Contents lists available at ScienceDirect

Computers in Human Behavior: Artificial Humans

journal homepage: www.journals.elsevier.com/computers-in-human-behavior-artificial-humans

Does mind perception explain the uncanny valley? A meta-regression analysis and (de)humanization experiment

Karl F. MacDorman

Luddy School of Informatics, Computing, and Engineering, Indiana University, 535 West Michigan Street, Indianapolis, IN, 46202, USA

ARTICLE INFO

Keywords:
Anthropomorphism
Dehumanization
Human–computer interaction
Sentience
Uncanny valley of mind

ABSTRACT

Gray and Wegner (2012) proposed that when robots look human, their appearance prompts attributions of experience, including sensations and feelings, which is uncanny. This theory, confusingly termed *mind perception*, differs from perceptual theories of the uncanny valley in that the robots' eeriness is not stimulus-driven. To explore this seminal theory, we conducted a meta-regression analysis of 10 experiments and a (de)humanization experiment. In the first part, experiments were identified in the literature that manipulated artificial entity's experience using descriptions. However, experiments with no observable stimuli yielded larger effects for experience and eeriness than those with robots and virtual reality characters. This finding undermines a theory that purports to explain how a robot's human likeness causes eeriness. Further, a second issue concerns Gray and Wegner's protocol based on a vignette design. Reading about an entity with experience activates thoughts that may not be activated when encountering it, and these thoughts may increase its eeriness. Therefore, the paper's second part focuses on an experiment we conducted with a novel humanization–dehumanization protocol. Participants' attitudes on robots' similarity to humans were gradually shifted to manipulate robots' perceived humanness, experience, and agency. However, the manipulation's effect on eeriness and coldness was mostly nonsignificant or counter to prediction. Differences in the robots' physical appearance had a much larger effect on their eeriness and coldness. In fact, as a mediator, experience mitigated the stimulus's overall effect of increasing eeriness. These results favor perceptual theories, rather than mind perception, in explaining the uncanny valley.

1. Introduction

In 1970, Mori (2012) made a groundbreaking observation: If a robot is given a moderately human appearance, people will feel more affinity for it, up to a point, shown as the first peak in his graph (Fig. 1). People would feel even more affinity if the robot could be made indistinguishable from a healthy person, placed at the second peak. However, between these two peaks, the robot risks appearing emotionally cold and eerie. Mori called this effect the uncanny valley.

Large-scale studies have plotted the uncanny valley: Participants rated 182 headshots of robot and human faces in Mathur et al. (2020, $n = 358$) and 251 full-body shots of robots in Kim, de Visser, and Phillips (2022, $n = 539$). A meta-analysis of 247 measured effects from 56 papers published between 2008 and 2021 determined that the uncanny valley has a large effect size (Diel, Weigelt, & MacDorman, 2022).

Many theories have been proposed to explain the uncanny valley (Diel & MacDorman, 2021; Kätsyri, Förger, Mäkäräinen, & Takala, 2015; Wang, Lilienfeld, & Rochat, 2015). One stands out for its influence on the literature, garnering much attention with recent advances in AI—Gray and Wegner's (2012) mind perception theory:

> We propose that humanlike robots are [unnerving] because their appearance prompts attributions of mind.… [M]achines become unnerving when people ascribe to them experience (the capacity to feel and sense) rather than agency (the capacity to act and do). (p. 125)

Building on Gray and Wegner's work, our research frames the mind perception theory thus: A human appearance prompts attributions of experience to an entity, which elicits feelings of eeriness when the entity is known to be artificial (Fig. 2).

1.1. A meta-regression analysis with moderation

In reviewing the mind perception literature, 10 experiments conducted by five independent research teams were identified. These experiments described an artificial entity with or without experience or other attributes in a vignette design. Their authors interpreted their

E-mail address: kmacdorm@iu.edu.

https://doi.org/10.1016/j.chbah.2024.100065
Received 9 February 2024; Received in revised form 19 March 2024; Accepted 19 March 2024

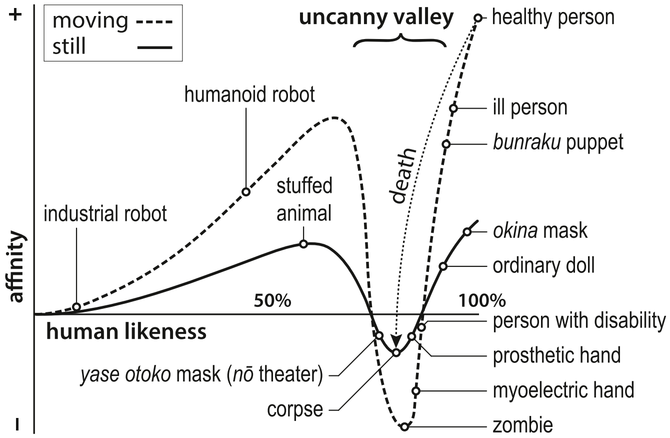


Fig. 1. Mori's (2012) graph depicts affinity for an entity as a function of its human likeness and whether it is still or moving.

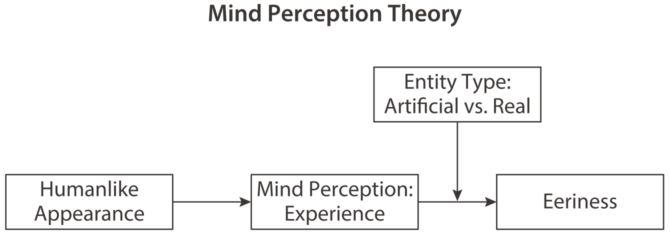


Fig. 2. The mind perception theory states that a humanlike appearance increases attributions of experience, which influences eeriness, mediated by entity type such that artificial entities increase eeriness.

findings as supporting the mind perception theory.

In Gray and Wegner (2012), a supercomputer described as having experience—able to feel "hunger, fear, and other emotions" (p. 127)—rated uncannier than a supercomputer described as having agency or the control condition. This effect was reproduced for a similarly described robot, chatbot, and smart speaker (Appel, Izydorczyk, Weber, Mara, & Lischetzke, 2020; Lu, 2021; Taylor, Weiss, & Marshall, 2020). The effect has been reproduced in studies that accompanied the descriptions with observable stimuli. Stein and Ohler (2017) presented participants with a human couple in virtual reality. Yam, Bigman, and Gray (2021) reversed the effect through dehumanization. Robot videos and physical robots were less uncanny after being described as lacking "the ability to feel" and "experience love, desire, or any emotion" (pp. 4 and 6).

A pattern was observed in the findings of these papers. Experiments that only provided a written description of the artificial entity reported larger effect sizes than those that also presented the entity's physical appearance. It is odd that the entity would be eerier when absent, given that mind perception theory purports to explain why a machine's humanlike appearance is eerie.

1.1.1. *Eerier when absent?*

Our research question is framed as follows: How does the presence or absence of the stimulus influence experience and eeriness in prior mind perception experiments?

To answer this question, a meta-regression analysis was performed with construct (experience or eeriness) and stimulus (present or absent) as focal moderator variables.

1.2. *(De)humanization experiment*

Although ascribing experience to an artificial entity increased its eeriness in mind perception experiments, Gray and Wegner's (2012) protocol may have a flaw: In the experimental condition, participants were told whether an entity (or its class) possessed or lacked experience before rating it on eeriness (or a synonymous construct like uncanniness). So, participants in the experience group were, in a sense, primed. Since they had just read a description that ascribed experience to the entity, when they rated the entity, experience-related concepts would be more active than when participants in other groups rated the same entity. However, this kind of priming is missing from typical encounters with technology in society. Therefore, we must determine whether heightened attributions of experience still increase eeriness when experience-related concepts are not activated in this way.

To investigate this phenomenon, a novel protocol is proposed: Participants rate android robots shown in videos. Then, through readings and writing assignments, their attitudes are influenced over five weeks. The dehumanization treatment shifts participants' attitudes toward believing that robots and computers cannot think or feel. The humanization treatment has the opposite effect. Dehumanization is predicted to reduce attributions of experience to the robots and, as a result, their eeriness, whereas humanization is predicted to increase attributions of experience and, as a result, their eeriness. The robots are rated again one week after the treatment; yet, due to this washout period, experience-related concepts are not directly activated by the vignette design as with Gray and Wegner (2012). Therefore, the experiment should determine whether—with only the stimuli directly activating experience-related concepts—heightened attributions of experience increase eeriness.

1.2.1. *Hypotheses*

Mori (2012) identifies the uncanny valley with eeriness (Japanese: *bukimi*, 不気味) and low or negative affinity (*shinwakan*, 親和感). Eeriness is the experiential quality of the uncanny. Affinity is an important covariate, identified in the social psychology literature as a *warmth–coldness* dimension, the primary dimension of person perception (Fiske, Cuddy, Glick, & Xu, 2002; Fiske, Cuddy, & Glick, 2007). This paper operationalizes the uncanny valley effect as self-reported eeriness and coldness ratings using validated indices for these constructs (Ho & MacDorman, 2010, 2017). Gray and Wegner's (2012) theory was interpreted as making this prediction: Hypothesis 1. Ascribing experience to android robots elicits feelings of eeriness and coldness.

However, other processes may cause eeriness. If eeriness is mainly stimulus-driven, observed differences among the androids should have a larger effect on eeriness than differences in their experience. Thus, the alternative view makes this prediction: Hypothesis 2. The robots' physical appearance has a larger effect on their eeriness and coldness than ascribing experience to them.

2. A meta-regression analysis with moderation

2.1. *Method*

The analysis aimed to determine how the presence or absence of the stimulus influences experience and eeriness in mind perception experiments. The approach was to identify relevant literature, calculate effect sizes and their variances, and perform a moderation analysis by fitting a mixed-effects meta-regression model.

2.1.1. *Inclusion criteria*

Experiments were included if they manipulated descriptions of an artificial entity at different experience levels as the independent variable, measured experience and uncanniness (or synonymous constructs) as dependent variables, and provided enough information to calculate their effect sizes and variances.

2.1.2. *Study search and selection*

The search used Google Scholar, retrieving papers published from 2012 to 2023 that cited Gray and Wegner (2012), which yielded 579 results. The titles and abstracts of these papers were read, and their contents were searched using terms like *experiment*, *mind perception*, and *uncanny valley*.

2.1.3. Data preparation, analysis, and reporting

The selected studies had control and experimental conditions. In the latter, the artificial entity was described as having or lacking experience. For each condition, its mean score, standard deviation, and group size were recorded for the experience and eeriness constructs.

Hedges' g and its variance were calculated to correct for positive bias in studies with small group sizes, n_1 and n_2:

$$g = d\left(1 - \frac{3}{4\,(n_1 + n_2) - 9}\right) \tag{1}$$

$$v_g = v_d\left(1 - \frac{3}{4\,(n_1 + n_2) - 9}\right)^2 \tag{2}$$

To estimate g, Cohen's d and its variance were first calculated. Since within-group correlations were unavailable in the repeated measures study (Lu, 2021), following Lakens (2013), d_{av} and its variance were used:

$$d_{av} = \frac{m_1 - m_2}{\frac{1}{2}(s_1 + s_2)} \tag{3}$$

$$v_{d_{av}} = \frac{n_1 + n_2}{n_1 n_2} + \frac{d^2}{2(n_1 + n_2)} \tag{4}$$

Moderation analysis was performed by fitting a mixed-effects meta-regression model by restricted maximum-likelihood estimation. Focal moderator variables were construct (experience or eeriness) and stimulus (present or absent).

Following Viechtbauer and Cheung (2010), an effect was considered influential if it met at least one of these four criteria:

$$|DFFITS| > 3\sqrt{\frac{p}{k - p}} \tag{5}$$

where p is the number of model coefficients and k is the number of effects, the Cook's distance is in the upper 50% of the F-distribution with p and $k - p$ degrees of freedom,

$$hat\ value > \frac{3p}{k} \tag{6}$$

or any *DFBETA* is greater than 1.

Effect sizes were reported as Hedges' g and its 95% confidence interval, interpreted with thresholds of 0.2 for small, 0.5 for medium, and 0.8 for large.

Effect size calculation, meta-regression analysis, and influential effects assessment were performed using the R package *metaphor*.

2.2. Results

The search yielded 10 experiments, published in five papers and a thesis, that met the inclusion criteria (Table 1). These experiments manipulated experience in a vignette design. Lu (2021) was a repeated measures experiment, and the others were between-group experiments. The results of seven were coded as stimulus absent because no stimuli were present (Appel et al., 2020; Gray & Wegner, 2012; Lu, 2021; Taylor et al., 2020). The results of three experiments were coded as stimulus present: Stein and Ohler (2017) used a couple in virtual reality, and Yam et al. (2021) used robot videos and physical robots.

Moderation analysis was performed using a mixed-effects meta-regression model with construct (experience or eeriness) and stimulus (present or absent) as moderator variables, $k = 20$, Akaike information criterion (AIC) $= 68.36$, $QE(17) = 138.72$, $p < 0.0001$, $\tau^2 = 0.81$, $I^2 = 0.97$, $R^2 = 0.533$, $QM(2) = 12.01$, $p = 0.0025$. The experience and uncanniness effects from Gray and Wegner (2012) exceeded thresholds for *DFFITS* and Cook's distance. Thus, they were identified as influential points (Fig. 3).

A better fitting model excluded these outliers, $k = 18$, $AIC = 20.66$, $QE(15) = 79.67$, $p < 0.0001$, $\tau^2 = 0.10$, $I^2 = 0.80$, $R^2 = 0.778$, $QM(2) = 45.49$, $p < 0.0001$ (Fig. 4). AIC was substantially lower, $\Delta AIC = -47.70$. The residual heterogeneity decreased from $\tau^2 = 0.81$ to $\tau^2 = 0.10$. The amount of heterogeneity explained increased from $R^2 = 0.533$ to $R^2 = 0.778$, with the revised model explaining 77.8% of the observed heterogeneity.

The model showed that, for experience, the effect size was 232% larger when the stimulus was absent, $g = 1.79$ [1.52, 2.07] (95% CI), $SE = 0.14$, $z = 12.71$, $p < 0.0001$, than when it was present, $g = 0.77$ [0.44, 1.10], $SE = 0.17$, $z = 4.61$, $p < 0.0001$. For eeriness, the effect size was 836% larger when the stimulus was absent, $g = 1.14$ [0.88, 1.44], $SE = 0.14$, $z = 8.37$, $p < 0.0001$, than when it was present, $g = 0.12$ [–0.20, 0.45], $SE = 0.17$, $z = 0.81$, $p = 0.457$. When the stimulus was present, the effect size for eeriness was nonsignificant.

The model showed that the presence of the stimulus led to a significant and substantial overall reduction in effect sizes: $g = -1.02$ [–1.37, -0.67], $SE = 0.18$, $z = -5.68$, $p < 0.0001$. The experience construct led to an overall increase in effect sizes relative to the eeriness construct: g

Table 1
Selected experiments by construct and stimulus with mean scores, standard deviations, and group sizes.

Study	Construct	Stimulus	M_1	SD_1	M_2	SD_2	n_1	n_2
Gray and Wegner (2012) [2]	experience	absent	4.53	0.52	1.03	0.13	15	15
Appel et al. (2020) [1A]	experience	absent	3.21	1.31	1.02	0.11	31	31
Appel et al. (2020) [1B]	experience	absent	3.68	1.25	1.38	0.76	37	37
Appel et al. (2020) [3]	experience	absent	3.16	1.24	1.15	0.45	169	165
Appel et al. (2020) [2]	experience	absent	3.01	1.21	1.22	0.70	112	143
Lu (2021)	experience	absent	3.90	2.00	1.30	0.80	90	90
Taylor et al. (2020)	experience	absent	3.15	1.53	1.17	0.63	29	29
Stein and Ohler (2017)	experience	present	2.55	1.06	1.78	0.80	23	23
Yam et al. (2021) [2]	experience	present	2.13	1.41	1.52	0.92	390	390
Yam et al. (2021) [3]	experience	present	2.30	1.58	1.91	1.26	173	174
Gray and Wegner (2012) [2]	eeriness	absent	3.27	0.61	1.22	0.24	15	15
Taylor et al. (2020)	eeriness	absent	3.67	1.42	1.88	1.09	29	29
Appel et al. (2020) [1A]	eeriness	absent	2.75	1.32	1.36	0.67	31	31
Lu (2021)	eeriness	absent	3.80	1.00	2.40	1.20	90	90
Appel et al. (2020) [1B]	eeriness	absent	2.46	1.04	1.70	0.89	37	37
Appel et al. (2020) [3]	eeriness	absent	2.28	1.00	1.65	0.71	169	165
Appel et al. (2020) [2]	eeriness	absent	2.59	1.25	1.85	0.89	112	143
Stein and Ohler (2017)	eeriness	present	3.43	0.96	2.77	0.59	23	23
Yam et al. (2021) [2]	eeriness	present	2.37	1.26	2.01	1.13	390	390
Yam et al. (2021) [3]	eeriness	present	2.37	1.26	2.01	1.13	173	174

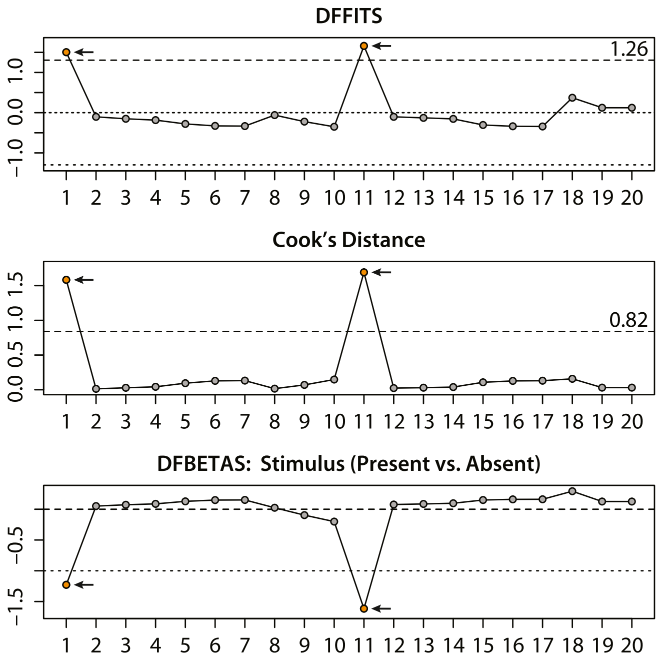


Fig. 3. *DFFITS* (threshold = 1.26), Cook's distance (threshold = 0.82), and *DFBETAS* for the effects in the meta-regression with moderation. Outliers are indicated by arrows.

$= 0.65$ [0.31, 0.99], $SE = 0.17$, $z = 3.74$, $p = 0.0002$.

2.3. Findings

Each study that met our inclusion criteria reported the same result: Ascribing experience to machines causes eeriness. Appel et al. (2020) found ascribing agency also causes eeriness, though the effect was weaker (Experiments 1A, 1B, 2, and 3). Lu (2021) found a stronger effect for agency than experience. Thus, their findings were contrary to Gray and Wegner's (2012) claim that experience but not agency causes eeriness.

Our meta-regression analysis with moderation, however, found that the presentation of an artificial entity's physical appearance significantly decreased the effect of the vignette's experience manipulation. When the stimulus was absent, the effect size was medium-to-large, $g = 0.77$, and significant, $p < 0.0001$, but negligible, $g = 0.12$, and nonsignificant, $p = 0.457$, when present. These findings indicate that the eeriness caused by the artificial entity's appearance masks the effects of mind perception. Given that Mori (2012) defines the uncanny valley as a relation between an entity's *human likeness* and the observer's feelings of affinity and eeriness, the analysis indicates an issue either with the experimental manipulation or with mind perception as an explanation of the uncanny valley.

In sum, these findings indicate the importance of mind perception for disembodied AI. Mind perception's impact will only increase as AI advances (Stein & MacDorman, 2024). However, they also raise concerns about mind perception as an explanation of Mori's uncanny valley.

3. (De)humanization experiment

3.1. Method

The meta-regression findings, that ascribing experience to artificial entities elicits eeriness and their observable presence diminishes it, indicate the vignette was the proximal cause of eeriness, not the entities' human likeness. The vignette likely activated experience-related concepts that increased the entities' perceived eeriness. Nevertheless, an experimental manipulation of experience is important to determine its causal effect on eeriness.

A novel protocol is proposed to shift participant attitudes progressively and then, after a washout period, determine whether this attitude change increased artificial entities' experience and eeriness. The advantage of this approach is that—true to the uncanny valley—the proximal cause of eeriness remains the artificial entities' physical appearance. This tests whether mind perception theory's prediction, that ascribing experience to artificial entities elicits eeriness and coldness, Hypothesis 1, or whether aspects of its physical appearance have a larger effect, Hypothesis 2.

3.1.1. Participants

Out of 133 graduate classmates majoring in human–computer interaction, 127 were included in the data analysis (81 women and 46 men). They ranged in age from 20 to 54 (median = 25, interquartile = 23–27). Of these, 11 were African American or Black, 84 were Asian, 29 were White, 2 were Hispanic, and 1 was Multiracial. Countries of nationality were Burkina Faso, 1; China, 1; India, 71; Nepal, 1; Nigeria, 2; the Philippines, 1; Taiwan, 4; and the United States, 46. Six participants were excluded for not completing the study. Cohort 1 had 40 participants, and Cohort 2 had 87 participants.

A power analysis was performed for the manipulation checks and hypotheses. Published effect sizes for human–robot similarity and humanness ratings, converted to d, ranged from 0.39 to 0.75 (MacDorman & Entezari, 2015). Given three conditions, an effect size $d = 0.33$, and a repeated-measures correlation $\rho = 0.5$, each group requires 40 participants for $p < 0.05$ significance at 0.90 power.

The Indiana University Office of Research Administration accepted this study protocol (No. 12100096830) under 45 CFR 46.101(b) (1–2).

3.1.2. Procedure

The experiment had a cohort crossover design with two independent variables: condition and stimulus. The order of the dehumanization and humanization treatments was cross-balanced. Cohort 1 completed Condition 1 pretest, dehumanization treatment, Condition 2 posttest, humanization treatment, and Condition 3 posttest. Cohort 2 completed Condition 1 pretest, humanization treatment, Condition 2 posttest, dehumanization treatment, and Condition 3 posttest. Thus, the dependent variables were measured three times for each cohort: once in the pretest and again in each posttest.

The pretest and posttests were identical. They involved rating videos of three different androids, presented sequentially, on eeriness, coldness, experience, agency, and humanness indices, and then completing a human–robot similarity index. The robots were rated on eeriness first to eliminate order effects on this critical dependent variable.

The dehumanization treatment advocated the position that humans are unique and different from machines like computers and robots. The humanization treatment opposed that position, arguing that humans and machines could be equivalent in essential respects. Each treatment spanned five weeks.

This study was conducted as a course module. Cohort 1 took the module in the fall of 2021, and Cohort 2 took the module in the spring and fall of 2023. Cohort 2 was larger owing to increasing enrollments and the inclusion of a spring course section.

3.1.3. Independent variables

Condition was the independent variable for Hypothesis 1. Condition and stimulus were the independent variables for Hypothesis 2.

The stimuli were videos of the three eeriest androids from MacDorman and Entezari (2015). Stimulus A is David Ng's Animatronic Head, making facial expressions while moving its eyes, 21 s; Stimulus B is Le Trung's Aiko, protesting when her arm is being hurt, 31 s; and Stimulus C is Hanson Robotics' Jules, anticipating the pain of missing its creator, 46 s (Fig. 5).

There were three conditions and two treatments. Condition 1 was

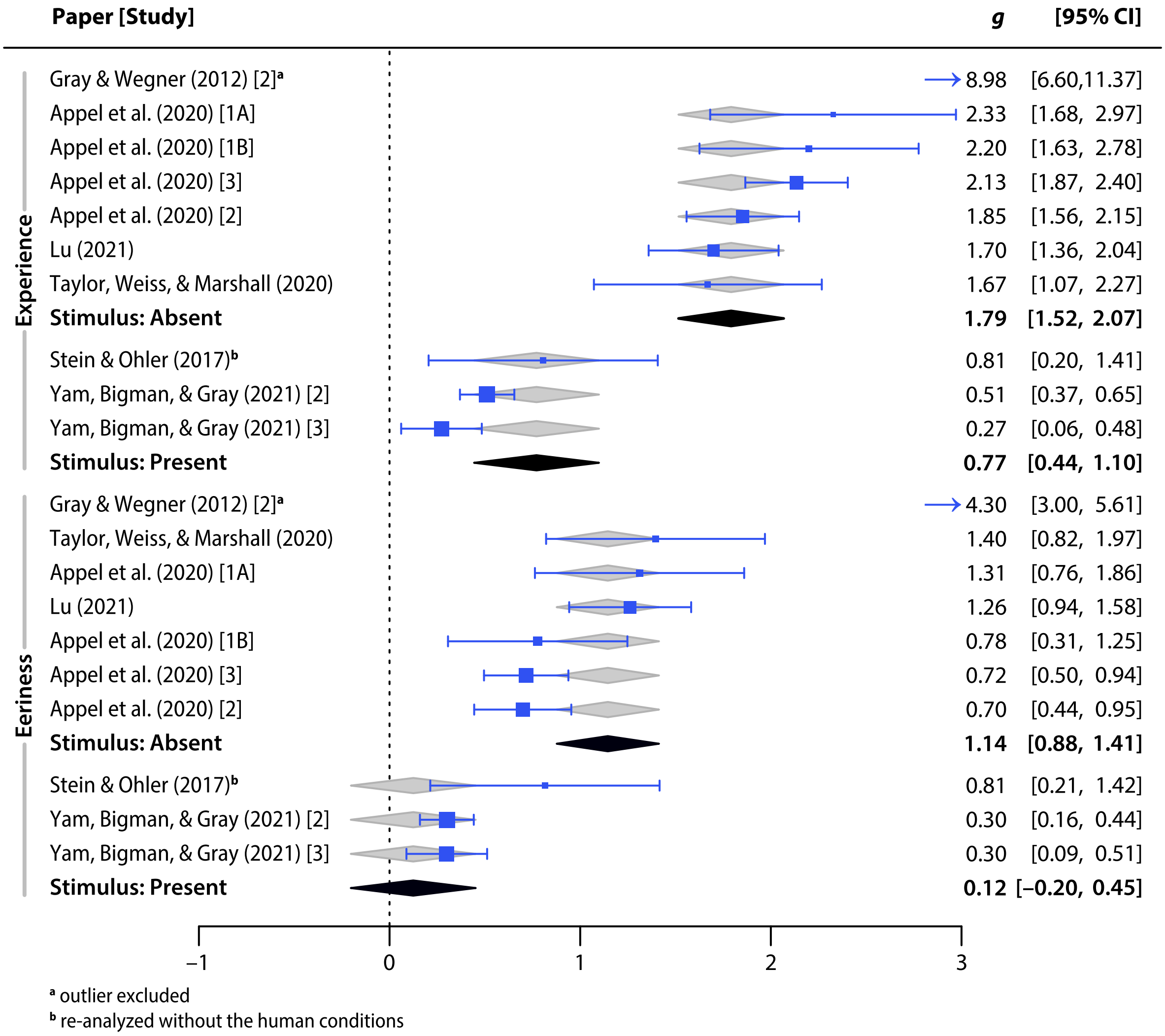


Fig. 4. The revised mixed-effects meta-regression model with construct (experience or eeriness) and stimulus (present or absent) as moderator variables. The experiment number is indicated in brackets.

before the first treatment, Condition 2 was after the first treatment but before the second, and Condition 3 was after the second treatment. Cohort 1 received the dehumanization treatment first, while Cohort 2 received the humanization treatment first.

Both treatments involved reading five or six articles, one per week, completing a 10-question multiple choice quiz, and writing a 50-to-100-word essay in response to a prompt encouraging support for the position taken in the article. The dehumanization treatment included articles advocating human uniqueness: Friedman and Kahn (1992), Searle (1990), Turkle (2007), and in abridged form, Block (1981), Lucas (1961), and Jefferson (1949). The humanization treatment included articles advocating human–robot equivalence: Chalmers (1996, pp. 253–259), Dennett (1997), and in abridged form, Calverley (2008), Newell and Simon (1976), and Putnam (1964). The readings, summarized in Appendix A, were selected to shift attitudes on human–robot similarity, both in terms of experience and agency.

3.1.4. *Dependent variables*

The indices eeriness, coldness, and humanness (5 items each) averaged 7-point semantic differential scales. To illustrate, the item humanlike–human-made from the humanness index ranged from humanlike (+3) to human-made (−3). The humanness and eeriness indices were taken from Ho and MacDorman (2017). The coldness index is Ho and MacDorman's (2010) warmth index, reverse scaled.

The indices experience (4 items), agency (4 items), and human–robot similarity (11 items) averaged 7-point Likert scales, ranging from agree (+3) to disagree (−3). For example, this item is from the experience index: "The android can feel pain." This reverse-scaled item is from the human–robot similarity index: "It is absurd to consider a human being and a robot to be the same kind of thing." The experience and agency indices were adapted from Bigman and Gray (2018). The human–robot similarity index was from MacDorman and Entezari (2015). All indices are listed in Appendix B.

3.1.5. *Data analysis*

Test statistics were two-tailed and interpreted at the $p < 0.05$ significance level. Mixed-effects models were fitted by maximum-likelihood estimation. Contrasts used type III sum of squares, and p-values for contrasts were Westfall-corrected for multiple comparisons. Effect size thresholds for η_p^2 were 0.01 for small, 0.06 for medium, and 0.14 for large.

The power analysis used the R package pwr. Descriptive statistics, reliability analyses, and correlation used psych. Hypothesis testing used nlme, performance, multcomp, and effectsize. Regression used base R. Mediation analyses used mediation. Structural equation models used lavaan.

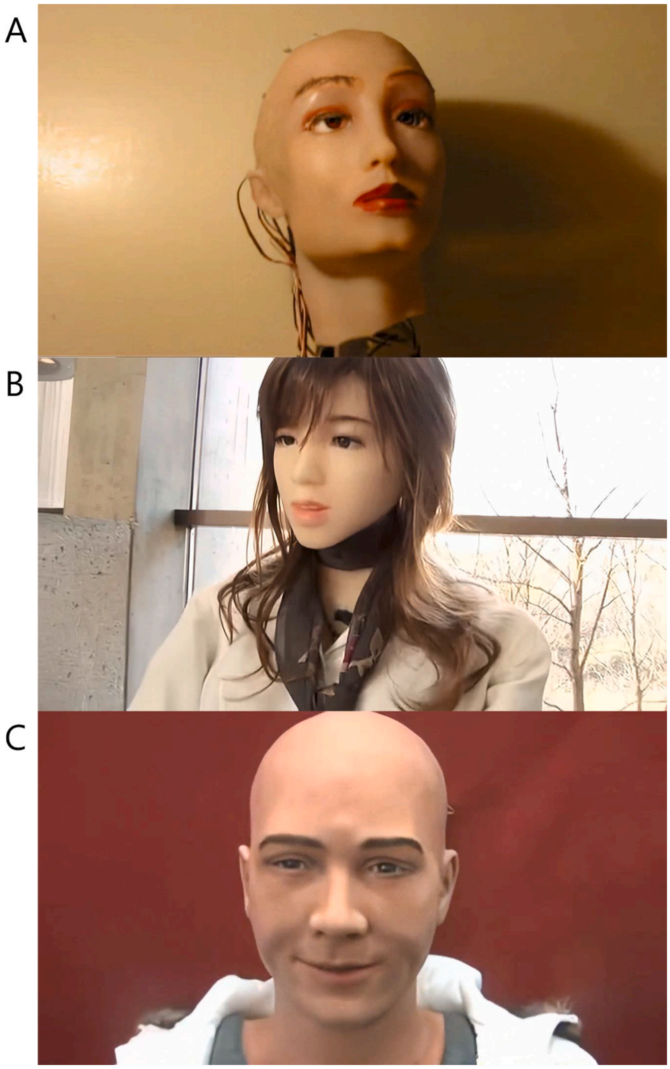


Fig. 5. Stimuli: (A) Animatronic Head, (B) Aiko, and (C) Jules.

3.2. *Results: experiment*

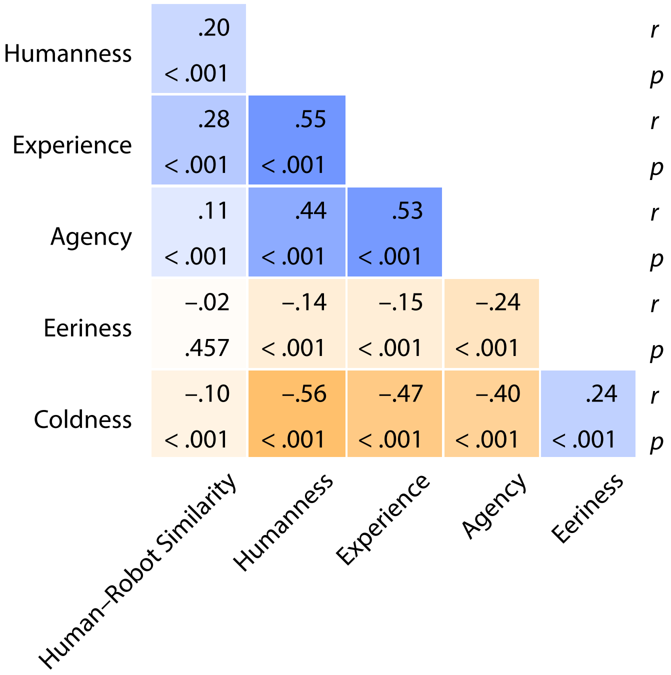
3.2.1. *Index reliability and correlation*

The psychometric properties of the dependent variables appear in Table 2. All indices were reliable (0.82 ≤ McDonald's $\omega_t \leq 0.92$). Fig. 6 lists correlations among dependent variables.

3.3. *Manipulation checks*

The dehumanization treatment was intended to shift attitudes to reduce the similarity between humans and robots and the perceived humanness of robot stimuli. The humanization treatment was intended to have the opposite effect.

For both cohorts, separate mixed-effects models with condition as the fixed factor and participant as the random factor revealed that condition had a significant effect on human–robot similarity, $p < 0.0001$, with a large effect size (Table 3). The planned contrasts revealed that the dehumanization treatment, advocating human uniqueness, decreased human–robot similarity (Cohort 1: $M_{\text{diff}} = -0.93$, $p < 0.0001$; Cohort 2: $M_{\text{diff}} = -0.38$, $p < 0.001$) and the humanization treatment, advocating human–robot equivalence, increased human–robot similarity (Cohort 1: $M_{\text{diff}} = 0.93$, $p < 0.0001$; Cohort 2: $M_{\text{diff}} = 0.59$, $p < 0.0001$). Fig. 7 visualizes these effects.



		Human–Robot Similarity	Humanness	Experience	Agency	Eeriness
Humanness	*r*	.20				
	p	< .001				
Experience	*r*	.28	.55			
	p	< .001	< .001			
Agency	*r*	.11	.44	.53		
	p	< .001	< .001	< .001		
Eeriness	*r*	−.02	−.14	−.15	−.24	
	p	.457	< .001	< .001	< .001	
Coldness	*r*	−.10	−.56	−.47	−.40	.24
	p	< .001	< .001	< .001	< .001	< .001

Fig. 6. Pairwise Pearson's correlation and Holm-adjusted *p*-value for the dependent variables.

Mixed-effects models with condition as the fixed factor and participant grouped by stimulus as the random factor revealed that condition had a significant effect on humanness, *ps* < 0.0001, and experience, *ps* ≤ 0.008 (Table 3). However, not all contrasts were significant. For both cohorts, the second treatment had a nonsignificant effect on humanness, and for Cohort 2, the humanization treatment had a nonsignificant effect on experience ($p = 0.076$). As intended, the counterbalanced treatments produced a fall-rise effect in Cohort 1 and a rise-fall effect in Cohort 2.

3.4. *Hypothesis 1*

Hypothesis 1 predicted that ascribing experience to a robot causes it to appear eerie and cold. Mixed-effects models with condition as the fixed factor and participant grouped by stimulus as the random factor revealed that condition had a significant effect on eeriness for Cohort 2, with significantly lower eeriness after the dehumanization treatment, $p = 0.0497$, with an effect size below the small threshold, $g = -0.14$ (Table 3). Condition had a significant effect on coldness for both cohorts ($p < 0.001$). However, planned contrasts revealed that the humanization

Table 2
Psychometric properties of the dependent variables.

DV	Items	Stimuli	Obs.	*M*	*SD*	ω_t	Skewness	Kurtosis
Human–Robot Similarity	11	NA	381	−0.88	1.17	0.82	0.38	−0.08
Humanness	5	3	1143	−1.37	1.45	0.88	0.65	−0.69
Experience	4	3	1143	−0.89	1.84	0.92	0.32	−1.13
Agency	4	3	1143	0.61	1.77	0.87	−0.55	−0.67
Eeriness	5	3	1143	0.37	1.38	0.84	−0.08	−0.42
Coldness	5	3	1143	−0.24	1.31	0.91	−0.23	−0.18

Table 3
One-way mixed-effects models of condition on the dependent variables with planned contrasts.

Order	DV	*RMSE*	*F*	df	*p*	η_p^2	Cont.	M_{diff}	*SE*	*z*	*p*	*g*
D–H	Human–Robot Similarity	0.51	16.45	2, 78	<0.0001	0.30	2–1	−0.93	0.19	−5.03	<0.0001	−0.76
							3–2	0.93	0.19	5.04	<0.0001	0.80
H–D	Human–Robot Similarity	0.40	14.61	2, 172	<0.0001	0.15	2–1	0.59	0.11	5.36	<0.0001	0.56
							3–2	−0.38	0.11	−3.47	<0.001	−0.35
D–H	Humanness	0.44	13.39	2, 238	<0.0001	0.10	2–1	−0.51	0.10	−5.08	<0.0001	−0.36
							3–2	0.16	0.10	1.58	0.113	0.11
H–D	Humanness	0.56	13.79	2, 520	<0.0001	0.05	2–1	0.37	0.08	4.89	<0.0001	0.26
							3–2	−0.06	0.08	−0.77	0.439	−0.04
D–H	Experience	0.94	31.54	2, 238	<0.0001	0.21	2–1	−1.13	0.15	−7.79	<0.0001	−0.65
							3–2	0.35	0.15	2.43	0.015	0.21
H–D	Experience	1.06	4.85	2, 520	0.008	0.02	2–1	0.19	0.10	1.78	0.076	0.10
							3–2	−0.32	0.10	−3.11	0.004	−0.18
D–H	Agency	1.16	27.80	2, 238	<0.0001	0.19	2–1	−1.01	0.16	−6.54	<0.0001	−0.62
							3–2	0.02	0.16	0.11	0.914	0.01
H–D	Agency	1.00	10.48	2, 520	<0.0001	0.04	2–1	−0.07	0.10	−0.66	0.508	−0.04
							3–2	−0.36	0.10	−3.60	<0.001	−0.21
D–H	Eeriness	0.52	2.42	2, 238	0.091	0.02	2–1	−0.24	0.11	−2.19	0.057	−0.16
							3–2	0.09	0.11	0.85	0.394	0.06
H–D	Eeriness	0.65	3.58	2, 520	0.028	0.01	2–1	−0.01	0.08	−0.14	0.886	−0.01
							3–2	−0.18	0.08	−2.25	0.0497	−0.14
D–H	Coldness	0.26	7.19	2, 238	<0.001	0.06	2–1	0.22	0.08	2.84	0.010	0.19
							3–2	0.06	0.08	0.79	0.432	0.05
H–D	Coldness	0.44	8.45	2, 520	<0.001	0.03	2–1	0.12	0.07	1.73	0.084	0.08
							3–2	0.16	0.07	2.38	0.036	0.11

Note: D–H indicates the dehumanization–humanization order presented to Cohort 1.
H–D indicates the humanization–dehumanization order presented to Cohort 2.

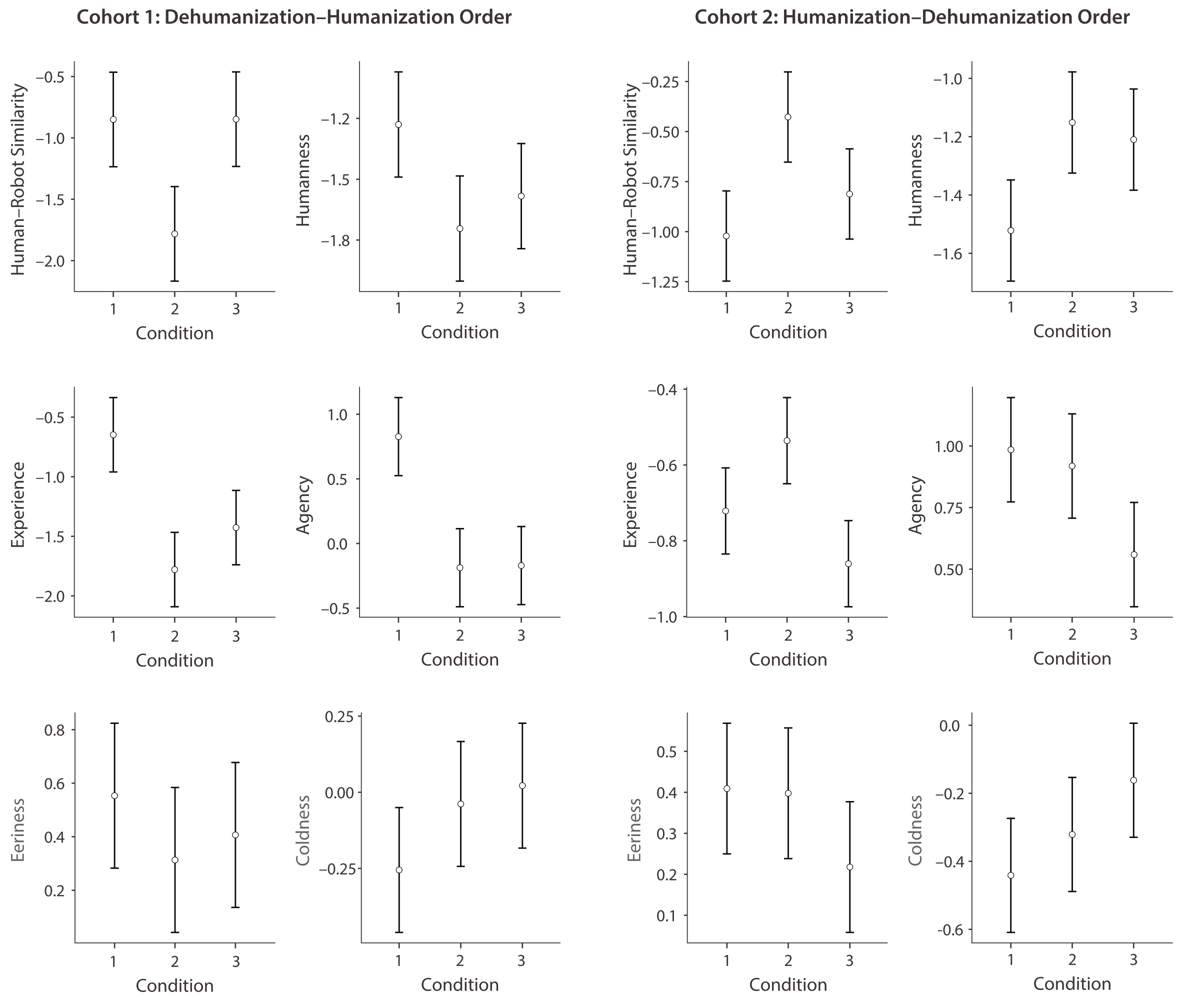


Fig. 7. Means and 95% confidence intervals of the dependent variables by condition.

treatment had a nonsignificant effect on coldness. Although dehumanization had a significant effect on coldness, the direction was counter to prediction for both cohorts. Dehumanization slightly increased coldness. Given that only three of eight contrasts were significant and two of those were counter to prediction, Hypothesis 1 was not supported.

3.5. *Hypothesis 2*

Hypothesis 2 predicted that the android robots' physical appearance has a larger effect on eeriness and coldness than ascribing experience to them. A two-way mixed effects model with condition × stimulus as the fixed factors and participant as the random factor revealed that stimulus had a larger effect on eeriness than condition (Table 4). Eeriness had a larger effect size for stimulus than for condition (Cohort 1: $\eta_p^2 = 0.26$ vs. 0.01, Cohort 2: $\eta_p^2 = 0.24$ vs. 0.01). Likewise, coldness had a larger effect size for stimulus than for condition (Cohort 1: $\eta_p^2 = 0.47$ vs. 0.03; Cohort 2: $\eta_p^2 = 0.51$ vs. 0.02). Thus, Hypothesis 2 was supported.

3.6. *Regression analyses*

As expected, viewing humans and robots as similar predicted experience and agency. A regression analysis revealed that human–robot similarity was a significant predictor of experience, $b = 0.44$, $SE = 0.04$, $\beta = 0.28$, $t(1141) = 9.88$, $p < 0.0001$, and explained 7.9% of the variance, $R^2 = 0.079$, $F(1, 1141) = 97.5$. Human–robot similarity was also a significant predictor of agency, $b = 0.17$, $SE = 0.04$, $\beta = 0.11$, $t(1141) = 3.84$, $p = 0.0001$, and explained 1.3% of the variance, $R^2 = 0.013$, $F(1, 1141) = 14.7$.

Contrary to Gray and Wegner (2012), experience was a significant negative predictor of eeriness, $b = -0.11$, SE = 0.02, $\beta = -0.15$, $t(1141) = -5.06$, $p < 0.0001$, and explained 2.2% of the variance, $R^2 = 0.022$, $F(1, 1141) = 25.6$. Agency was also a significant negative predictor of eeriness, $b = -0.19$, $SE = 0.02$, $\beta = -0.24$, $t(1141) = -8.51$, $p < 0.0001$, and explained 6% of the variance, $R^2 = 0.060$, $F(1, 1141) = 72.5$. These negative estimates may relate to the stimuli lying mainly on the human side of the uncanny valley, where increases in human traits coincide with a reduction in the uncanny valley effect.

However, if both experience and agency were predictors in the model, experience became nonsignificant, $b = -0.02$, $SE = 0.03$, $\beta = -0.02$, $t(1140) = -0.73$, $p = 0.465$ (a pattern repeated in the full structural equation model). The overall model was significant, $R^2 = 0.060$, $F(2, 1140) = 36.5$, $p < 0.0001$.

Experience was a significant negative predictor of coldness, $b = -0.33$, $SE = 0.02$, $\beta = -0.47$, $t(1141) = -17.8$, $p < 0.0001$, and explained 21.8% of the variance, $R^2 = 0.218$, $F(1, 1141) = 318$. Agency was also a significant negative predictor of coldness, $b = -0.30$, $SE = 0.02$, $\beta = -0.40$, $t(1141) = -14.72$, $p < 0.0001$, and explained 16% of the variance, $R^2 = 0.160$, $F(1, 1141) = 217$.

In the combined regression model, both experience, $b = -0.25$, $SE = 0.02$, $\beta = -0.35$, $t(1140) = -11.69$, $p < 0.0001$, and agency, $b = -0.16$, $SE = 0.02$, $\beta = -0.21$, $t(1140) = -6.93$, $p < 0.0001$, were significant negative predictors of coldness, explaining 24.9% of the variance, $R^2 = 0.249$, $F(2, 1140) = 189$.

3.7. *Mediation analyses*

A causal mediation analysis was performed using quasi-Bayesian confidence intervals with 1143 observations over 1000 simulations. The analysis revealed that experience partially mediated the effect of stimulus on eeriness: The mediation effect, $ACME = -0.05$ [–0.07, −0.03] (95% CI), $p < 0.0001$, direct effect, $ADE = 0.23$ [0.13, 0.32], $p < 0.0001$, total effect, $TE = 0.18$ [0.09, 0.28], $p < 0.0001$, and proportion of the effect mediated, $PM = -0.25$ [–0.59, −0.12], $p < 0.0001$, were all significant. The negative proportion of the effect mediated indicated that the mediator experience mitigated the stimulus's overall effect of increasing eeriness.

Agency also partially mediated the effect of stimulus on eeriness with a suppressor effect: The mediation effect, $ACME = -0.05$ [–0.09, −0.03], $p < 0.0001$, direct effect, $ADE = 0.24$ [0.15, 0.34], $p < 0.0001$, total effect, $TE = 0.19$ [0.09, 0.29], $p < 0.0001$, and proportion of the effect mediated, $PM = -0.29$ [–0.74, −0.13], $p < 0.0001$, were all significant. Agency had a suppressor effect because it reduced the direct positive relation between stimulus and eeriness, indicating that perceiving agency in the stimulus makes it less eerie than it would be otherwise.

Experience partially mediated the effect of stimulus on coldness: The mediation effect, $ACME = -0.10$ [–0.15, −0.06], $p < 0.0001$, direct effect, $ADE = -0.63$ [–0.70, −0.55], $p < 0.0001$, total effect, $TE = -0.73$ [–0.82, −0.65], $p < 0.0001$, and proportion of the effect mediated, $PM = 0.14$ [0.09, 0.19], $p < 0.0001$, were all significant. The mediator experience enhanced the total negative effect of the stimulus on coldness. In other words, ascribing experience to the stimulus reduced its coldness beyond the direct effect of the stimulus alone.

Agency partially mediated the effect of stimulus on coldness: The mediation effect, $ACME = -0.07$ [–0.10, −0.04], $p = 0.002$, direct effect, $ADE = -0.66$ [–0.74, −0.59], p < 0.0001, total effect, $TE = -0.73$ [–0.81, −0.65], $p < 0.0001$, and proportion of the effect mediated, $PM = 0.09$ [0.05, 0.14], $p = 0.002$, were all significant. Ascribing agency to the stimulus had an enhancer effect, decreasing its perceived coldness.

3.8. *Structural equation models with mediation*

A structural equation model explored how stimulus influenced perceived eeriness and coldness, mediated by experience and agency. Stimulus was the dummy-coded exogenous variable. Experience, agency, eeriness, and coldness were latent variables, and their corresponding index items were measurement variables. The video of the robot Aiko, which rated intermediate on humanness and lowest on eeriness, was selected as the reference (Table 5).

A structural equation model explored how stimulus influenced

Table 4
Two-way mixed-effects models of condition and stimulus on eeriness and coldness.

Order	IV	DV	*RMSE*	*F*	df	*p*	η_p^2
D–H	Condition	Eeriness	0.99	1.56	2, 312	0.212	0.01
D–H	Stimulus	Eeriness		54.22	2, 312	<0.0001	0.26
D–H	Condition × Stimulus	Eeriness		0.27	4, 312	0.895	0.00
H–D	Condition	Eeriness	0.98	2.79	2, 688	0.062	0.01
H–D	Stimulus	Eeriness		108.66	2, 688	<0.0001	0.24
H–D	Condition × Stimulus	Eeriness		0.71	4, 688	0.588	0.00
D–H	Condition	Coldness	0.57	4.01	2, 312	0.019	0.03
D–H	Stimulus	Coldness		137.24	2, 312	<0.0001	0.47
D–H	Condition × Stimulus	Coldness		1.55	4, 312	0.189	0.02
H–D	Condition	Coldness	0.77	5.88	2, 688	0.003	0.02
H–D	Stimulus	Coldness		360.41	2, 688	<0.0001	0.51
H–D	Condition × Stimulus	Coldness		1.98	4, 688	0.096	0.01

Table 5
Dependent variables by stimulus.

Stimulus	Humanness		Experience		Agency		Eeriness		Coldness	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Animatronic Head	−1.98	1.10	−1.66	1.44	−0.55	1.74	1.13	1.38	0.54	0.88
Aiko	−1.68	1.15	−0.86	1.80	0.93	1.42	−0.20	1.24	0.10	1.00
Jules	−0.43	1.54	−0.14	1.94	1.46	1.50	0.17	1.15	−1.36	1.20

perceived eeriness and coldness, partially mediated by experience and agency (Fig. 8). Most global fit indices indicated adequate fit: $\chi^2(158) = 1534.86$, $p < 0.0001$, $RMSEA = 0.09$ [0.08, 0.09] (90% CI), $CFI = 0.90$, $TLI = 0.88$, $SRMR = 0.10$.

Significant direct effects of Animatronic Head on eeriness, $b = 1.21$, $SE = 0.11$, $\beta = 0.45$, $z = 11.39$, $p < 0.0001$, and Jules on eeriness, $b = 0.35$, $SE = 0.09$, $\beta = 0.13$, $z = 3.75$, $p = 0.0002$, were observed, indicating videos of both robots predicted eeriness. The indirect effects of Animatronic Head and Jules on eeriness, mediated by experience, were nonsignificant ($p = 0.147$ and $p = 0.151$, respectively) and only became significant if agency was removed from the model (see simpler model below). The indirect effect of Animatronic Head on eeriness, mediated by agency, was significant, $b = 0.10$, $SE = 0.04$, $\beta = 0.04$, $z = 2.46$, $p = 0.014$, indicating an enhancer effect. Ascribing agency to the Animatronic Head significantly increased eeriness beyond its direct impact. Given that Aiko is the reference and Animatronic Head is the least humanlike and most eerie, the enhancer effect appears on the descent from Aiko into the valley. The indirect effect of Jules on eeriness, mediated by agency, was significant and negative, $b = -0.04$, $SE = 0.02$, $\beta = -0.01$, $z = -2.20$, $p = 0.028$. Ascribing agency to Jules had a mitigating effect, reducing the effect of Jules on eeriness. Although still eerier than Aiko, Jules is the most humanlike of the three robots. The total effect of Animatronic Head on eeriness was significant, $b = 1.34$, $SE = 0.10$, $\beta = 0.50$, $z = 13.59$, $p < 0.0001$, as was the total effect of Jules on eeriness, $b = 0.29$, $SE = 0.09$, $\beta = 0.11$, $z = 3.19$, $p = 0.001$.

Significant direct effects of Animatronic Head on coldness, $b = 0.25$, $SE = 0.09$, $\beta = 0.09$, $z = 2.91$, $p = 0.004$, and Jules on coldness, $b = -1.41$, $SE = 0.08$, $\beta = -0.49$, $\mathrm{z} = -16.84$, $\mathrm{p} < 0.0001$, were observed. The indirect effect of Animatronic Head on coldness, mediated by experience, was significant and positive, $b = 0.17$, $SE = 0.03$, $\beta = 0.06$, $z = 5.41$, $p < 0.0001$, indicating an enhancer effect: Ascribing experience to the Animatronic Head increased its coldness. The indirect effect of Jules on coldness was significant and negative, $b = -0.15$, $SE = 0.03$, $\beta = -0.05$, $z = -4.85$, $p < 0.0001$, indicating a mitigating effect: Ascribing experience to Jules reduced its coldness. The indirect effect of Animatronic Head on coldness, mediated by agency, was significant and positive, $b = 0.08$, $SE = 0.04$, $\beta = 0.03$, $z = 2.21.0$, $p = 0.027$, indicating an enhancer effect: Ascribing agency to the Animatronic Head increased its coldness. The indirect effect of Jules on coldness, mediated by agency, was significant and negative, $b = -0.03$, $SE = 0.01$, $\beta = -0.01$, $z = -2.02$, $p = 0.044$, indicating a mitigating effect: Ascribing agency to Jules reduced its coldness. The total effect of Animatronic Head on coldness was significant, $b = 0.50$, $SE = 0.08$, $\beta = 0.17$, $z = 6.11$, $p < 0.0001$, as was the total effect of Jules on coldness, $b = -1.59$, $SE = 0.09$, $\beta = -0.55$, $z = -18.25$, $p < 0.0001$.

Since the direct effect of experience on eeriness was nonsignificant in the full structural equation model, a simpler model was constructed to explore how stimulus influenced eeriness through experience. In this model, experience and eeriness were the only latent variables. Most global fit indices indicated an acceptable fit: $\chi^2(40) = 506.45$, $p < 0.0001$, $RMSEA = 0.10$ [0.09, 0.11] (90% CI), $CFI = 0.92$, $TLI = 0.90$, $SRMR = 0.05$.

Significant direct effects of Animatronic Head, $b = 1.31$, $SE = 0.10$, $\beta = 0.48$, $z = 13.04$, $p < 0.0001$, and Jules, $b = 0.33$, $SE = 0.09$, $\beta = 0.12$, $z = 3.59$, $p < 0.0001$, on eeriness were observed. The indirect effect of Animatronic Head on eeriness, mediated by experience, was significant, $b = 0.05$, $SE = 0.02$, $\beta = 0.02$, $z = 2.35$, $p = 0.019$, as was the indirect effect of Jules on eeriness, $b = -0.04$, $SE = 0.02$, $\beta = -0.02$, $z = -2.28$, $p = 0.022$. These partial mediations indicate experience enhanced Animatronic Head's effect on eeriness and mitigated Jules' effect on eeriness. The total effect of Animatronic Head on eeriness, mediated by experience, was significant, $b = 1.36$, $SE = 0.10$, $\beta = 0.50$, $z = 13.64$, $p < 0.0001$, as was the total effect of Jules on eeriness, $b = 0.29$, $SE = 0.09$, $\beta = 0.11$, $z = 3.21$, $p = 0.001$.

3.9. Findings

The dehumanization treatment shifted attitudes toward viewing humans and robots as less similar, and the humanization treatment shifted attitudes toward viewing them as more similar. The first treatment significantly influenced the android robots' perceived humanness: dehumanization decreased it, and humanization increased it. Dehumanization significantly reduced the android robots' experience and agency, but humanization only significantly increased their experience after dehumanization and had no significant effect on agency.

Hypothesis 1 states that ascribing experience to android robots elicits

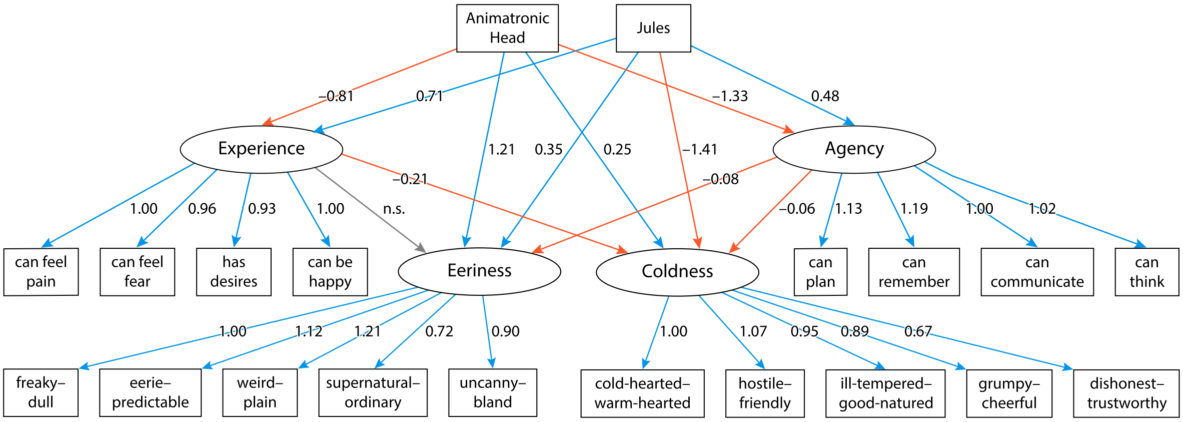


Fig. 8. A structural equation model showing the stimuli's effect of experience, agency, eeriness, and coldness with Aiko as the reference. (For readability, correlations and variances are omitted.)

feelings of eeriness and coldness. Hypothesis 2 states that the android robots' physical appearance has a larger effect on their eeriness and coldness than ascribing experience to them. Only Hypothesis 2 was supported.

The mediation analysis revealed that experience's partial mediation of eeriness appears as a mitigating effect, and agency's partial mediation appears as a suppressor effect. This is likely because two out of three of the stimuli were on the human side of the uncanny valley: The video of Animatronic Head was in the uncanny valley (eeriness = 1.13, coldness = 0.54), and Aiko (eeriness = −0.20, coldness = 0.10) and Jules (eeriness = 0.17, coldness = −1.36) were rising out of the valley (Table 5).

This pattern became more apparent in the structural equation models with Aiko as the reference. As mediators, experience and agency enhanced the effect of Animatronic Head on eeriness and coldness and mitigated the effect of Jules on eeriness and coldness. For coldness, mediation by experience was stronger than by agency. For eeriness, surprisingly, mediation by agency was stronger.

4. Data availability

The datasets, R scripts, and output of the analyses are available at https://osf.io/adn2q/.

5. Discussion

Gray and Wegner (2012) proposed "mind perception" as a theory of the uncanny valley. According to their paper, artificial entities appear eerie because their humanlike appearance prompts attributions of mind, specifically experience. Ten experiments were interpreted by their authors as supporting this theory. However, a meta-regression with moderation uncovered a troubling pattern: Experiments with no observable stimuli beyond a description generally had much larger effects for experience and eeriness than experiments with virtual reality characters, robot videos, and physical robots.

This result presents a paradox: Although attributing mind to a machine is eerie, its presence as a stimulus masks this effect. Given that the uncanny valley concerns human likeness, and an artificial entity must be perceptible to evaluate its human likeness, we sought an alternative means of testing mind perception's predictions—one using observable stimuli as the proximal cause. This required rethinking the experimental approach and how the uncanny valley effect is elicited.

A limitation with the vignette design used in these studies was identified. The vignette directly activates experience-related concepts by describing the artificial entity as having experience. This is a limitation because the vignette is the proximal cause of eeriness, not the entity's appearance.

This realization inspired the design of a novel protocol to determine whether attributions of mind could cause eeriness with only the entity's appearance as the proximal cause. In this new setup, android robots were first rated, and then attitudes regarding human–robot similarity were gradually shifted over weeks. After a one-week washout period, the robots were rated again. This methodology was designed to assess whether the robots' appearance could elicit attributions of eeriness directly, without the influence of explicit descriptions of experience, thereby providing a better-controlled test of mind perception theory's predictions.

The experimental results did not support Gray and Wegner's (2012) theory. Their theory predicts that ascribing experience to robots increases their eeriness and coldness (Hypothesis 1). In the dehumanization treatment, participants engaged with materials advocating for human uniqueness by reading articles and writing essays, significantly reducing the android robots' humanness, experience, and agency. However, contrary to the theory's predictions, robot dehumanization had a nonsignificant effect on the robots' perceived eeriness and significantly *increased* their coldness. The humanization treatment advocated human–robot equivalence. It significantly increased the android robots' experience and agency. However, robot humanization had a nonsignificant effect on the robots' eeriness and coldness. Thus, Hypothesis 1, predicted by Gray and Wegner's (2012) theory, was not supported.

An alternative view, examined below, is that automatic stimulus-driven perceptual processing elicits eeriness (MacDorman & Chattopadhyay, 2016). This view predicts that aspects of the stimulus other than experience have a larger effect on eeriness and coldness (Hypothesis 2). This hypothesis was supported.

The long-standing debate on whether affective reactions require cognition resurfaces here and elsewhere in discussions of the uncanny valley (Shin, Kim, & Biocca, 2019; Zajonc, 1980). Our study's findings favor automatic stimulus-driven perceptual processing as the primary catalyst for eeriness. This processing is fast and effortless. It operates relatively early in perception without the observer's intention or conscious control; it occurs independently of the current tasks or goals.

The recognition and interpretation of faces by the brain's specialized networks exemplifies this automatic processing. These networks have developed through extensive exposure to human faces and exhibit distinct response patterns to artificial faces, as revealed by studies using electroencephalography (EEG) and functional magnetic resonance imaging (Vaitonytė, Alimardani, & Louwerse, 2023). As a result, encountering faces with nonhuman features may trigger feedback error signals, as evidenced by changes in event-related potential (ERP) amplitudes, specifically in the LPP and N170 components (Cheetham, Wu, Pauli, & Jäncke, 2015; Schindler, Zell, Botsch, & Kissler, 2017).

Several factors may contribute to this heightened response. For instance, even though nonhuman faces are a novel category, the brain may attempt to process them configurally, as if they were real human faces (Diel & MacDorman, 2021). Then, slight deviations from facial norms in these nonhuman faces are amplified by perceptual narrowing (Chattopadhyay & MacDorman, 2016; Moore, 2012). Perceptual narrowing improves our ability to distinguish among familiar stimuli but also makes us more sensitive to anomalies in them (Maurer & Werker, 2013). The P200 findings in ERP studies indicate that altered or less familiar faces elicit stronger responses (Mustafa & Magnor, 2016). Moreover, inconsistencies arise when the brain integrates facial features into a whole, especially if some features are processed more efficiently than others. This results in lags and timing errors (MacDorman & Chattopadhyay, 2016). These discrepancies heighten unease with faces that appear less than fully human (Saygin, Chaminade, Ishiguro, Driver, & Frith, 2012).

Attributions differ from perceptual processing in that they are more deliberate, slower, occur later, and involve a broader range of cognitive processing. To attribute mind to a machine may require us to make inferences from its behavior and the situation to draw a reasoned conclusion. Brain imaging studies have implicated both early and later processing in the uncanny valley effect. The fusiform gyrus, ventromedial prefrontal cortex (vmPFC), and temporoparietal junction (TPJ) show differential activations associated with perceiving human likeness in nonhuman agents (Cheetham, Suter, & Jäncke, 2011; Rosenthal-von der Pütten, Krämer, Maderwald, Brand, & Grabenhorst, 2019; Wang & Quadflieg, 2015). Although the fusiform gyrus is involved early in identifying face features, the vmPFC and TPJ engage later in emotional evaluation and social reasoning.

Although an explanation of the uncanny valley that relies on attributions of mind ties the effect to these slower, more deliberate processes, such an explanation may not be necessary. For example, exposures to humanlike robots as brief as 50 ms have elicited an uncanny valley effect comparable to exposures lasting several seconds (Yam, Gong, & Xu, 2024). This observation shows how the uncanny valley effect could often stem from more immediate and less consciously controlled perceptual dynamics—rather than the slower, more deliberate cognitive evaluations involved in attributions of mind.

A thought experiment may illustrate this distinction: Imagine a girl named Hazel is planning a road trip with her mother. She introduces her mother to a long-time friend, hoping to bring the friend along. Soon, her mother agrees, and they set off together. At a roadside hotel, Hazel is awakened by her cell phone. Distraught, her mother explains that she and Hazel's friend got up early to do some shopping, but along the way, her friend was struck by a school bus. The collision revealed that Hazel's friend, whom she believed to be human, was an android. This news left Hazel with an eerie sensation. The school children on the bus were unnerved for a different reason: They had just witnessed humanlike and mechanical parts violently scattered across the road. Cutting the trip short, Hazel and her mother collected the android's parts and returned home. At first, in nightmares, but later in dreams, Hazel was visited by her friend. She eventually accepted that the bonds of friendship could extend beyond the human.

This thought experiment is meant to delineate the uncanniness of perception experienced by the school children (Mori, 2012) and the uncanniness of mind attribution experienced by Hazel. The android's shattered human likeness on the road elicited the former; implications surrounding a friend's nonhuman identity resulted in the latter—even in the absence of the android (Cha et al., 2020; MacDorman, Vasudevan, & Ho, 2009; Złotowski, Yogeeswaran, & Bartneck, 2017).

This distinction between perception and attribution leads us to reconsider the uncanny valley effect in light of mind perception theory. Mind perception predicts that, as a robot becomes more humanlike, its appearance elicits more attributions of experience, causing uncanniness. A conceptual flaw in this characterization is that, with increasing human likeness, this effect does not produce a valley. It produces a downward slope up until the point at which the robot becomes indistinguishable from a human being.

Empirically, focusing solely on the *descent* into the uncanny valley leads to a false conclusion. Gray and Wegner (2012) found that experience partially mediated the stimulus's effect on eeriness when comparing Kaspar's machine-like back view with the robot's humanlike front view in 12-s videos (Experiment 1). However, they used the Sobel test, which does not indicate the type of mediation. Assuming experience had an enhancer effect in their experiment, in our present experiment, experience had a mitigating effect coming out of the valley on the human side. Moreover, in the full structural equation model, only the mitigating effect of agency reached significance, suggesting a critical role for agency when directly perceiving robots rather than just imagining them in a vignette.

Our interpretation of these results contrasts with Gray and Wegner's (2012): Experience and agency partially mediate eeriness, but the reference determines whether they have an enhancer or mitigating effect. For the descent into the valley, experience and agency are enhancers, and for the ascent from the valley, they are mitigators. Based on this perspective, as the robot's human likeness increases, experience and agency enhance eeriness up to the uncanny valley, but beyond the uncanny valley, these factors mitigate eeriness. Coldness functions differently: It monotonically decreases as experience and agency increase (Fig. 8). This is unsurprising, as experience is closely related to the warmth–coldness dimension (Fiske et al., 2007, cited by Gray & Wegner, 2012).

However, the stimuli in our experiment are limited because there are only three robots, and they are not varied systematically. Therefore, the evidence is insufficient to generalize whether experience and agency increase eeriness when descending into the uncanny valley or have a mitigating effect when ascending from it. Our experiment needs to be repeated with stimuli that are systematically varied in human likeness (as in Chattopadhyay & MacDorman, 2016) or numerous (as in Kim et al., 2022 and Mathur et al., 2020).

Bifurcation is a straightforward way to resolve the issues with mind perception theory. First, mind perception can be developed as a theory of disembodied AI without linking it to Mori's uncanny valley (Stein & MacDorman, 2024). Second, to explain Mori's uncanny valley, mind perception must be situated within a broader theory.

One of the simplest and earliest is that an android robot is uncanny because it elicits a model of a human being but violates some of the model's predictions regarding human norms (MacDorman & Ishiguro, 2006). Urgen, Kutas, and Saygin (2018) found empirical support for this theory. When a mechanical robot, android, and human performed the same movement, N400 amplitudes were highest when observing the android, likely owing to the incongruence of its human appearance paired with a nonbiological movement. N400 amplitudes were also larger for photorealistic computer-generated characters than less realistic characters or real humans, indicating that these agents violate predictions of anticipated human movement (Mustafa, Guthe, Tauscher, Goesele, & Magnor, 2017).

In our experiment, the Animatronic Head's human appearance causes us to expect human facial expressions, which involve muscles pulling skin across bones. Instead, we see what looks more like bones moving under skin, a violation of human norms. Experience and agency enhance this effect because this kind of norm violation in an entity that is conscious and cognizant is horrifying. In contrast, Jules having a warm conversation is much less disturbing, despite the wires coming out of its head. So, for Jules, experience and agency mitigate eeriness by increasing our warm feelings toward the robot.

In conclusion, contrary to Fig. 2, the depths of the uncanny may not be occupied by a humanlike entity that we know to be a machine. The depths may instead be occupied by a figure with a mix of features—human and nonhuman, living and inanimate—perhaps akin to the corpse's position in Mori's graph (Fig. 1). When experience and agency awaken in it, the corpse becomes something even more disturbing: a zombie. These examples could indicate how mind perception contributes to the uncanny valley effect.

CRediT authorship contribution statement

Karl F. MacDorman: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Summary of readings

Dehumanization treatment. These readings advocate human uniqueness. Their philosophical positions include anti-behaviorism (Block, 1981), anti-reductionism (Friedman & Kahn, 1992), dualism (Jefferson, 1949), anti-computationalism (Lucas, 1961), biological naturalism (Searle, 1990), and technological skepticism and humanism (Turkle, 2007). Block (1981) and Searle (1990) assume agency in computational systems and argue against mind or experience. Lucas argues against agency, examining what a formal system cannot do. Friedman and Kahn (1992), Jefferson (1949), and Turkle (2007) focus on both experience and agency and the risks to people in mistakenly ascribing these qualities to machines.

Block (1981) shows that passing the Turing test does not indicate intelligence in a computer because mechanisms as simple as a giant lookup table or as unrelated to cognition as a particle simulator could also pass it. Friedman and Kahn (1992) argue that present-day computers cannot be moral agents because they lack intentionality, which causally affects human actions. Jefferson (1949) contends that machines lack genuine thoughts, emotions, and human creativity. Lucas (1961) reasons that human insight and free will transcend rule-governed machines because truths can be self-evident but unprovable within a formal system, as Kurt Gödel showed. Searle (1990) argues that if a human cannot understand Chinese by manipulating Chinese symbols according to rules, neither can a digital computer. Turkle (2007) observes that despite their engaging design, relational artifacts lack understanding and empathy, leading to deceptive, inauthentic relationships.

Humanization treatment. These readings advocate human–robot equivalence. They broadly align with the functionalist position in the philosophy of mind. This position asserts that mental states are determined by their function in a cognitive system, not their internal constitution. Consequently, mental states can, in principle, be realized by media other than biological brains, such as silicon chips. Since functionalism views experience and agency as inseparable from cognitive function, these readings advocate for equivalence on both dimensions, though with varying emphasis.

Calverley (2008) argues that robots exhibiting the capacity for intentional action and autonomy can be considered legal persons. Chalmers (1996, pp. 253–259) shows in a proof by contradiction why gradually replacing a person's neurons with functionally equivalent silicon chips should not cause subjective experience to fade. Dennett (1997) posits that robots have the potential to become conscious because humans are conscious, and both humans and robots are complex physical mechanisms governed by the same laws of nature. Newell and Simon (1976) assert that both human minds and intelligent programs are examples of physical symbol systems, which are necessary and sufficient for general intelligent action. Putnam (1964) argues that to avoid discriminatory moral consequences, we should presume sophisticated androids have subjective experiences just as we presume humans do.

Appendix B. Reliability analysis

An omega total reliability analysis was conducted for the (de)humanization experiment's indices. Tables B1–B6 list the indices, their items, factor loadings (λ), commonalities (h^2), item–rest correlation (ρ), and Cronbach's $\boldsymbol{\alpha}$ if dropped. Reliability and model fit measures are also provided. Human–robot similarity is from MacDorman and Entezari (2015), experience and agency are from Bigman and Gray (2018), humanness and eeriness are from Ho and MacDorman (2017), coldness is the warmth index, reverse scaled, from Ho and MacDorman (2010).

Table B1
Human–Robot Similarity Index, Reverse Scaled

Item	λ	h^2	$\rho_{\text{i–r}}$	$\boldsymbol{\alpha}_{\text{drop}}$
2. Even if a robot might one day seem human, it would never be anything like a real human being. [r]	0.70	0.50	0.70	0.79
1. It is absurd to consider a human being and a robot to be the same kind of thing. [r]	0.66	0.44	0.68	0.79
4. Human beings are fundamentally different from robots. [r]	0.64	0.41	0.65	0.79
3. Someday, robots will be able to feel pain and heartache just like human beings do.	−0.61	0.37	0.66	0.79
6. Human beings have a soul, which a robot could never have. [r]	0.54	0.29	0.57	0.80
5. Reproduce human brain processes in a robot, and the robot would be conscious.	−0.52	0.27	0.61	0.80
7. Since only human beings are created in God's image, no robot could ever be.	0.50	0.25	0.54	0.81
8. In a sense, human beings are nothing more than highly sophisticated, self-replicating robots.	−0.49	0.24	0.58	0.80
11. The internal workings of human beings and robots are governed by the same physical processes.	−0.42	0.18	0.51	0.81
10. It would be satisfactory if someday we could not tell robots from human beings.	−0.41	0.17	0.51	0.81
9. When taking on human occupations, robots also take on moral responsibility for their actions.	−0.41	0.16	0.52	0.81

[r] reverse scaled.

$n = 127$, $N = 381$, $\omega_t = 0.82$, $\boldsymbol{\alpha} = 0.81$, % Var. = 0.30, $RMSEA = 0.13$, $TLI = 0.68$, $fit = 0.70$.

Table B2
Humanness Index

Item	λ	h^2	$\rho_{\text{i–r}}$	$\boldsymbol{\alpha}_{\text{drop}}$
Real–Synthetic	0.87	0.76	0.88	0.83
Humanlike–Human-made	0.84	0.71	0.87	0.83
Biological Movement–Mechanical Movement	0.82	0.68	0.85	0.84
Living–Inanimate	0.79	0.63	0.83	0.85
Mortal–Without Definite Lifespan	0.52	0.27	0.67	0.90

$n = 127$, $N = 1143$, $\omega_t = 0.88$, $\boldsymbol{\alpha} = 0.88$, % Var. = 0.61, $RMSEA = 0.11$, $TLI = 0.96$, $fit = 0.93$.

Table B3
Experience Index

Item	λ	h^2	$\rho_{\text{i–r}}$	$\boldsymbol{\alpha}_{\text{drop}}$
… can feel fear.	0.90	0.80	0.91	0.88
… can be happy.	0.88	0.77	0.91	0.89
… can feel pain.	0.85	0.72	0.89	0.90
… has desires.	0.81	0.66	0.87	0.90

$n = 127$, $N = 1143$, $\omega_t = 0.92$, $\boldsymbol{\alpha} = 0.92$, % Var. = 0.74, $RMSEA = 0.29$, $TLI = 0.84$, $fit = 0.97$.

Table B4
Agency Index

Item	λ	h^2	$\rho_{\text{i–r}}$	α_{drop}
… can remember things.	0.86	0.74	0.88	0.81
… can plan actions.	0.84	0.70	0.87	0.81
… can think.	0.73	0.53	0.82	0.85
… can communicate with others.	0.72	0.52	0.81	0.85

$n = 127$, $N = 1143$, $\omega_t = 0.87$, $\alpha = 0.87$, % Var. = 0.62, $RMSEA = 0.08$, $TLI = 0.98$, $fit = 0.93$.

Table B5
Eeriness Index

Item	λ	h^2	$\rho_{\text{i–r}}$	α_{drop}
Weird–Plain	0.82	0.67	0.84	0.78
Eerie–Predictable	0.77	0.59	0.81	0.79
Freaky–Dull	0.72	0.52	0.78	0.80
Uncanny–Bland	0.67	0.45	0.76	0.81
Supernatural–Ordinary	0.58	0.33	0.69	0.83

$n = 127$, $N = 1143$, $\omega_t = 0.84$, $\alpha = 0.84$, % Var. = 0.51, $RMSEA = 0.09$, $TLI = 0.95$, $fit = 0.87$.

Table B6
Coldness Index

Item	λ	h^2	$\rho_{\text{i–r}}$	α_{drop}
Friendly–Hostile	−0.89	0.80	0.91	0.88
Good-natured–Ill-tempered	−0.89	0.80	0.90	0.88
Cheerful–Grumpy	−0.81	0.65	0.85	0.89
Warm-hearted–Cold-hearted	−0.81	0.65	0.86	0.90
Trustworthy–Dishonest	−0.71	0.51	0.78	0.91

$n = 127$, $N = 1143$, $\omega_t = 0.91$, $\alpha = 0.91$, % Var. = 0.68, $RMSEA = 0.07$, $TLI = 0.99$, $fit = 0.96$.

References

Appel, M., Izydorczyk, D., Weber, S., Mara, M., & Lischetzke, T. (2020). The uncanny of mind in a machine: Humanoid robots as tools, agents, and experiencers. *Computers in Human Behavior, 102*, 274–286. https://doi.org/10.1016/j.chb.2019.07.031

Bigman, Y., & Gray, K. (2018). People are averse to machines making moral decisions. *Cognition, 181*, 21–34. https://doi.org/10.1016/j.cognition.2018.08.003

Block, N. (1981). Psychologism and behaviourism. *Philosophical Review, 90*(1), 5–43. https://doi.org/10.2307/2184371

Calverley, D. J. (2008). Imagining a nonbiological machine as a legal person. *AI & Society, 22*, 523–537. https://doi.org/10.1007/s00146-007-0092-7

Cha, Y., Baek, S., Ahn, G. S., Lee, H., Lee, B., Shin, J., et al. (2020). Compensating for the loss of human distinctiveness: The use of social creativity under human–machine comparisons. *Computers in Human Behavior, 103*, 80–90. https://doi.org/10.1016/j.chb.2019.08.027

Chalmers, D. J. (1996). *The conscious mind (“Fading qualia,” pp. 253-259)*. Oxford, UK: Oxford University Press.

Chattopadhyay, D., & MacDorman, K. F. (2016). Familiar faces rendered strange: Why inconsistent realism drives characters into the uncanny valley. *Journal of Vision, 16* (11), 1–25. https://doi.org/10.1167/16.11.7, 7.

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*, 1–14. https://doi.org/10.3389/fnhum.2011.00126. Article 126.

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), 1–15. https://doi.org/10.3389/fpsyg.2015.00981

Dennett, D. C. (1997). Consciousness in human and robot minds. In M. Ito, Y. Miyashita, & E. T. Rolls (Eds.), *Cognition, computation, and consciousness* (pp. 17–29). New York: Oxford University Press. https://doi.org/10.1037/10247-002.

Diel, A., & MacDorman, K. F. (2021). Creepy cats and strange high houses: Support for configural processing in testing predictions of nine uncanny valley theories. *Journal of Vision, 21*(4), 1–20. https://doi.org/10.1167/jov.21.4.1

Diel, A., Weigelt, S., & MacDorman, K. F. (2022). A meta-analysis of the uncanny valley’s independent and dependent variables. *ACM Transactions on Human–Robot Interaction, 11*(1). https://doi.org/10.1145/3470742

Fiske, S. T., Cuddy, A. J., & Glick, P. (2007). Universal dimensions of social cognition: Warmth and competence. *Trends in Cognitive Sciences, 11*, 77–83. https://doi.org/10.1016/j.tics.2006.11.005

Fiske, S. T., Cuddy, A. J., Glick, P., & Xu, J. (2002). A model of (often mixed) stereotype content: Competence and warmth respectively follow from perceived status and competition. *Journal of Personality and Social Psychology, 82*, 878–902. https://doi.org/10.1037/0022-3514.82.6.878

Friedman, B., & Kahn, P. H., Jr. (1992). Human agency and responsible computing: Implications for computer system design. *Journal of Systems and Software, 17*(1), 7–14. https://doi.org/10.1016/0164-1212(92)90075-U

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

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*(6), 1508–1518. https://doi.org/10.1016/j.chb.2010.05.015

Ho, C.-C., & MacDorman, K. F. (2017). Measuring the uncanny valley effect: Refinements to indices for perceived humanness, attractiveness, and eeriness. *International Journal of Social Robotics, 9*(1), 129–139. https://doi.org/10.1007/s12369-016-0380-9

Jefferson, G. (1949). The mind of mechanical man. *British Medical Journal, 1*(4616), 1105–1110.

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. https://doi.org/10.3389/fpsyg.2015.00390

Kim, B., de Visser, E., & Phillips, E. (2022). Two uncanny valleys: Re-evaluating the uncanny valley across the full spectrum of real-world human-like robots. *Computers in Human Behavior, 135*, Article 107340. https://doi.org/10.1016/j.chb.2022.107340

Lakens, D. (2013). Calculating and reporting effect sizes to facilitate cumulative science: A practical primer for *t*-tests and ANOVAs. *Frontiers in Psychology, 4*(863). https://doi.org/10.3389/fpsyg.2013.00863

Lu, E. M. (2021). *Behind the uncanny valley of mind: Investigating the effects of agency and experience in chatbot interactions (Identity No. 1282271)*. Master’s thesis. Eindhoven University of Technology. TU/e Repository https://research.tue.nl/en/studentTheses/behind-the-uncanny-valley-of-mind.

Lucas, J. R. (1961). Minds, machines and Gödel. *Philosophy, 36*, 112–127. https://doi.org/10.1017/S0031819100057983

MacDorman, K. F., & Chattopadhyay, D. (2016). Reducing consistency in human realism increases the uncanny valley effect; increasing category uncertainty does not. *Cognition, 146*, 190–205. https://doi.org/10.1016/j.cognition.2015.09.019

MacDorman, K. F., & Entezari, S. (2015). Individual differences predict sensitivity to the uncanny valley. *Interaction Studies, 16*(2), 141–172. https://doi.org/10.1075/is.16.2.01mac

MacDorman, K. F., & Ishiguro, H. (2006). The uncanny advantage of using androids in social and cognitive science research. *Interaction Studies, 7*(3), 297–337. https://doi.org/10.1075/is.7.3.03mac

MacDorman, K. F., Vasudevan, S. K., & Ho, C.-C. (2009). Does Japan really have robot mania? Comparing attitudes by implicit and explicit measures. *AI & Society, 23*, 485–510. https://doi.org/10.1007/s00146-008-0181-2

Mathur, M. B., Reichling, D., Lunardini, F., Geminiani, A., Antonietti, A., Ruijten, P., et al. (2020). Uncanny but not confusing: Multisite study of perceptual category confusion in the uncanny valley. *Computers in Human Behavior, 103*, 21–30. https://doi.org/10.1016/j.chb.2019.08.029

Maurer, D., & Werker, J. F. (2013). Perceptual narrowing during infancy: A comparison of language and faces. *Developmental Psychobiology, 56*(2), 154–178. https://doi.org/10.1002/dev.21177

Moore, R. K. (2012). A Bayesian explanation of the 'uncanny valley' effect and related psychological phenomena. *Scientific Reports, 2*(864), 1–5. https://doi.org/10.1038/srep00864

Mori, M. (2012). The uncanny valley (K. F. MacDorman & N. Kageki, trans.). *IEEE Robotics and Automation, 19*(2), 98–100. https://doi.org/10.1109/MRA.2012.2192811

Mustafa, M., Guthe, S., Tauscher, J. P., Goesele, M., & Magnor, M. (2017). How human am I? EEG-based evaluation of virtual characters. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (pp. 5098–5108). https://doi.org/10.1145/3025453.3026043

Mustafa, M., & Magnor, M. (2016). EEG based analysis of the perception of computer-generated faces. In *Proceedings of the 13th European Conference on Visual Media Production* (pp. 1–10). https://doi.org/10.1145/2998559.2998563 (Article No. 4).

Newell, A., & Simon, H. A. (1976). Computer science as empirical inquiry: Symbols and search. *Communications of the ACM, 19*(3), 113–126. https://doi.org/10.1145/360018.360022

Putnam, H. (1964). Robots: Machines or artificially created life? *Journal of Philosophy, 61* (21), 668–691.

Rosenthal-von der Pütten, A. M., Krämer, N. C., Maderwald, S., Brand, M., & Grabenhorst, F. (2019). Neural mechanisms for accepting and rejecting artificial social partners in the uncanny valley. *Journal of Neuroscience, 39*(33), 6555–6570. https://doi.org/10.1523/JNEUROSCI.2956-18.2019

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*(4), 413–422. https://doi.org/10.1093/scan/nsr025

Schindler, S., Zell, E., Botsch, M., & Kissler, J. (2017). Differential effects of face-realism and emotion on event-related brain potentials and their implications for the uncanny valley theory. *Scientific Reports, 7*(45003), 1–13. https://doi.org/10.1038/srep45003

Searle, J. (1990). Is the brain's mind a computer program? *Scientific American, 262*(1), 20–25. https://doi.org/10.1038/scientificamerican0190-26

Shin, M., Kim, S.-J., & Biocca, F. (2019). The uncanny valley: No need for any further judgments when an avatar looks eerie. *Computers in Human Behavior, 94*, 100–109. https://doi.org/10.1016/j.chb.2019.01.016

Stein, J.-P., & MacDorman, K. F. (2024). After confronting one uncanny valley, another awaits. *Nature Reviews Electrical Engineering.* https://doi.org/10.1038/s44287-024-00041-w

Stein, J.-P., & Ohler, P. (2017). Venturing into the uncanny valley of mind—the influence of mind attribution on the acceptance of human-like characters in a virtual reality setting. *Cognition, 160*, 43–50. https://doi.org/10.1016/j.cognition.2016.12.010

Taylor, J., Weiss, S. M., & Marshall, P. J. (2020). "Alexa, how are you feeling today?" Mind perception, smart speakers, and uncanniness. *Interaction Studies, 21*(3), 329–352. https://doi.org/10.1075/is.19015.tay

Turkle, S. (2007). Authenticity in the age of digital companions. *Interaction Studies, 8*(3), 501–517. https://doi.org/10.1075/is.8.3.11tur

Urgen, B. A., Kutas, M., & Saygin, A. P. (2018). Uncanny valley as a window into predictive processing in the social brain. *Neuropsychologia, 114*, 181–185. https://doi.org/10.1016/j.neuropsychologia.2018.04.027

Vaitonytė, J., Alimardani, M., & Louwerse, M. M. (2023). Scoping review of the neural evidence on the uncanny valley. *Computers in Human Behavior Reports, 9*, Article 100263. https://doi.org/10.1016/j.chbr.2022.100263

Viechtbauer, W., & Cheung, M. W.-L. (2010). Outlier and influence diagnostics for meta-analysis. *Research Synthesis Methods, 1*(2), 112–125. https://doi.org/10.1002/jrsm.11

Wang, S., Lilienfeld, S. O., & Rochat, P. (2015). The uncanny valley: Existence and explanations. *Review of General Psychology, 19*(4), 393–407. https://doi.org/10.1037/gpr0000056

Wang, Y., & Quadflieg, S. (2015). In our own image? Emotional and neural processing differences when observing human–human vs human–robot interactions. *Social Cognitive and Affective Neuroscience, 10*(11), 1515–1524. https://doi.org/10.1093/scan/nsv043

Yam, K. C., Bigman, Y., & Gray, K. (2021). Reducing the uncanny valley by dehumanizing humanoid robots. *Computers in Human Behavior, 125*, Article 106945. https://doi.org/10.1016/j.chb.2021.106945

Yam, J., Gong, T., & Xu, H. (2024). A stimulus exposure of 50 ms elicits the uncanny valley effect. *Heliyon, 10*(6), Article e27977. https://doi.org/10.1016/j.heliyon.2024.e27977

Zajonc, R. B. (1980). Feeling and thinking: Preferences need no inferences. *American Psychologist, 35*(2), 151–175. https://doi.org/10.1037/0003-066X.35.2.151

Złotowski, J., Yogeeswaran, K., & Bartneck, C. (2017). Can we control it? Autonomous robots threaten human identity, uniqueness, safety, and resources. *International Journal of Human-Computer Studies, 100*, 48–54. https://doi.org/10.1016/j.ijhcs.2016.12.008