomenon is still influential in contemporary research.

In 1970, Mori defined the uncanny feeling as a lack of affinity in humans' reaction to realistic prosthetic hands. Nevertheless, there has been little research on the emotional components of the uncanny feeling. Ho and colleagues (2008) surveyed 143 Indonesian adult participants who had little experience with robots prior to the study. Researchers presented video clips of 17 robots and a person in a random order while asking participants to rate 31 statements about their emotional reactions (e.g., fear, disgust, shock, and happiness) to these characters on 7-point Likert Scales. By regressing ratings on eeriness and strangeness against those on the emotional reactions, researchers found that five emotions—fear, disgust, nervousness, dislike, and shock—combined to explain 60% of the variance in eeriness. In contrast, the five top predictors of strangeness accounted for only 37% of its variance. Based on the higher coefficient of determination in predicting eeriness than strangeness, Ho and colleagues argued that eeriness—a mix of fear, disgust, nervousness, dislike, and shock—captures our emotional reactions to uncanny robots better than strangeness. Although some robots can elicit the uncanny feeling, it does not follow that all robots are uncanny or that our emotional reactions to robots are all about uncanniness. In this study, the researchers assumed without adequate justification that the 17 robots are uncanny and that the term (i.e., eeriness) most accounted by basic emotional reactions would best characterize uncanniness.

The meaning of "uncanniness" also lies in the translation of "Bukimi No Tani (不気味の谷現象)." Its accurate translation, according to MacDorman (n.d., quoted in Hsu, 2012), is "valley of eeriness" rather than the far better known translation—"uncanny valley," which first appeared in Jasia Reichardt's 1978 book *Robots: Fact, Fiction, and Prediction* (Hsu, 2012). The reason for using "uncanny valley" as the translated title was that translators MacDorman and Minato thought the term was familiar to English speakers (Hsu, 2012). We note that the Japanese term "Bukimi (不気味)" was first translated into "strangeness" (Mori, 1970/2005, p. 34), which was later replaced by the terms such as "eeriness" and "creepiness" (Mori et al., 2012, p. 100). As the translation "uncanny valley" has gained worldwide acceptance, the English word "uncanny" has become the standard interpretation of the Japanese word "Bukimi." Because "uncanny" is also close in meaning, although not equivalent, to the German term "Unheimliche," it is therefore also linked to its aforementioned interpretations in the German literature, including Jentsch (1906/1997)'s *On the Psychology of the Uncanny* and Freud (1919/1964)'s *The Uncanny.*

Summary

In recognition of the inconsistent findings in the literature, we have reviewed the measurements and definitions of likability and human likeness, and proposed an innovative approach to quantitatively testing the uncanny valley hypothesis. Researchers have measured human likeness using both subjective ratings and objective behavioral measures, such as gaze behavior. They have proposed different interpretations of the dependent variable originally known as "shinwakan" in Japanese and have used subjective ratings to measure it. Nevertheless, there exist no standard measurements for both the independent variable and the dependent variable in the uncanny valley hypothesis. This methodological limitation is partly attributable to the ill-defined construct of human likeness and the multifaceted nature of the uncanny feeling. Future research should provide clearer definitions and reliable measurements of these constructs to test the uncanny valley hypothesis. To do so, we suggest studying the cognitive underpinnings of the uncanny phenomenon. As previously discussed, the uncanny phenomenon may be linked to perception of humanness. Therefore, understanding how people perceive each other may help to define human likeness above and beyond the physical similarities between humans and nonhuman entities. It may also allow researchers to better understand the uncanny feeling by studying its cognitive underpinnings.

Previous studies have proposed hypotheses derived from evolutionary, social, and cognitive approaches. Although these hypotheses provide plausible accounts of the uncanny phenomenon from different perspectives, they have neglected to verify the underlying assumption that observers would spontaneously perceive a human replica that closely resembles humans as a person. Because of previously mentioned limitations in their methodologies, there is a lack of compelling evidence to support these hypotheses. Here, we hypothesize that two cognitive processes—anthropomorphism and dehumanization—may underlie the uncanny phenomenon. In particular, we argue that dehumanization of an anthropomorphized human replica elicits the uncanny phenomenon, a notion we term the *Dehumanization* hypothesis.

The Dehumanization Hypothesis

Dehumanization refers to "perceiving a person or group as lacking humanness—the attributes that define what it is to be human" (Haslam & Loughnan, 2014, p. 401). Haslam and Loughnan (2014) noted that over four decades the study of dehumanization has extended to numerous phenomena, including the extreme cases of genocide, slavery, and the subtler cases of social stereotyping toward immigrants and ethnic minority groups. Systematic investigation into dehumanization began in the 1970s and the 1990s in social psychology, with special attention to violence (Bandura, Underwood, & Fromson, 1975; Kelman, 1973; Staub, 1989), moral judgment (Bandura, 1999; Opotow, 1990), and intergroup conflict (Struch & Schwartz, 1989). Nevertheless, recent theories have incorporated subtler forms of dehumanization, one of which is known as "infrahumanization" (Bain, Park, Kwok, & Haslam, 2009; Leyens et al., 2001). Stemming from the anthropological concept of ethnocentrism, infrahumanization implies that people tend to perceive out-group members as possessing fewer humanlike properties but more animal-like features than in-group members (Viki et al., 2006). Infrahumanization occurs even without intergroup conflict (Leyens, Demoulin, Vaes, Gaunt, & Paladino, 2007).

Haslam (2006) proposed a dual model to integrate the extreme and subtle forms of dehumanization and proposed two forms of dehumanization: The denial of human uniqueness renders a person animalistic such that he or she is considered lacking in self-control and high levels of intelligence. In contrast, the denial of human nature renders a person mechanistic such that he or she is considered lacking in emotions and warmth. The distinction between the two forms of dehumanization—animalistic and mechanistic—according to Angelucci and colleagues (Angelucci, Bastioni, Graziani, & Rossi, 2014), may shed new insights onto the uncanny phenomenon in androids and other human replicas, although which form of dehumanization is responsible for the uncanny phenomenon awaits further examination. This view echoes the writings of Złotowski, Proudfoot, and Bartneck (2013), who argued that understanding how dehumanization influences perception of humanness can provide new insights on human-robot interactions and validation of the uncanny valley.

Based on the literature on anthropomorphism and dehumanization, we develop these aforementioned ideas and propose that attributing humanlike characteristics (e.g., humans' subjective experience), by itself, does not explain the uncanny feeling; instead, perceiving an anthropomorphized human replica as lacking humanness, namely, the process of dehumanization, is at work in the uncanny phenomenon.

An anthropomorphized human replica appears to be the assumption underlying all the previously mentioned hypotheses: the Pathogen Avoidance, the Mortality Salience, the Evolutionary Aesthetics, the Violation of Expectation, the Categorical Uncertainty, and even the Mind Perception. In previous studies on the uncanny valley hypothesis, however, researchers have neglected to corroborate this assumption by showing that human observers spontaneously anthropomorphize human replicas and perceive them as real people. Nevertheless, this assumption is plausible, given the propensity we have to anthropomorphize inanimate or nonhuman entities in literature, the arts, sciences, and in perception (Guthrie, 1993). According to anthropologist Stewart Guthrie (1993), both anthropomorphism (e.g., attributing human characteristics to the nonhuman) and animism (e.g., attributing life to the inanimate) stem from a perceptual strategy, probably evolved in part through natural selection, to "search for organization and significance, and both consist in overestimating them" (p. 62).

Guthrie's theory of anthropomorphism allows us to build on this assumption and argue that dehumanization may extend from humans to human replicas (e.g., androids and avatars). It follows that any feature of a human replica, including its eyes, emotional expressions, voice, and movements, that reveals its mechanistic nature could question the replica's humanness such as the capacity for emotions and warmth, and lead to dehumanization, thereby diminishing its likability and eliciting the uncanny feeling. Critically, the uncanny human replica is not perceived to be a typical robot; it becomes a dehumanized "robotlike" human who lacks humanness.

The Dehumanization hypothesis is not necessarily in conflict with previous hypotheses or findings. It is compatible with the findings that human observers' attribution of humanlike experiences to nonhuman robots is linked to the uncanny feeling (Gray & Wegner, 2012). According to the Dehumanization hypothesis, however, we interpret these findings from a different perspective. The more human observers attribute humanlike characteristics to (i.e., anthropomorphize) a human replica, the more likely detecting its mechanistic features triggers the dehumanization process that would lead to the uncanny feeling.

In trying to understand the cognitive processes that may underlie the uncanny phenomenon, we first review relevant findings in face perception literature, based on which we then propose an innovative paradigm to test the Dehumanization hypothesis.

A large body of literature has demonstrated humans' rapid detection of faces, whether they are humans, dolls, or schematic line drawings. Nevertheless, humans are also sensitive to cues signaling animacy in faces. Looser and Wheatley (2010) presented continua of morphed images created from human (animate) and mannequin (inanimate) faces. They found that participants used static cues in the morphed images to infer animacy. They perceived animacy consistently at a location close to the midpoint on the continuum but was shifted toward the human endpoint. Farid and Bravo (2012) showed that human observers reliably discriminated between computer-generated (inanimate) and photographic (human) images and that their performance was largely robust to variations in image resolution, compression, and color.

Researchers have argued that this hyperacuity to animacy may be associated with the activity of brain regions including fusiform gyrus (FG) and posterior superior temporal sulcus (pSTS). Facing two computer-generated characters with identical humanlike appearance, the FG responded more strongly to the perceived animate agent that demonstrated goal-directed actions than to the one that lacked goal-directedness (Shultz & McCarthy, 2014). Using multivariate pattern analysis of blood-oxygen-level dependent (BOLD) responses, Looser, Guntupalli, and Wheatley (2013) examined the neural correlates of animacy perception in static faces by delineating two stages of face processing substantiated by the "core face perception system" (Haxby, Hoffman, & Gobbini, 2000, p. 224). They demonstrated that the inferior occipital gyrus (IOG) prioritizes face form detection during an early stage of face processing, whereas the lateral fusiform gyrus (latFG) and right superior temporal sulcus prioritize animacy perception during a later stage. The authors argued that face form detection may endow people with an evolutionary advantage in perceiving faces in everyday objects (Guthrie, 1993), because false alarms resulting from the misidentification of a falling rock as a predator are associated with higher long-term fitness than misses resulting from the failure to correctly identify a predator. Nevertheless, the authors noted that repeatedly false alarms can be costly from the standpoint of time and energy expenditure, so animacy perception may serve to optimize allocation of social–cognitive resources by reducing false alarms.

These arguments were corroborated by the findings that event-related potential (ERP) components index animacy discrimination during face processing (Balas & Koldewyn, 2013; Wheatley, Weinberg, Looser, Moran, & Hajcak, 2011). Although both animate (human) and inanimate (doll) faces provoked early electrophysiological responses (N170/VPP), only human faces sustained a positive potential beyond 400 ms following the stimulus onset (Wheatley et al., 2011). Balas and Koldewyn (2013) found larger P100 amplitude in response to artificial faces than to human faces.

We argue that the two stages of face processing (Looser et al., 2013) may map closely onto the Dehumanization hypothesis by providing neuroimaging evidence that supports the mental processes we proposed to underpin the uncanny phenomenon. As Looser and colleagues (2013) argued, the IOG may facilitate the propensity to anthropomorphize a realistic human replica, giving observers an impression that it is a real person. The latFG and STS may then facilitate the scrutiny of the anthropomorphized replica by allowing humans to detect whether it is animate and deserves social–cognitive resources. Once animacy perception reveals that the replica lacks "life" or a mind, which is highly likely given observers' hypersensitivity to animacy (Farid & Bravo, 2012), the replica may be dehumanized, eliciting the uncanny feeling.

We have thus far reviewed literature on animacy perception and suggested a link between the two stages of face processing (Looser et al., 2013) and the proposed Dehumanization hypothesis. It is well established that the first stage prioritizes face form detection, whereas the second stage prioritizes animacy perception (Looser et al., 2013; Wheatley et al., 2011). Nevertheless, these findings do not preclude the possibility that animacy perception occurs during the first stage (e.g., within 400 ms following stimulus onset). In fact, Balas and Koldewyn (2013) found a larger peak amplitude at the P100 component in response to artificial than to human faces, suggesting that animacy perception may occur as early as 100 ms. As the authors noted, this finding seems inconsistent with findings that showed reliably larger P100 responses to faces than to objects (e.g., cars) (Dering, Martin, Moro, Pegna, & Thierry, 2011; Dering, Martin, & Thierry, 2009; Thierry, Martin, Downing, & Pegna, 2007). These findings suggest that during an early stage of face processing, artificial faces are perceived as more humanlike and animate than are human faces. Nevertheless, this inconsistency could be explained if we assume that the artificial faces were anthropomorphized during the early stage of face processing. Future research should examine the extent to which artificial faces that vary in human likeness may be anthropomorphized and the possibility that animacy perception occurs during both stages of face processing.

Converging evidence suggests that the two-stage and/or the anthropomorphism–dehumanization process helps to account for the perception of artificial faces, including those that elicit the uncanny feeling. Therefore, the proposed process should be a necessary but not sufficient condition for perceiving uncanniness in artificial faces. To test this hypothesis, we propose an innovative paradigm. In line with previous studies, we focus on highly humanlike androids or computer-generated characters and use static images of their faces. In contrast to some studies, we require that the human replicas possess no experimenter-manipulated disproportionate, mismatched, or unattractive features. An example of the stimuli that meet those criteria includes those recently used in Mathur and Reichling (2016, see Figure 1).

First, we propose to ask participants to rate the human likeness and uncanniness of the selected human replicas. Based on the ratings on uncanniness, we categorize them into the reportedly uncanny human replicas and those that do not elicit the uncanny feeling (e.g., mechanical-looking robots). In addition, we include objects and faces of real people as controls.

We then propose to present stimuli from each category (e.g., uncanny androids, normally looking robots, humans, and nonhuman objects) for either 100 or 400 ms in counterbalanced order. After each stimulus disappears, participants must classify it as either human or nonhuman. According to previous findings, the duration of presentation affects the different stages of face processing with regard to anthropomorphism and dehumanization processes. By manipulating the presentation duration, we should be able to dissociate these processing stages while participants perceive images from each category of stimuli. At the same time, we will ask participants to rate their affective responses and perceived animacy to the stimuli under each presentation condition. We predict that if the uncanniness is attributable to mind perception (as suggested by Gray & Wegner, 2012), participants should rate the previously reported uncanny androids as high in perceived animacy (e.g., comparable with real people) and high uncanniness when they are presented for 400 ms but low in perceived animacy for 100 ms. In contrast, if uncanniness is attributable to dehumanization (as we suggest in the Dehumanization hypothesis), participants should rate the uncanny androids as high in perceived animacy (e.g., comparable with real people) when they are presented for 100 ms, but low in perceived animacy for 400 ms. They should also rate uncanniness as high when the uncanny androids are presented for 400 ms but not for 100 ms.

Conclusions and Future Directions

In summary, existing evidence neither completely refutes nor corroborates the uncanny valley hypothesis. Whether the uncanny valley exists is still an open question. The inconsistent findings not only reflect a lack of systematic methodologies across studies to manipulate human likeness and measure likability, but also highlight the need to gain deeper understandings of the independent and dependent variables in this hypothesis.

Existing explanations of the uncanny phenomenon, although providing plausible accounts of the perceptual and cognitive processes by which the uncanny feeling may arise, have not tested the assumption that when the uncanny phenomenon occurs, human observers would spontaneously perceive the human replica as a person. Future research should test this assumption during human–android interactions. Researchers and theorists have recently highlighted the need to examine how human observers perceive human replicas instead of focusing merely on their physical appearance (Angelucci et al., 2014; Gray & Wegner, 2012). By testing the Dehumanization hypothesis, we hope that researchers will be able to better understand the cognitive underpinnings of the uncanny phenomenon, and eventually define and manipulate human likeness to further explore the uncanny valley hypothesis.

Finally, by addressing the existence of the uncanny valley and trying to account for the uncanny feeling elicited by some realistic human replicas, we proposed to distinguish the uncanny valley hypothesis from the uncanny phenomenon. We argue that they focus on different problems. The uncanny phenomenon focuses on the psychological question concerning humans' perception of human replicas and their emotional reactions in terms of the uncanny feeling. In contrast, the uncanny valley hypothesis raises an engineering problem regarding the possibility of creating androids that increasingly resemble humans and may eventually be indistinguishable from real people. As Brenton et al. (2005) argued, "The existence of an uncanny reaction does not validate the valley model" (p. 2). We maintain that this distinction would open more opportunities for future research to examine the uncanny phenomenon without being limited by a need to detect the uncanny valley, whose very existence remains in question.

Further investigation is needed to unpack the uncanny phenomenon. Still, the uncanny valley remains a popular topic with practitioners outside of academia, particularly in computer animation and games. Will robot designers and computer engineers eventually overcome the uncanny valley? The answer to this question lies in rapidly developing technologies in computer graphics.

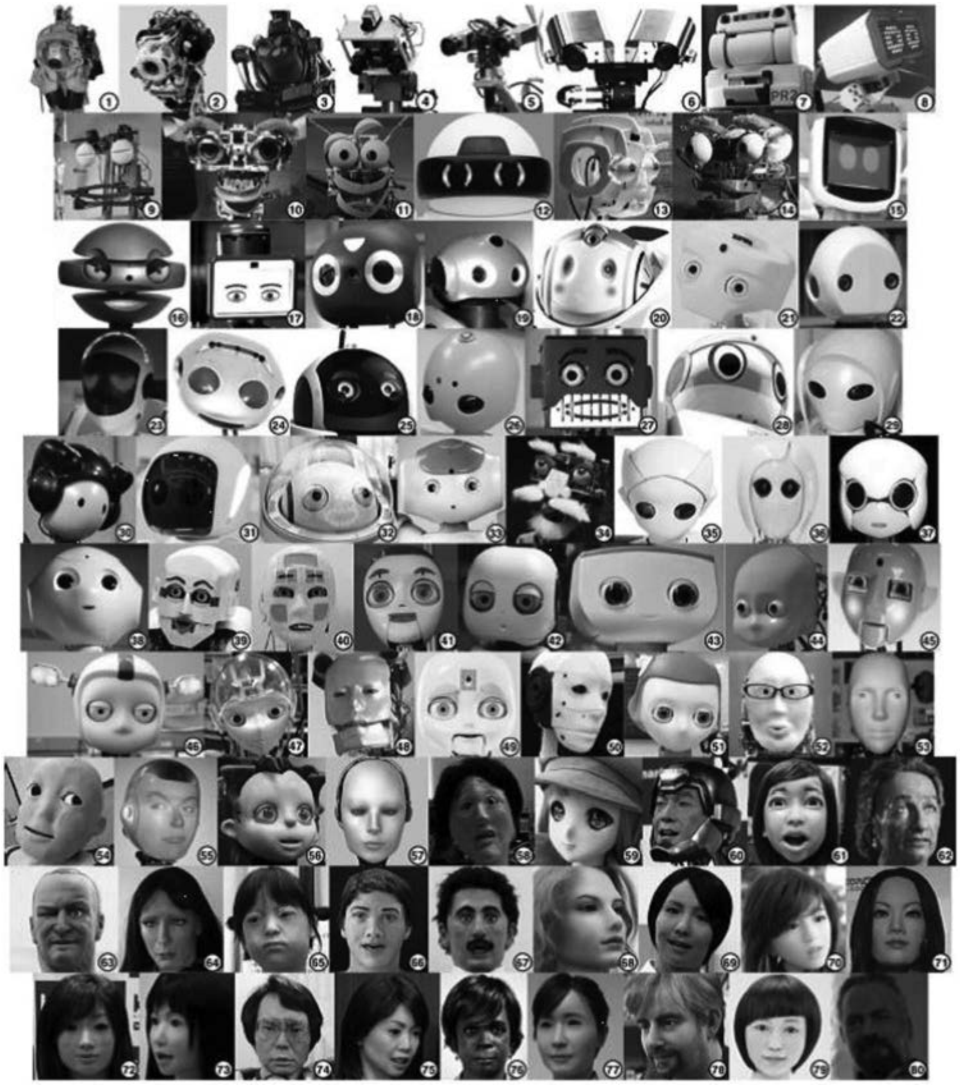

Figure 1. Human replica face images numbered from low to high human likeness based on participants' ratings. From "Navigating a Social World with Robot Partners: A Quantitative Cartography of the Uncanny Valley," by M. B. Mathur and D. B. Reichling, 2016, *Cognition*, *146*, pp. 22–32. Copyright 2015 by Elsevier. Adapted with permission.

Almost three decades ago, Pixar's *Tin Toy* (Lasseter, 1988) became the first computer-animated film to win an Academy Award for Best Animated Short Film. The pseudorealistic baby character "Billy," however, appeared uncanny to many audiences (Hsu & InnovationNewsDaily, 2012), and by today's standards, it is far from being realistic. Mori (1970/2005) appeared justified to argue that designers should avoid creating highly realistic characters. Nevertheless, recent leaps in computer graphics such as real-time rendering have allowed game developers to design lifelike characters, including Nvidia's Digital Ira (Colaner, 2013) and Square Enix's crying human face (White, 2015). Rob Pieke, an expert in computer graphics and the software lead of MPC Film, believes that to create completely photorealistic computer-generated characters is only a matter of time (De Semlyen, 2014). If so, the need to understand the nature and boundary conditions of the uncanny phenomenon will become all the more pressing.

References

Angelucci, A., Bastioni, M., Graziani, P., & Rossi, M. G. (2014). A philosophical look at the uncanny valley. In J. Seibt, R. Hakli, & M. Nørskov (Eds.), *Sociable robots and the future of social relations: Proceedings of robo-philosophy, 2014* (pp. 165–169).

Archer, J. (1997). Why do people love their pets? *Evolution and Human Behavior*, *18*, 237–259. http://dx.doi.org/10.1016/S0162-3095(99)80001-4

Bain, P., Park, J., Kwok, C., & Haslam, N. (2009). Attributing human uniqueness and human nature to cultural groups: Distinct forms of subtle dehumanization. *Group Processes & Intergroup Relations*, *12*, 789–805. http://dx.doi.org/10.1177/1368430209340415

Balas, B., & Koldewyn, K. (2013). Early visual ERP sensitivity to the species and animacy of faces. *Neuropsychologia, 51,* 2876–2881. http://dx.doi.org/10.1016/j.neuropsychologia.2013.09.014

Bandura, A. (1999). Moral disengagement in the perpetration of inhumanities. *Personality and Social Psychology Review, 3,* 193–209. http://dx.doi.org/10.1207/s15327957pspr0303_3

Bandura, A., Underwood, B., & Fromson, M. E. (1975). Disinhibition of aggression through diffusion of responsibility and dehumanization of victims. *Journal of Research in Personality, 9,* 253–269. http://dx.doi.org/10.1016/0092-6566(75)90001-X

Barrett, J. L., & Keil, F. C. (1996). Conceptualizing a nonnatural entity: Anthropomorphism in God concepts. *Cognitive Psychology, 31,* 219–247. http://dx.doi.org/10.1006/cogp.1996.0017

Bartneck, C., Kanda, T., Ishiguro, H., & Hagita, N. (2007). Is the uncanny valley an uncanny cliff? *Proceedings of the 16th IEEE International Symposium on Robot and Human Interactive Communication,* RO-MAN 2007, Jeju, Korea (pp. 368–373). http://dx.doi.org/10.1109/ROMAN.2007.4415111

Bartneck, C., Kanda, T., Ishiguro, H., & Hagita, N. (2009). My robotic doppelgänger—A critical look at the uncanny valley. *The 18th IEEE International Symposium on Robot and Human Interactive Communication*, Toyama, Japan (pp. 269–276). http://dx.doi.org/10.1109/ROMAN.2009.5326351

Bartneck, C., Kulić, D., Croft, E., & Zoghbi, S. (2009). Measurement instruments for the anthropomorphism, animacy, likeability, perceived intelligence, and perceived safety of robots. *International Journal of Social Robotics*, *1,* 71–81. http://dx.doi.org/10.1007/s12369-008-0001-3

Brenton, H., Gillies, M., Ballin, D., & Chatting, D. (2005). *The uncanny valley: Does it exist? Proceedings of conference of human computer interaction, workshop on human animated character interaction*.

Burleigh, T. J., Schoenherr, J. R., & Lacroix, G. L. (2013). Does the uncanny valley exist? An empirical test of the relationship between eeriness and the human likeness of digitally created faces. *Computers in Human Behavior*, *29,* 759–771. http://dx.doi.org/10.1016/j.chb.2012.11.021

Cheetham, M., Suter, P., & Jäncke, L. (2011). The human likeness dimension of the "uncanny valley hypothesis": Behavioral and functional MRI findings. *Frontiers in Human Neuroscience*, *5,* 126. http://dx.doi.org/10.3389/fnhum.2011.00126

Cheetham, M., Suter, P., & Jäncke, L. (2014). Perceptual discrimination difficulty and familiarity in the Uncanny Valley: More like a "Happy Valley." *Frontiers in Psychology, 5,* 1219. http://dx.doi.org/10.3389/fpsyg.2014.01219

Cheetham, M., Wu, L., Pauli, P., & Jäncke, L. (2015). Arousal, valence, and the uncanny valley: Psychophysiological and self-report findings. *Frontiers in Psychology, 6,* 981. http://dx.doi.org/10.3389/fpsyg.2015.00981

Colaner, S. (2013, May 20). NVIDIA's impressive "Digital Ira" FaceWorks demo app now available for download. *Hothardware*. Retrieved from http://hothardware.com/news/nvidias-impressive-digital-ira-faceworks-demo-app-now-available-for-download

Daily Mail Reporter. (2011, August 19). The King is dead: Burger King creepy mascot dethroned in image makeover for struggling chain. *Daily Mail*. Retrieved from http://www.dailymail.co.uk/news/article-2028059/The-king-dead-Burger-King-creepy-mascot-dethroned-image-makeover-struggling-chain.html

Dering, B., Martin, C. D., Moro, S., Pegna, A. J., & Thierry, G. (2011). Face-sensitive processes one hundred milliseconds after picture onset. *Frontiers in Human Neuroscience, 5,* 93. http://dx.doi.org/10.3389/fnhum.2011.00093

Dering, B., Martin, C. D., & Thierry, G. (2009). Is the N170 peak of visual event-related brain potentials car-selective? *NeuroReport: For Rapid Communication of Neuroscience Research, 20,* 902–906. http://dx.doi.org/10.1097/WNR.0b013e328327201d

De Semlyen, P. (2014). In the future . . . Cinema will cross the uncanny valley. *Empireonline*. Retrieved from http://www.empireonline.com/features/future-of-vfx

Epley, N., Waytz, A., & Cacioppo, J. T. (2007). On seeing human: A three-factor theory of anthropomorphism. *Psychological Review, 114,* 864–886. http://dx.doi.org/10.1037/0033-295X.114.4.864

Farid, H., & Bravo, M. J. (2012). Perceptual discrimination of computer-generated and photographic faces. *Digital Investigation*, *8,* 226–235. http://dx.doi.org/10.1016/j.diin.2011.06.003

Freud, S. (1964). The uncanny. In J. Strachey & A. Freud (Eds. And Trans.), *The standard edition of the complete psychological works of Sigmund Freud* (pp. 217–256). London, UK: Hogarth Press. (Original work published 1919)

Geller, T. (2008). Overcoming the uncanny valley. *IEEE Computer Graphics and Applications, 28,* 11–17. http://dx.doi.org/10.1109/MCG.2008.79

Gray, H. M., Gray, K., & Wegner, D. M. (2007). Dimensions of mind perception. *Science, 315,* 619. http://dx.doi.org/10.1126/science.1134475

Gray, K., & Wegner, D. M. (2012). Feeling robots and human zombies: Mind perception and the uncanny valley. *Cognition, 125,* 125–130. http://dx.doi.org/10.1016/j.cognition.2012.06.007

Green, R. D., MacDorman, K. F., Ho, C.-C., & Vasudevan, S. (2008). Sensitivity to the proportions of faces that vary in human likeness. *Computers in Human Behavior, 24,* 2456–2474. http://dx.doi.org/10.1016/j.chb.2008.02.019

Guthrie, S. (1993). *Faces in the clouds*. New York, NY: Oxford University Press.

Hanson, D. (2005). Expanding the aesthetic possibilities for humanoid robots. In IEEE-RAS international conference on humanoid robots, Tsukuba, Japan.

Hanson, D., Olney, A., Pereira, I. A., & Zielke, M. (2005). *Upending the uncanny valley*. Proceedings of the Twentieth National Conference on Artificial Intelligence (AAAI-05). Presented at the Twentieth National Conference on Artificial Intelligence (AAAI-05), Pittsburgh, PA.

Haslam, N. (2006). Dehumanization: An integrative review. *Personality and Social Psychology Review, 10,* 252–264. http://dx.doi.org/10.1207/s15327957pspr1003_4

Haslam, N., & Loughnan, S. (2014). Dehumanization and infrahumanization. *Annual Review of Psychology*, *65,* 399–423. http://dx.doi.org/10.1146/annurev-psych-010213-115045

Haxby, J. V., Hoffman, E. A., & Gobbini, M. I. (2000). The distributed human neural system for face perception. *Trends in Cognitive Sciences, 4,* 223–233. http://dx.doi.org/10.1016/S1364-6613(00)01482-0

Ho, C.-C., & MacDorman, K. F. (2010). Revisiting the uncanny valley theory: Developing and validating an alternative to the Godspeed indices. *Computers in Human Behavior*, *26,* 1508–1518. http://dx.doi.org/10.1016/j.chb.2010.05.015

Ho, C.-C., MacDorman, K. F., & Pramono, Z. A. (2008). *Human emotion and the uncanny valley: A GLM, MDS, and Isomap analysis of robot video ratings. In Proceedings of the 3rd ACM/IEEE international conference on Human robot interaction, HRI '08* (pp. 169–176). Presented at the HRI '08, New York, NY. http://dx.doi.org/10.1145/1349822.1349845

Hoffmann, E. T. A. (1994). *The sandman* (Trans. John Oxenford). 19th Century German Stories (Ed. Robert Godwin-Jones). (Original work published 1817)

Hsu, J. (2012, June 12). Robotics' Uncanny Valley gets new translation. *Livescience*. Retrieved from http://www.livescience.com/20909-robotics-uncanny-valley-translation.html

Hsu, J., & InnovationNewsDaily. (2012, April 3). Why "Uncanny Valley" human look-alikes put us on edge. *Scientific American*. Retrieved from http://www.scientificamerican.com/article/why-uncanny-valley-human-look-alikes-put-us-on-edge/

Jentsch, E. (1997). On the psychology of the uncanny (1906). *Angelaki, 2,* 7–16. http://dx.doi.org/10.1080/09697259708571910

Kaba, F. (2013). Hyper-realistic characters and the existence of the uncanny valley in animation films. *International Review of Social Sciences and Humanities, 4,* 188–195.

Kätsyri, J., Förger, K., Mäkäräinen, M., & Takala, T. (2015). A review of empirical evidence on different uncanny valley hypotheses: Support for perceptual mismatch as one road to the valley of eeriness. *Frontiers in Psychology, 6,* 390. http://dx.doi.org/10.3389/fpsyg.2015.00390

Kelman, H. (1973). Violence without moral restraint: Reflections on the dehumanization of victims and victimizers. *Journal of Social Issues, 29,* 25–61. http://dx.doi.org/10.1111/j.1540-4560.1973.tb00102.x

Lasseter, J. (1988). *Tin toy* [film]. Emeryville: Pixar.

Lewkowicz, D. J., & Ghazanfar, A. A. (2012). The development of the uncanny valley in infants. *Developmental Psychobiology, 54,* 124–132. http://dx.doi.org/10.1002/dev.20583

Leyens, J.-P., Demoulin, S., Vaes, J., Gaunt, R., & Paladino, M. P. (2007). Infra-humanization: The wall of group differences. *Social Issues and Policy Review, 1,* 139–172. http://dx.doi.org/10.1111/j.1751-2409.2007.00006.x

Leyens, J.-P., Rodriguez-Perez, A., Rodriguez-Torres, R., Gaunt, R., Paladino, M.-P., Vaes, J., & Demoulin, S. (2001). Psychological essentialism and the differential attribution of uniquely human emotions to ingroups and outgroups. *European Journal of Social Psychology, 31,* 395–411. http://dx.doi.org/10.1002/ejsp.50

Looser, C. E., Guntupalli, J. S., & Wheatley, T. (2013). Multivoxel patterns in face-sensitive temporal regions reveal an encoding schema based on detecting life in a face. *Social Cognitive and Affective Neuroscience, 8,* 799–805. http://dx.doi.org/10.1093/scan/nss078

Looser, C. E., & Wheatley, T. (2010). The tipping point of animacy. How, when, and where we perceive life in a face. *Psychological Science, 21,* 1854–1862. http://dx.doi.org/10.1177/0956797610388044

MacDorman, K. (2006). *Subjective ratings of robot video clips for human likeness, familiarity, and eeriness: An exploration of the Uncanny Valley*. *ICCS/CogSci-2006 Long Symposium: Toward Social Mechanisms of Android Science* (pp. 26–29).

MacDorman, K. F., & Entezari, S. O. (2015). Individual differences predict sensitivity to the uncanny valley. *Interaction Studies: Social Behaviour and Communication in Biological and Artificial Systems, 16,* 141–172.

MacDorman, K. F., Green, R. D., Ho, C.-C., & Koch, C. T. (2009). Too real for comfort? Uncanny responses to computer generated faces. *Computers in Human Behavior, 25,* 695–710. http://dx.doi.org/10.1016/j.chb.2008.12.026

MacDorman, K. F., & Ishiguro, H. (2006). The uncanny advantage of using androids in cognitive and social science research. *Interaction Studies: Social Behaviour and Communication in Biological and Artificial Systems, 7,* 297–337. http://dx.doi.org/10.1075/is.7.3.03mac

MacDorman, K. F., Minato, T., Shimada, M., Itakura, S., Cowley, S., & Ishiguro, H. (2005). *Assessing human likeness by eye contact in an android testbed*. CogSci2005 - XXVII Annual Conference of the Cognitive Science Society.

Macy, R., & Schrader, V. (2008). Pediophobia: A new challenge facing nursing faculty in clinical teaching by simulation. *Clinical Simulation in Nursing, 4,* e89–e91. http://dx.doi.org/10.1016/j.ecns.2008.07.001

Mathur, M. B., & Reichling, D. B. (2016). Navigating a social world with robot partners: A quantitative cartography of the Uncanny Valley. *Cognition, 146,* 22–32. http://dx.doi.org/10.1016/j.cognition.2015.09.008

McCarthy, A., & Muir, D. (2003). *Eye movements as social signals during thinking: Age differences*. In *Biennial Meeting of the Society for Research in Child Development*.

Misselhorn, C. (2009). Empathy with inanimate objects and the uncanny valley. *Minds and Machines, 19,* 345–359. http://dx.doi.org/10.1007/s11023-009-9158-2

Mitchell, W. J., Szerszen, K. A., Sr., Lu, A. S., Schermerhorn, P. W., Scheutz, M., & Macdorman, K. F. (2011). A mismatch in the human realism of face and voice produces an uncanny valley. *i-Perception, 2,* 10–12. http://dx.doi.org/10.1068/i0415

Monahan, J. L. (1998). I don't know it but I like you: The influence of nonconscious affect on person perception. *Human Communication Research, 24,* 480–500. http://dx.doi.org/10.1111/j.1468-2958.1998.tb00428.x

Moore, R. K. (2012). A Bayesian explanation of the 'Uncanny Valley' effect and related psychological phenomena. *Scientific Reports, 2,* 864.

Mori, M. (1970/2005). The uncanny valley. (K. F. MacDorman, & T. Minato, Trans.). *Energy, 7,* 33–35.

Mori, M., MacDorman, K. F., & Kageki, N. (2012). The uncanny valley [from the field]. *Robotics & Automation Magazine, IEEE, 19,* 98–100. http://dx.doi.org/10.1109/MRA.2012.2192811

Nass, C., Steuer, J., & Tauber, E. R. (1994, April). *Computers are social actors*. In *Proceedings of the SIGCHI conference on Human factors in computing systems* (pp. 72–78). ACM.

Opotow, S. (1990). Moral exclusion and injustice: An introduction. *Journal of Social Issues, 46,* 1–20. http://dx.doi.org/10.1111/j.1540-4560.1990.tb00268.x

Piwek, L., McKay, L. S., & Pollick, F. E. (2014). Empirical evaluation of the uncanny valley hypothesis fails to confirm the predicted effect of motion. *Cognition, 130,* 271–277. http://dx.doi.org/10.1016/j.cognition.2013.11.001

Poliakoff, E., Beach, N., Best, R., Howard, T., & Gowen, E. (2013). Can looking at a hand make your skin crawl? Peering into the uncanny valley for hands. *Perception, 42,* 998–1000. http://dx.doi.org/10.1068/p7569

Pyszczynski, T., Greenberg, J., & 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,* 835–845. http://dx.doi.org/10.1037/0033-295X.106.4.835

Ramey, C. H. (2006). An inventory of reported characteristics for home computers, robots, and human beings: Applications for android science and the uncanny valley. In *Proceedings of the ICCS/CogSci-2006 Long Symposium 'Toward Social Mechanisms of Android Science', Vancouver, Canada*.

Reichardt, J. (1978). *Robots: fact, fiction, and prediction*. London, United Kingdom: Thames & Hudson Ltd.

Rozin, P., & Fallon, A. E. (1987). A perspective on disgust. *Psychological Review, 94,* 23–41. http://dx.doi.org/10.1037/0033-295X.94.1.23

Saygin, A. P., Chaminade, T., Ishiguro, H., Driver, J., & Frith, C. (2012). The thing that should not be: Predictive coding and the uncanny valley in perceiving human and humanoid robot actions. *Social Cognitive and Affective Neuroscience, 7,* 413–422. http://dx.doi.org/10.1093/scan/nsr025

Schneps, L., & Colmez, C. (2013). *Math on trial: How numbers get used and abused in the courtroom*. New York, NY: Basic Books.

Seyama, J., & Nagayama, R. S. (2007). The uncanny valley: Effect of realism on the impression of artificial human faces. *Presence (Cambridge, Mass.), 16,* 337–351. http://dx.doi.org/10.1162/pres.16.4.337

Shimada, M., Minato, T., Itakura, S., & Ishiguro, H. (2007, August). *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* (pp. 374–379). IEEE.

Shultz, S., & McCarthy, G. (2014). Perceived animacy influences the processing of human-like surface features in the fusiform gyrus. *Neuropsychologia, 60,* 115–120. http://dx.doi.org/10.1016/j.neuropsychologia.2014.05.019

Soanes, C., & Stevenson, A. (Eds.). (2005). *Oxford dictionary of English* (2nd ed.). New York, NY: Oxford University Press.

Staub, E. (1989). *The roots of evil: The origins of genocide and other group violence*. New York, NY: Cambridge University Press.

Steckenfinger, S. A., & Ghazanfar, A. A. (2009). Monkey visual behavior falls into the uncanny valley. *PNAS Proceedings of the National Academy of Sciences of the United States of America, 106,* 18362–18366. http://dx.doi.org/10.1073/pnas.0910063106

Struch, N., & Schwartz, S. H. (1989). Intergroup aggression: Its predictors and distinctness from in-group bias. *Journal of Personality and Social Psychology, 56,* 364–373. http://dx.doi.org/10.1037/0022-3514.56.3.364

Thierry, G., Martin, C. D., Downing, P., & Pegna, A. J. (2007). Controlling for interstimulus perceptual variance abolishes N170 face selectivity. *Nature Neuroscience, 10,* 505–511.

Tinwell, A., Grimshaw, M., Nabi, D. A., & Williams, A. (2011). Facial expression of emotion and perception of the uncanny valley in virtual characters. *Computers in Human Behavior, 27,* 741–749. http://dx.doi.org/10.1016/j.chb.2010.10.018

Tinwell, A., Nabi, D. A., & Charlton, J. P. (2013). Perception of psychopathy and the Uncanny Valley in virtual characters. *Computers in Human Behavior, 29,* 1617–1625. http://dx.doi.org/10.1016/j.chb.2013.01.008

Veale, T. (2004). Incongruity in humor: Root cause or epiphenomenon? *Humor: International Journal of Humor Research, 17,* 419–428. http://dx.doi.org/10.1515/humr.2004.17.4.419

Viki, G. T., Winchester, L., Titshall, L., Chisango, T., Pina, A., & Russell, R. (2006). Beyond secondary emotions: The infrahumanization of outgroups using human-related and animal-related words. *Social Cognition, 24,* 753–775. http://dx.doi.org/10.1521/soco.2006.24.6.753

Wheatley, T., Weinberg, A., Looser, C., Moran, T., & Hajcak, G. (2011). Mind perception: Real but not artificial faces sustain neural activity beyond the N170/VPP. *PLoS ONE, 6,* e17960. http://dx.doi.org/10.1371/journal.pone.0017960

White, G., McKay, L., & Pollick, F. (2007). Motion and the uncanny valley. *Journal of Vision, 7,* 477.

White, J. (2015, April 30). Square Enix, Microsoft, NVIDIA point way through Uncanny Valley at Microsoft BUILD. *NVIDIA*. Retrieved from http://blogs.nvidia.com/blog/2015/04/30/build-demo/

Wolfe, P. (1997). *In the Zone: The twilight world of Rod Serling*. Bowling Green, OH: Popular Press.

Yamada, Y., Kawabe, T., & Ihaya, K. (2013). Categorization difficulty is associated with negative evaluation in the "uncanny valley" phenomenon. *Japanese Psychological Research, 55,* 20–32. http://dx.doi.org/10.1111/j.1468-5884.2012.00538.x

Zemeckis, R. (2004). *The polar express* [film]. Los Angeles, CA: Warner Bros.

Złotowski, J., Proudfoot, D., & Bartneck, C. (2013). *More human than human: Does the uncanny curve really matter?* Proceedings of the HRI2013 Workshop on Design of Humanlikeness in HRI from uncanny valley to minimal design, Tokyo, pp. 713.

Received August 9, 2015
Revision received October 11, 2015
Accepted October 12, 2015 ■