International Journal of Human–Computer Interaction

ISSN: 1044-7318 (Print) 1532-7590 (Online) Journal homepage: www.tandfonline.com/journals/hihc20

Electrophysiological Correlates of the Uncanny Valley

Motonori Yamaguchi, Maemi J. Bautista & Ian Daly

To cite this article: Motonori Yamaguchi, Maemi J. Bautista & Ian Daly (30 Jan 2026): Electrophysiological Correlates of the Uncanny Valley, International Journal of Human–Computer Interaction, DOI: 10.1080/10447318.2026.2621278

To link to this article: https://doi.org/10.1080/10447318.2026.2621278

INTERNATIONAL JOURNAL OF HUMAN–COMPUTER INTERACTION
https://doi.org/10.1080/10447318.2026.2621278

Electrophysiological Correlates of the Uncanny Valley

Motonori Yamaguchi[a], Maemi J. Bautista[a] and Ian Daly[b]

[a]Centre for Brain Science, Department of Psychology, University of Essex, Colchester, UK; [b]Brain-Computer Interfaces and Neural Engineering Laboratory, University of Essex, Colchester, UK

ABSTRACT

The present study investigated the electrophysiological correlates of the uncanny valley effect while human observers viewed pictures of robot faces that varied in the human-likeness of their appearance. We found characteristic non-linear trajectories of the N400 amplitude against the self-rated human-likeness score, similar to the trajectory found for the self-rated likability score. The late positive potential (LPP) and frontal-midline theta showed parabolic trajectories against the human-likeness scores, which peaked at most uncanny faces but did not follow the non-linear uncanny-valley pattern. The frontal alpha asymmetry and the N170 did not produce this trajectory. These results corroborated the proposal that the uncanny valley effect stems from expectancy violation for human faces, but they surprisingly provided little evidence for affective reactions to uncanny robot faces. The LPP and frontal-midline theta could reflect increased cognitive control to process categorically ambiguous faces rather than the uncanniness of the faces. These findings suggest that the uncanny valley effect is primarily a perceptual or cognitive phenomenon, whereas its emotional impact might be less prominent than commonly believed.

KEYWORDS

Human–robot interaction; social robotics; electrophysiology; affect; valence

1. Introduction

When we encounter images of artificially created agents, such as virtual avatars and robots, their human-like appearance can trigger a deep sense of unease. This phenomenon was first documented by Mori (1970) and was introduced in English as the *uncanny valley* (Reichardt, 1978; also see Mori et al., 2012). Mori observed that people's affinity for objects increases as their appearance becomes more human-like until a certain threshold, beyond which their affinity drops sharply (see Figure 1). Despite the broad acceptance of the uncanny valley hypothesis, however, much remains unknown about how we respond to images that fall with the uncanny valley (for reviews, see, Kätsyri et al., 2015; Wang et al., 2015). Whereas most studies used subjective ratings of robots or digital images, there remains much to understand about neural responses to images at different locations along the uncanny valley (Vaitonytė et al., 2023). The present study used electroencephalography (EEG) to monitor brain activities while people observed images of robots that varied the degree of uncanniness. We evaluated multiple potential neural correlates of the uncanny valley.

1.1. Electrophysiological responses to the uncanny valley

Although the concept of the uncanny valley was proposed more than a half century ago, empirical studies of the phenomenon are a relatively recent development (Diel et al., 2022; Kätsyri, et al., 2015; Vaitonytė et al., 2023; Wang et al., 2015), and many accounts have been proposed. For instance, some

CONTACT Motonori Yamaguchi motonori.yamaguchi@essex.ac.uk, cog.yamaguchi@gmail.com Centre for Brain Science, Department of Psychology, University of Essex, Wivenhoe Park, Colchester, CO4 3SQ, UK

Supplemental data for this article can be accessed online at https://doi.org/10.1080/10447318.2026.2621278.

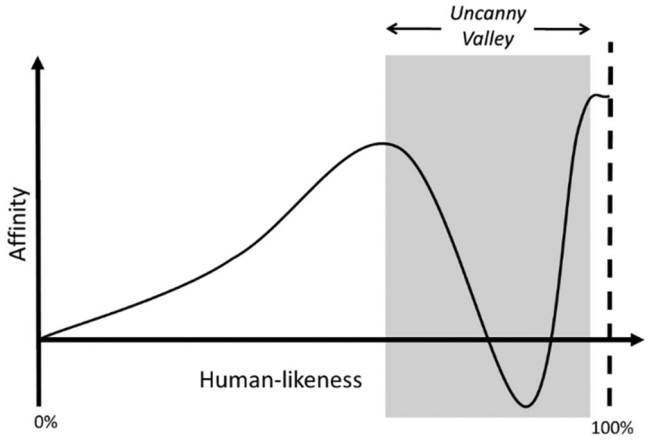

Figure 1. An illustration of the uncanny valley hypothesis by Mori (1970).

argue that the uncanniness of faces results from facial distortions that are indicative of a transmissible disease, provoke a fear of death, or fail aesthetic standards for mate selection (MacDorman & Ishiguro, 2006). Other theories suggest that near-human appearance evokes the expectancy of human features or characteristics, which are then violated by non-human features or a lack of human characteristics (e.g., MacDorman & Chattopadhyay, 2016; Moore, 2012). Still others argue that categorical confusion or atypicality of exemplars creates the feeling of uncanniness (e.g., Burleigh et al., 2013; Yamada et al., 2013) or that "seeing the mind" in robots (e.g., the capacity to feel pain) causes the uncanny feeling (Gray & Wegner, 2012). Some of these accounts have been tested empirically (Gray & Wegner, 2012; MacDorman, 2024; Mathur et al., 2020; Yamada et al., 2013), but other accounts have not been submitted to a rigorous evaluation. Arguably, accounts that have received more support are those that emphasize the presence of atypical (e.g., MacDorman & Chattopadhyay, 2016; Seyama & Nagayama, 2007; Tinwell et al., 2013) or distorted (Diel & MacDorman, 2021) facial features (e.g., having unusually large eyes), but some argue that human-likeness is a multidimensional construct (Burleigh et al., 2013), so the uncanny valley may not be determined by a single factor (Hanson, 2006). To examine the multiple potential sources of the uncanny valley, the present study examined neural correlates that are known to be manifestations of different psychological reactions and states.

Only a handful of neuro-imaging studies have been conducted on the uncanny valley. Vaitonytė et al.'s (2023) scoping review found 13 neuro-imaging studies, of which seven used EEG. The earliest EEG study was conducted by Urgen et al. (2013) who examined sensorimotor mu and theta oscillations when observing actions of a mechanical robot, an android, and a human actor. The researchers considered the mu oscillations to reflect the mirror neuron system that simulated observed actions in the observer's mind, but they did not find a significant modulation across the three agents, suggesting that the type of agent did not change the observers' perception of actions. On the other hand, they found larger theta power for the mechanical robot than for the android or human actor. The researchers interpreted this outcome as reflecting greater memory processing to encode the action of the mechanical robot than those of the android or the human actor (but see the next section for an alternative interpretation of the theta oscillations). Subsequently, Urgen et al. (2018) examined event-related potentials (ERPs) with the movies of the three types of agents as in their earlier study and found a larger N400 component when their participants observed an android with mechanistic movements than when they observed a human with natural human movements or a mechanistic robot with mechanistic movements. The researchers explained that the larger N400 resulted from the mismatch between the human-like appearance and the mechanistic movement that violated the observer's expectancy. Mustafa et al. (2017) also had their participants observe computer-generated (CG) images of human faces with varied degrees of realism and found that the N400 was smaller for human faces and comic-style CG faces, both of which did not involve any violation of expectancy for given categories of stimuli, than for human-like CG faces that were rated more uncanny. It should be noted that, as in Urgen et al.'s (2018) study, Mustafa et al.'s stimuli also included mouth movements to simulate speeches (but without the actual speech sounds). These results, thus, could have emerged from an expectancy violation caused by

the mismatch between the agents and their movements rather than direct reactions to uncanny faces, because the N400 can be enhanced for unnatural body movements (Syrov et al., 2022).

Other ERP components examined for the uncanny valley include the N170 and the late positive potential (LPP). The N170 is thought to be an index of facial processing, which is larger for human faces than for non-human faces (e.g., Gajewski & Stoerig, 2011) but also for faces with emotional expressions than for those with neutral expressions (Dubal et al., 2011; Hinojosa et al., 2015). Schindler et al. (2017) created CG facial images and found that the N170 increased from a cartoon face to more human-like faces but then decreased as the images approached a real human face. The LPP is thought to reflect the arousal component of emotional stimuli (rather than valence). Shindler et al. found that the LPP increased for more human-like faces and was largest for a real human face. Interestingly, another study found that the LPP was smaller for human faces than for computer-generated faces (Cheetham et al., 2015). These two findings contradict one another, and it is also noteworthy that the monotonic increase or decrease of the LPP across different degrees of human-likeness does not seem to mirror the non-linear relationship between affinity and human-likeness that is thought to be a salient characteristic of the uncanny valley.

Furthermore, Mustafa and Magnor (2016) used CG faces and found a larger P200 for uncanny faces than for real human faces, but their sample size was very small ($N = 10$). Bagdasarian et al. (2020) further examined the steady-state visual evoked potential (SSVEP) with CG faces. They only found that the SSVEP amplitude for a real human face was more strongly correlated with subjective ratings of the realness of the face than for CG faces, but the implication of their results for the uncanny valley is unclear.

1.2. The uncanny valley as an affective phenomenon

Regardless of the origin of the uncanny valley effect, most researchers would agree that its eventual consequences are affective. Mori's (1970) original definition of the uncanny valley hypothesis uses a Japanese term, *Shinwa-kan*, which was made of the two characters that roughly mean the feeling of "familiarity" and "compatibility." It has been translated into different English terms, such as "comfort," "likability," "warmth," and so on (e.g., Kätsyri et al., 2015; MacDorman et al., 2009; Mitchell et al., 2011; Reichardt, 1978), while Mori preferred "affinity" as its English translation (Mori et al., 2012). Others also used a reverse-coded negative connotation, such as "eeriness" (MacDorman & Ishiguro, 2006), or a compound scale, such as "pleasant vs. creepy" (Mathur & Reichling, 2016). Although the interpretations varied across studies, all seems to involve some sort of emotionality as the manifestation of the uncanny valley effect. In a recent study, two behavioral measures of affective responses, affective priming and the single-category implicit association test, were found to produce the uncanny valley effect as plotted against the human-likeness of robot faces as well (Yamaguchi, 2025).

Arguably, two common electrophysiological correlates of affective responses are related to the alpha (8-12 Hz) and theta (4-7 Hz) power bands. First, the frontal cortical activities have been shown to produce distinct involvements of the two hemispheric regions for positive and negative emotions (Davidson, 1992), whereby positive emotion results in a greater left-sided activation whereas negative emotion results in a greater right-sided activation. These hemispheric asymmetries are often observed in the alpha band (the *frontal alpha asymmetry*; Ahern & Schwartz, 1985; Davidson et al., 1979). To our knowledge, there has not been any study of the uncanny valley that examined the frontal alpha asymmetry.

Second, the mid- or anterior cingulate cortex is thought to regulate negative affect, whose activities are thought to be manifested in the theta band (the *frontal-midline theta*; Cavanagh & Shackman, 2015; Sammler et al., 2007). The aforementioned study by Urgen et al. (2013) examined the theta power while their participants viewed videos of three different agents (human, mechanical robot, and android). They found increased theta power for the mechanical robot compared to the android or human at the frontal and central sites, and they proposed that greater memory processing to encode the action of the mechanical robot was involved. An alternative interpretation of their finding is that the action of the mechanical robot induced a more negative reaction to the observer. In either case, their study only included three exemplars, so the implication for the uncanny valley effect may be limited.

1.3. The present study

The previous EEG studies provided some promising candidates for neural responses to the uncanny valley (e.g., Cheetham et al., 2015; Mustafa et al., 2017; Schindler et al., 2017; Urgen et al., 2018), but all of these findings are limited by the relatively narrow range of stimuli used in these studies, which prevented researchers from testing the expected non-linear relationship between an electrophysiological correlate and the human-likeness of their exemplars. Thus, we extended previous research over a wide range of images with varying levels of human-likeness, which were originally created for a survey study (self-report; Mathur et al., 2020) and were validated with behavioral measures (Yamaguchi, 2025). The present study examined three ERP components (N400, N170, and LPP)[1] and two EEG oscillatory components (frontal alpha asymmetry and frontal-midline theta).

2. Method

2.1. Participants

Thirty-nine students and staff from a local university were recruited and were paid £24 (GBP) for their participation; all reported normal or corrected-to-normal vision and no color blindness. One participant did not complete the entire session, and five participants were excluded based on the recording quality (see the Electrophysiological Data Processing section). In the end, there were 32 participants (24 females, 8 males; mean age = 26.00, SD = 5.01, range = 19-42) in the analysis. The analysis involved comparisons among polynomial regression (Mathur et al., 2020; Mathur & Reichling, 2016; Tu et al., 2020; Yamaguchi, 2025), for which there was no established method to determine an appropriate sample size. Thus, we performed computer simulations to determine that the sample size was sufficient to obtain reliable results (see Supplementary Materials). All participants read the participant information sheet and signed an informed consent form before data collection commenced. The experiment was reviewed and approved by the University of Essex's Research Ethics Committee.

2.2. Transparency and data availability

All stimuli, experimental programs, and analysis scripts are available from the OSF project page (https://osf.io/pf3j7/). The raw continuous data are available from the Harvard Dataverse (https://doi.org/10.7910/DVN/ZTCWPG). The analysis was not pre-registered but followed closely of our previous study (Yamaguchi, 2025).

2.3. Apparatus and stimuli

The experiment was programmed and controlled by Inquisit 6 (Millisecond, LLC.) with a local computer running Windows 10. Stimuli were presented on a 24-in flat screen monitor (120 Hz refresh rate). The monitor was located outside the EEG recording cage and was viewed by participants through a window on one side of the cage. The stimuli were images of 50 robot faces, which were selected from Mathur et al.'s (2020) stimulus set. We chose to use Mathur et al.'s images because their data showed a clear non-linear trajectory of self-rated likability scores of these robot faces as plotted against their human-likeness scores, which beautifully captured the uncanny valley effect. In the first author's previous study (Yamaguchi, 2025), 99 robot faces were selected from the original stimulus set by excluding similar images, which replicated the uncanny valley effect in self-rated likability scores. The stimulus set was then reduced to 50 robot faces by arranging the images in the order of self-rated human-likeness scores and choosing every other images in that order, so the reduced stimulus set would still cover the entire human-likeness spectrum. The resulting stimulus set was used in two behavioral experiments in that study, which also exhibited similar non-linear trajectories of behavioral affective scores plotted against self-rated human-likeness scores. The same set of 50 robot faces were used in the present experiment and illustrated in Figure 2. The remaining 49 images were also used for practice trials, which never showed up during the recorded trials.

Figure 2. Images of 50 robot faces used in the experiment in the order of their MK scores (from more mechanistic to more human-like) from study 1 of Yamaguchi (2025).

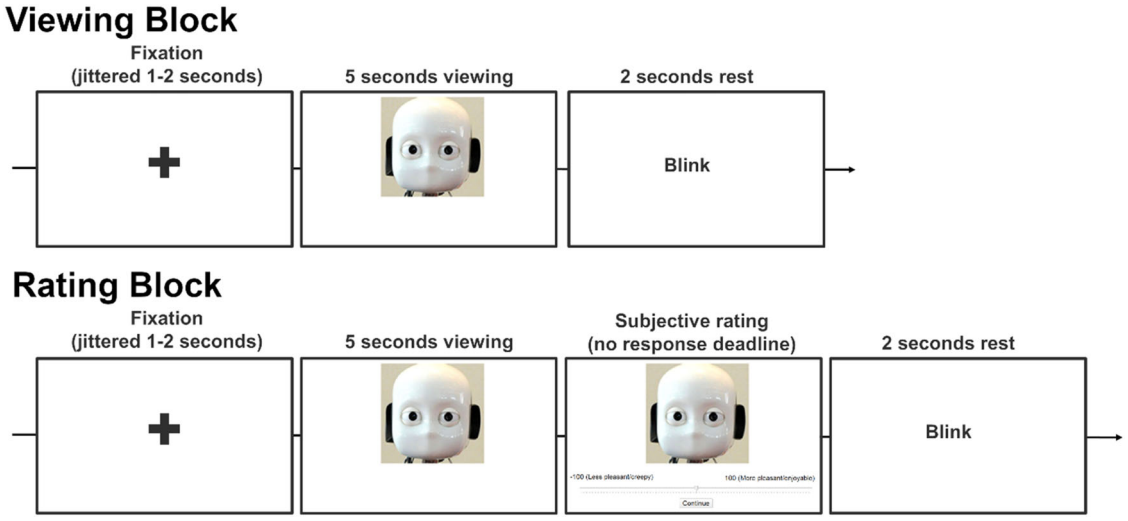

Figure 3. An illustration of the event sequence on a trial for the viewing-only blocks (top) and subjective human-likeness and likability rating blocks (bottom).

2.4. Procedure

The experiment was run individually. After the electrodes were set at satisfactory impedances ($< 15\text{k}\Omega$) with conductive gel, participants sat in a recording cage and were instructed on the task. They first performed practice trials with three separate blocks before they performed nine recording blocks. A recording session took no longer than 90 min. Figure 3 illustrates the sequence of displays on a trial.

The first practice block was a viewing-only block, in which participants were presented with an image of one of the robot faces and were asked to view it passively for five seconds. They were told to minimize the movements of their eyes, head, and body, as much as possible during this period. Each trial started with a fixation cross in the screen center. The duration of the fixation cross was randomly selected between 1 and 2 s, with 200-ms increments. An image of a robot face replaced the fixation cross and appeared to fit within 40% of the screen width and height in the center. After 5 s, the image was replaced with a word "Blink" for 2 s, which prompted participants to blink before the next trial started.

In the second and third blocks, participants viewed a robot face and rated the robot face for its likability or human-likeness, depending on the block. In the likability block, they were told to inspect the robot face within the 5-second viewing interval (without moving their eyes) and consider "how pleasant or creepy it would be if they were to interact with the robot in an everyday situation such as asking questions in a museum information booth." After the 5-second viewing interval, participants were presented with a slider and adjusted it with a computer mouse to indicate a likability within a range between −100 (creepy) and +100 (pleasant). In the human-likeness block, they rated "how machine- or human-like the robot looks;" participants adjusted the slider within a range between −100 (mechanical) and +100 (human-like). Following Mathur and Reichling (2016), we used the composite of pleasantness-creepiness rating as the affinity measure in this experiment, and we refer to it as "likability" score; we also follow the same study and refer to the human-likeness score as the mechanical-human (MH) score (also see Mathur et al., 2020).

After adjusting the slider, they clicked a button below the slider to proceed to the next trial. The initial position of the slider was set at zero, and participants were not allowed to proceed to the next trial until they moved the slider from the initial position. The trial ended with a 2-second blank display. The order of the likability and human-likeness blocks were counterbalanced across participants, and the order was maintained for the recording blocks. Participants could perform as many practice trials as they needed, with a minimum of 10 trials in each block. During the practice blocks, the experimenter monitored the movements of the eyes, head, and body and explained the importance of keeping still to their best efforts during the EEG recording while referring to their brain activities displayed on the screen. EEG signals were not recorded during the practice blocks.

After the practice blocks, participants performed nine test blocks alone in the cage. Each block consisted of 50 robot faces. The first three blocks were always the viewing only blocks, and the following six blocks were divided into sets of three likability blocks and of three human-likeness blocks. Each of the 50 robot faces were presented once in random order in each block.

2.5. EEG recording and data processing

The EEG was recorded by using SynAmps RT and Curry 8 (Compumedics Neuroscan) running on a Windows 10 computer, separate from the one presenting the stimuli. The data were sampled at 1000 Hz and were recorded continuously from 60 scalp (Ag/AgCl) electrodes (BrainCap, Brain Products) positioned according to the international 10 − 10 system (FP1/2/z, AF3/4/7/8/z, F1/2/3/4/5/6/7/8/z, FC1/2/3/4/5/6/7/8/z, C1/2/3/4/5/6/z, T7/8, CP1/2/3/4/5/6/z, TP7/8, P1/2/3/4/5/6/7/8/z, PO3/4/7/8/z, O1/2/z), with the ground electrode at AFz. The left earlobe was used as an online reference. Additional monopolar electrodes were also positioned at the two mastoids (which was used as offline re-references) and below the left eye (which was used along with FP1 to monitor vertical eye movement). In addition, three sets of bipolar electrodes were used at the outer canthi of both eyes (to monitor horizontal eye movement), at the corrugator spercilii above the right eye, and at the zygomaticus major on the left cheek (to monitor facial electromyography). We did not analyze the facial EMG data.

The EEG data were processed and analyzed by using custom scripts using EEGLAB (Delorme & Makeig, 2004), ERPLAB (Lopez-Calderon & Luck, 2014), and BioSig (Vidaurre et al., 2011) functions. The data were first re-referenced to the average of the two mastoids and were band-passed between 1 Hz and 45 Hz. They were then down-sampled to 200 Hz. Independent component analysis (ICA; Delorme et al., 2007) was applied to separate oculomotor, muscular, and line artifacts, from the scalp electrodes by using ICLabel (Pion-Tonachini et al., 2019). The data were epoched to form stimulus-locked segments from 100 ms pre-stimulus to 1000 ms post-stimulus, which were corrected based on the mean voltage for the 100-ms pre-stimulus interval as the baseline. The trials for which the peak-to-peak amplitude exceeded 100 uV (e.g., blinks and motion artifacts) or for which the channels of interest were flat were removed before averaging. Participants for whom the proportion of excluded trials exceeded 30% of all trials were removed from the subsequent analyses.

The ERP components of interest included the N400, N170, and LPP. They were computed, respectively, as the mean amplitudes in the post-stimulus intervals between 350 to 550 ms (N400), between 150 to 200 ms (N170), and between 600 to 1000 ms (LPP). We also computed the P300 amplitude, but

this is not reported below because of the temporal overlap with the N400 and the N170. The N400 was examined for the frontal-parietal midline (Fz, FCz, Cz, CPz), the N170 was examined for the lateral sites in the occipital region (PO7/8 and O1/2), and LPP were examined for the parietal midline (CPz, Pz, and POz).

For the frontal alpha asymmetry and front-midline theta, we first obtained the log-band-power for the alpha (8-12 Hz) and theta (4-7Hz) bands from the pre-processed EEG signals, with a 1-s smoothing window at each sampled time point (i.e., 5-ms intervals). The frontal alpha asymmetry was computed by subtracting the mean amplitudes at the left hemispheric sites (AF3, F3, and F7) from those at the respective right hemispheric sites (AF4, F4, and F8). The midline theta was computed for channels Fz, FCz, and Cz. They were then averaged for each trial over the interval from 400 ms pre-stimulus to 1000 ms post-stimulus. The grand average was then obtained for each robot face.

2.6. Data analysis

To evaluate the uncanny valley, we adopted the analysis strategy used by Mathur and Reichling (2016) and others (Mathur et al., 2020; Tu et al., 2020) for subjective likability ratings. Our previous study also used the same approach to evaluate behavioral measures of the uncanny valley (Yamaguchi, 2025). This analysis involved first computing the mean likability score for each robot face for each participant and then taking the mean and standard deviation of the likability scores for each face across participants. These likability scores for individual robot faces served as the unit of polynomial regression models, which were regressed on the polynomial terms of the MH scores sequentially from lower degree models to higher degree models, up to the fifth degree (i.e., MH, $MH^2, \ldots, MH^5$). The data points were weighted according to the inverse of their variances in the regression models, so model estimations weigh on more stable data points. Each model was compared to a lower-degree model and a higher-degree model in terms of F-tests, and the lowest-degree model that fit the data significantly better than a lower-degree model but was not significantly worse than a higher-degree model was considered to be the most parsimonious model for the given data.

We also computed the differences in the Bayesian Information Criteria (BIC) for each pair of the models, which are shown in the respective tables below; ΔBIC was determined by subtracting the BIC for a lower-degree model from the BIC for a higher-degree model, so $\Delta BIC > 0$ indicated a better fit for a higher-degree model than a lower-degree model. Based on previous findings, the uncanny valley effect would result in the most parsimonious model that is either the third- or fourth-degree polynomial model (e.g., Mathur et al., 2020; Yamaguchi, 2025).

In the present study, the likability and MH scores were derived from the subjective rating. For the electrophysiological data, the likability score was replaced by the respective ERP components (N400, N170, and LPP) and frequency band power (frontal alpha asymmetry and frontal-midline theta). The polynomial regression analyses were performed in R Studio (R Core Team, 2021) with the following packages: lme4 (Bates et al., 2015), tidyverse (Wickham et al., 2019), and ggpubr (Kassambara, 2022).

3. Results

The results are summarized in Tables 1 and 2. Figure 4 shows the grant averages of the stimulus-locked ERPs, and Figure 5 shows the scatterplots of the self-report and ERP measures against the MH scores along with the predictions of the most parsimonious models for the respective measures. Figure 6 shows the grand averages of the stimulus-locked oscillatory components, and Figure 7 shows their scatterplots against the MH scores. Overall, the grand means across all robot faces were −15.10 ($SD = 24.96$, range = [−71.77, 37.41]) for self-rated likability and −13.54 ($SD = 52.33$, range = [−84.57, 76.40]) for the MH scores.

3.1. Subjective likability rating

Self-report likability ratings showed that the most parsimonious model was the third-degree model (which meant that the third-degree model fit the data significantly better than the second-degree model

Table 1. Comparisons of polynomial regression models for the ERP (N400, N170, and LPP) and for self-rated likability.

Measure	Electrode location	Model comparison	*F*	*p*	η_p^2	ΔBIC
Likability	–	First vs. second	F(1,47) = 62.02	<0.001	0.407	22.19
		Second vs. third	**F(1,46) = 42.23**	**<0.001**	**0.467**	**27.51**
		Third vs. fourth	F(1, 45) = 1.57	0.214	0.033	−2.26
N400	Fz	First vs. second	F(1,47) = 3.47	0.069	0.062	−0.72
		Second vs. third	**F(1,46) = 7.80**	**0.008**	**0.148**	**4.11**
		Third vs. fourth	F(1,45) = 0.79	0.379	0.018	−3.02
	FCz	First vs. second	F(1,47) = 2.66	0.110	0.048	−1.47
		Second vs. third	**F(1,46) = 8.10**	**0.007**	**0.153**	**4.37**
		Third vs. fourth	F(1,45) = 0.04	0.334	0.021	−2.84
	Cz	First vs. second	F(1,47) = 2.52	0.119	0.046	−1.53
		Second vs. third	**F(1,46) = 6.97**	**0.011**	**0.135**	**3.32**
		Third vs. fourth	F(1,45) = 0.72	0.402	0.016	−3.11
	CPz	First vs. second	F(1,47) = 4.53	0.039	0.074	−0.09
		Second vs. third	**F(1,46) = 10.52**	**0.002**	**0.184**	**6.27**
		Third vs. fourth	F(1,45) = 2.58	0.115	0.055	−1.07
N170	PO7	First vs. second	F(1,47) = 1.46	0.234	0.030	−2.41
		Second vs. third	F(1,46) = 0.08	0.780	0.002	−3.83
		Third vs. fourth	F(1,45) = 0.12	0.736	0.002	−3.79
	PO8	First vs. second	F(1,47) = 0.43	0.515	0.009	−3.46
		Second vs. third	F(1,46) = 0.07	0.799	0.002	−3.84
		Third vs. fourth	F(1,45) = 1.04	0.313	0.022	−2.79
	O1	**First vs. second**	**F(1,47) = 6.39**	**0.015**	**0.122**	**2.59**
		Second vs. third	F(1,46) = 0.17	0.686	0.004	−3.73
		Third vs. fourth	F(1,45) = 0.54	0.467	0.012	−3.32
	O2	**First vs. second**	**F(1,47) = 4.74**	**0.035**	**0.096**	**1.11**
		Second vs. third	F(1,46) = 0.18	0.676	0.004	−3.71
		Third vs. fourth	F(1,45) = 0.09	0.760	0.002	−3.81
LPP	CPz	**First vs. second**	**F(1,47) = 17.98**	**<0.001**	**0.269**	**11.74**
		Second vs. third	F(1,46) = 0.60	0.442	0.013	−3.26
		Third vs. fourth	F(1,45) = 1.78	0.189	0.039	−1.94
	Pz	**First vs. second**	**F(1,47) = 23.15**	**<0.001**	**0.336**	**16.57**
		Second vs. third	F(1,46) = 0.60	0.442	0.003	−3.77
		Third vs. fourth	F(1,45) = 1.78	0.189	0.035	−2.11
	POz	**First vs. second**	**F(1,47) = 23.49**	**<0.001**	**0.335**	**16.48**
		Second vs. third	F(1,46) = 0.50	0.484	0.011	−3.38
		Third vs. fourth	F(1,45) = 1.52	0.224	0.033	−2.24

Note: Bold indicates the most parsimonious model. ΔBIC is the difference in BIC (positive values indicate better fit for a higher degree model). $\alpha = 0.05$.

Table 2. Comparisons of polynomial regression models for frequency components (frontal alpha asymmetry and frontal-midline theta).

Measure	Electrode location	Model comparison	*F*	*p*	η_p^2	ΔBIC
Alpha asymmetry	AF4 - AF3	First vs. second	F(1,47) = 0.11	0.740	0.002	−3.80
		Second vs. third	F(1,46) = 3.10	0.086	0.065	−0.53
		Third vs. fourth	F(1,45) < 0.01	0.958	<0.001	−3.91
	F4 - F3	First vs. second	F(1,47) = 0.11	0.746	0.002	−3.80
		Second vs. third	F(1,46) = 0.27	0.609	0.006	−3.62
		Third vs. fourth	F(1,45) = 1.31	0.259	0.029	−2.44
	F8 - F7	First vs. second	F(1,47) = 0.60	0.442	0.006	−3.63
		Second vs. third	F(1,46) = 0.26	0.611	0.002	−3.81
		Third vs. fourth	F(1,45) = 1.23	0.273	0.011	−3.33
Midline theta	Fz	**First vs. second**	**F(1,47) = 13.02**	**<0.001**	**0.094**	**8.65**
		Second vs. third	F(1,46) = 0.04	0.839	<0.001	−3.87
		Third vs. fourth	F(1,45) = 1.93	0.172	0.042	−1.75
	FCz	**First vs. second**	**F(1,47) = 17.54**	**<0.001**	**0.284**	**12.77**
		Second vs. third	F(1,46) < 0.01	0.974	<0.001	−3.91
		Third vs. fourth	F(1,45) = 0.84	0.364	0.019	−2.95
	Cz	**First vs. second**	**F(1,47) = 34.22**	**<0.001**	**0.438**	**24.87**
		Second vs. third	F(1,46) = 0.03	0.859	<0.001	−3.88
		Third vs. fourth	F(1,45) = 0.22	0.640	0.005	−3.66

Note: Bold indicates the most parsimonious model. ΔBIC is the difference in BIC (positive values indicate better fit for a higher degree model). $\alpha = 0.05$.

while the fourth-degree model did not improve the fit significantly). The prediction of the model (see Figure 5) indicated an initial rise of the likability score as the MH score increased, followed by a steep decline of the likability score at around the MH score of −50, and this is then followed by a steep recovery at around the MH score of 30. This replicated the uncanny valley effect in the self-rated likability score observed for the same stimulus set (Yamaguchi, 2025).

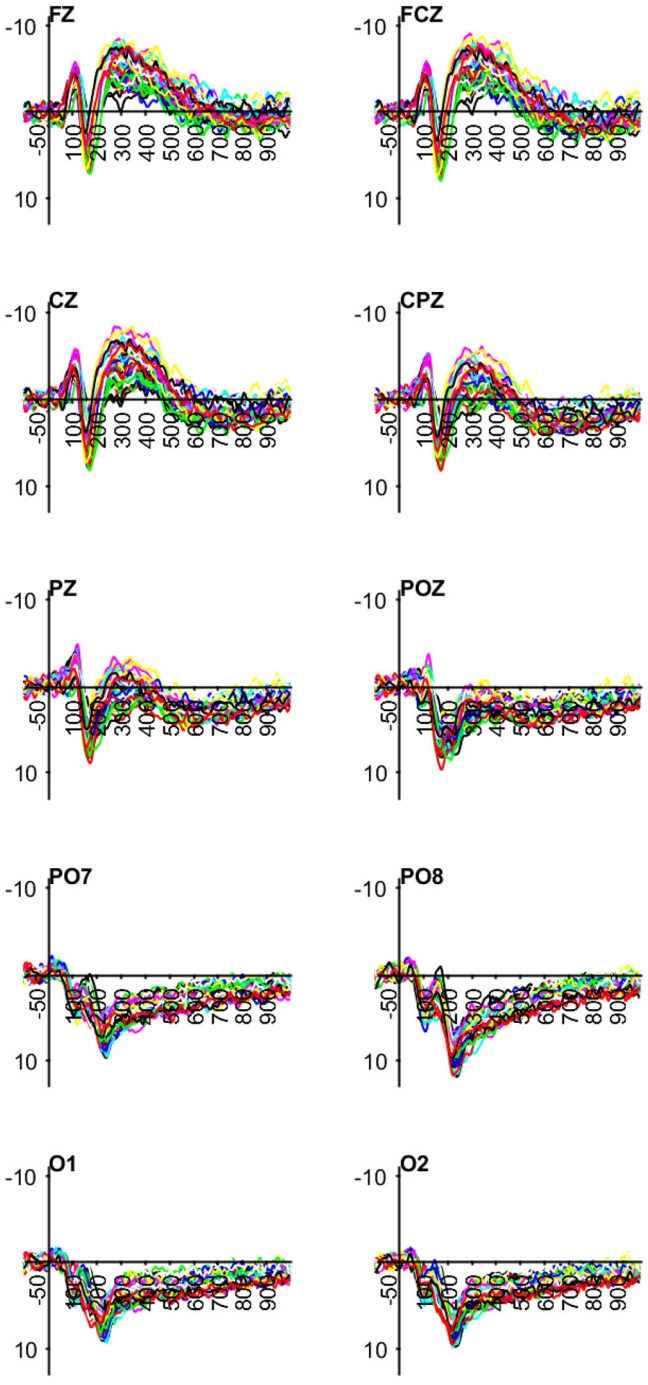

Figure 4. Grand average stimulus-locked ERPs for individual robot faces (*x*-axis is in milliseconds and *y*-axis is in microvolts; *y*-axis is inverted).

3.2. Event-related potentials

The N400 for all four sites showed that the most parsimonious model was the third-degree polynomial (see Table 1). Figure 5 shows that there was an initial *decrease* of the N400 amplitude as the MH score increased (note that the N400 is a negative ERP component, so a larger N400 means higher negative values), but it *increased* at around the MH score of −50 and then declined again at around the MH score of 50, similarly to the pattern observed for the self-report likability score.

In contrast, for the N170 and the LPP, the most parsimonious model was the second-degree model. For the N170 (see the four middle row panels in Figure 5), the second-degree model showed a slight, but steady, decrease of the N170 amplitude (note again that the N170 is a negative ERP component, so

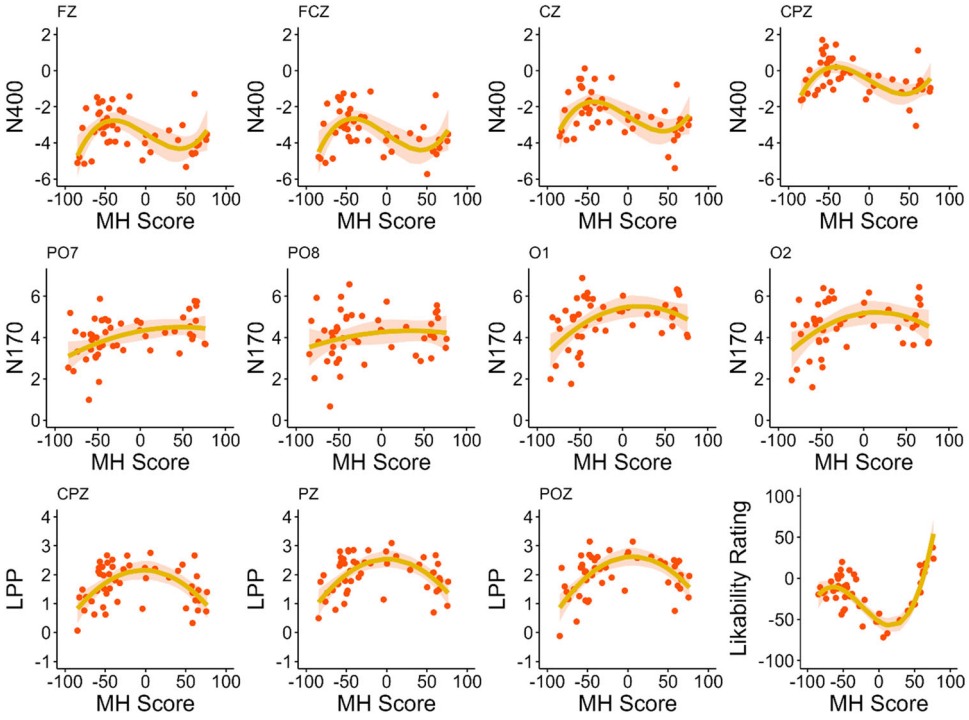

Figure 5. Scatter plots of the ERP components (N400, N170, and LPP) and subjective likability scores against the human-likeness scores (MH score). The orange lines are the predictions of the most parsimonious model fit to the respective data. The dot points represent individual robot faces. The shaded area represents the point-wise 95% confidence interval of the means.

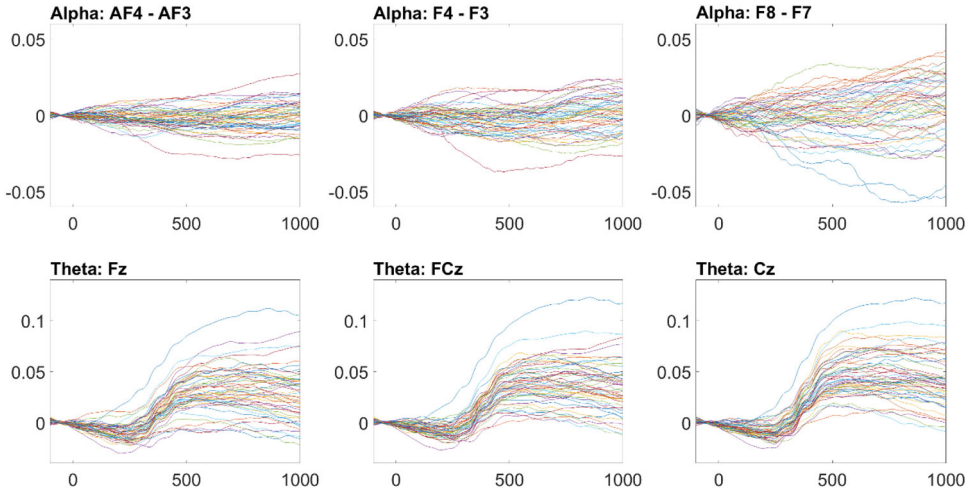

Figure 6. Grand average stimulus-locked frontal alpha asymmetry (top panels) and frontal-midline theta (bottom panels) for individual robot faces (*x*-axis is in milliseconds and *y*-axis is in log[μV^2/hz]).

a larger N170 means more negative) from lower to higher MH scores, which arrived at a peak near the MH score of 0. This pattern was statistically supported for the two occipital sites (O7/8) but not for more parietal sites (PO7/8). There were small decelerations in the change of the N170 amplitude for higher values of the MH score, which remained flat for near human-like faces.

For the LPP (see the three bottom-left panels in Figure 5), the amplitude increased quickly as the MH scores increased, which again reached a peak near the MH score of 0. It then decreased quickly as the MH score increased further. This outcome suggests that the LPP increased as the human-likeness increased for mechanical robot faces and then decreased for more human-like robot faces. It never recovered for near human-like faces. Instead, the LPP may depend on the categorical ambiguity as it reached at the highest point for faces that were neither mechanical nor human-like. A heightened

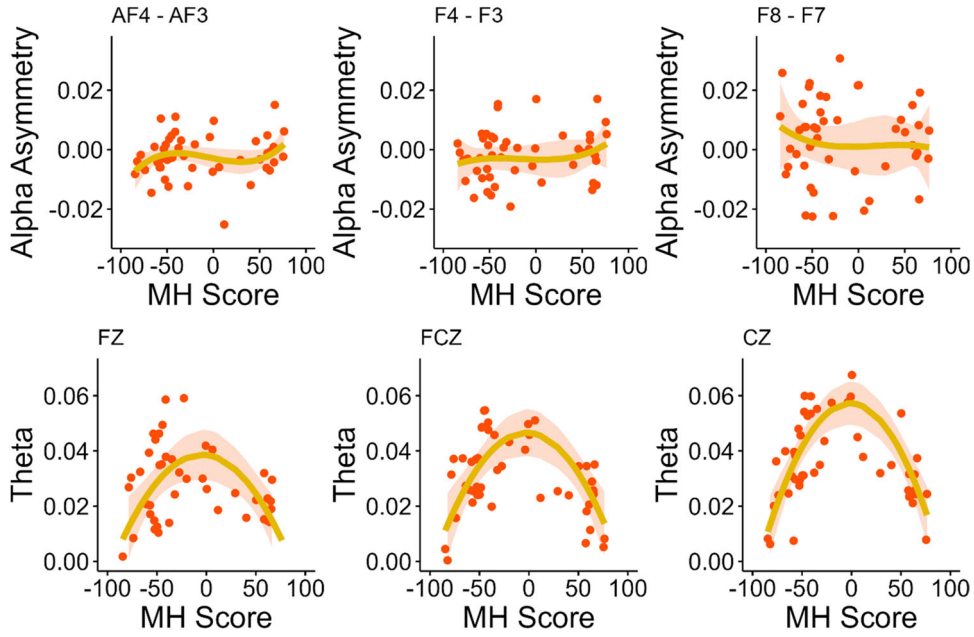

Figure 7. Scatter plots of the frequency components (frontal alpha asymmetry and midline theta) against the human-likeness scores (MH score). The orange lines are the predictions of the most parsimonious model or of the third-degree polynomial regression fit (in case there was no parsimonious model) to the respective data. The dot points represent individual robot faces. The shaded area represents point-wise 95% confidence interval of the means.

arousal might result from categorical ambiguity, which may also contribute partly to the uncanny feeling of robot faces as the bottom of the uncanny valley in the self-rated likability also resulted from this region.

3.3. Oscillatory EEG measures

For the frontal alpha asymmetry, no polynomial models up to the fifth degree satisfied the criteria for a parsimonious model for all three electrode pairs. For the F3/4 pair, the third-degree model showed a marginally better fit than the fourth-degree model, but it did not reach the level of statistical significance. For the AF3/4 and F3/4 electrode pairs, there were slight tendencies of the non-linear pattern that were similar to those observed for the self-rated likability ratings, and the F7/8 pair showed a reverse pattern. These tendencies were not statistically significant.

For the frontal-midline theta, all three sites showed that the second-degree model was the most parsimonious model. The frontal-midline theta increased as the MH score increased for more mechanical robot faces, reaching the peak at around the MH score of 0. It then decreased for higher MH scores for more human-like faces, which was similar to those observed for the LPP but differed from that of the self-report likability score.

4. Discussion

Several psychological explanations of the uncanny valley effect have been proposed, which attributed the phenomenon to perceptual properties (e.g., categorization difficulty, configural distortion; e.g., Diel & MacDorman, 2021; Yamada et al., 2013), cognitive processes (e.g., expectancy mismatch; MacDorman & Chattopadhyay, 2016) or socio-emotional reactions (e.g., mind attribution; Gray & Wegner, 2012). The present study used a range of robot faces that varied in their human-likeness to examine potential electrophysiological correlates of the uncanny valley effect. As in previous studies (Mathur et al., 2020; Tu et al., 2020; Yamaguchi, 2025), we replicated the non-linear relationship between the human-likeness and self-rated likability scores, which is similar to Mori's (1970) uncanny valley hypothesis. Very similar non-linear trajectories to that of the likability score were observed for the N400 at the midline electrode sites, suggesting that the N400 is a robust electrophysiological

correlate of the uncanny valley effect. This outcome corroborates two previous studies (Mustafa et al., 2017; Urgen et al., 2018) and the proposal that the uncanny valley originated from prediction violation, or expectancy mismatch, for human-like faces. The present results further support this account by showing the uncanny valley without a movement mismatch in the animated stimuli that was involved in the previous studies.

In contrast, the N170 decreased as the human-likeness of robot faces increased, but the decreasing trend leveled off as the human-likeness increased further. This outcome was similar to Schindler et al.'s finding but was different from the uncanny valley effect observed in the self-rated likability score or the N400. These results may be best understood as showing that the N170 reflected human-likeness, rather than the uncanniness. Thus, as the N170 is thought to reflect an early stage of visual facial processing (Schindler et al., 2017), a lack of the expected non-linear trajectory of the uncanny valley seems to suggest that the uncanny valley reflect later cognitive reactions to uncanny faces rather than early perceptual processes. We should also note that as the N170 is sensitive to a variety of facial and perceptual properties in the stimulus set used in the experiment, there is still a possibility that the uncontrolled facial properties masked the uncanny valley effect. Hence, the potential contributions of early perceptual processes to the uncanny valley should be subjected to future scrutiny in the future investigations.

Similarly, the LPP, which is thought to reflect reactions to emotional contents of stimuli (Schindler et al., 2017) or arousal (Cheetham et al., 2015), did not produce a non-linear trajectory similar to the self-rated likability score. Instead, it produced a parabolic relationship with the human-likeness, whereby the LPP peaked at the midpoint between mechanical robots and near-human-like faces. A similar result was also found for the frontal-midline theta power. Although theta power is associated with processing of emotional contents (e.g., Sammler et al., 2007), the relationship with theta power and emotional reaction may be more indirect. Instead, the frontal-midline theta appears to be associated more reliably with cognitive control (Cavanagh & Shackman, 2015) or memory retrieval (Karakaş, 2020). It is likely that the parabolic relationship of the frontal-midline theta and the LPP with the human-likeness score, which peaked near the midpoint of human-likeness score, reflected categorical confusion, as stronger cognitive control may be required to process ambiguous faces that activate conflicting categorical representations. Yet, because the peaks of the parabolic curves also coincided with the lowest point of the uncanny valley in self-rated likability scores, it is still possible that categorical ambiguity also contributed to creating uncanny feelings in the observers, although the curves differed from the expected trajectory at the lower end of human-likeness.

It is somewhat surprising that the frontal alpha asymmetry did not show a clear non-linear trajectory over human-likeness of robot faces. The uncanny valley is often assumed to reflect emotional or empathic reactions to human-like robots and characters, such as "eeriness" (MacDorman & Ishiguro, 2006), "comfort" (MacDorman et al., 2009), and "pleasantness" (Mathur & Reichling, 2016). Indeed, our previous study used a behavioral measure, *affective priming*, which is thought to capture affective responses, and showed that the behavioral measure produced the uncanny valley effect similar to that observed in the self-rated likability score. Nevertheless, the lack of a clear uncanny valley effect in the frontal alpha asymmetry suggests that the uncanniness of robot faces does not necessarily result in affective reactions. It may still be immature to conclude that there is no affective consequence of the uncanny valley because the failure to find a statistically reliable effect does not allow us to dismiss its existence, so it is possible that the affective effect is simply more difficult to detect statistically. Nevertheless, the present results raise the possibility that the affective impact might play a more minor role than commonly believed, as compared to a "surprisal" effect as inferred from the N400. Thus, human-like robot faces may cause a violation of the expectancy for human faces, and this could result in discomfort due to the surprisal reaction; nonetheless, they may not result in an immediate negative emotion toward these faces. This possibility would require further scrutiny, but it could offer important implications for the psychological theories of the uncanny valley.

4.1. Potential limitations and contributions to human–robot interaction

The present findings provided novel insights into what might, or might not, underly the uncanny valley phenomenon. As in recent studies (e.g., Mathur et al., 2020; Yamaguchi, 2025), the use of the

exemplars that covered a wide range of human-likeness is useful to determine which neural responses are most likely reflections of the uncanny valley effect. We should also note a limitation of using such a stimulus set because exemplars are heterogenous and are not tightly controlled in any specific facial or perceptual features. There can be hidden biases that are not explicitly tested in this or previous studies. It is possible that the uncanny valley curve observed in the series of studies using this stimulus set can be explained in terms of an unknown property of the images. Finding out such a feature in future investigations would contribute to a significant advance in our understanding of the uncanny valley.

Another limitation of the stimuli used in this experiment is that the images were static. Although the use of static images is advantageous in that it eliminates the possible interpretations of neural responses as reflecting unnatural movements (e.g., Urgen et al., 2018), it might also limit potential implications of the present results for human-robot interaction in the real world. As robots become more accessible and prevalent, physical interactions with robots might mediate people's perceptions of them and make uncanny appearance more acceptable, or vice versa. Therefore, it would become more important to examine whether people's familiarity and experience interacting with robots could bridge the uncanny valley.

A methodological contribution in the present study is the use of the analysis approach involving polynomial regression models that used item-level mean amplitudes of the ERPs and EEG oscillatory measures. This approach was also used in the previous study with behavioral measures (Yamaguchi, 2025), which noted a small number of repetitions per participant was sufficient to find a robust uncanny valley effect, because these dependent measures were averaged across all participants for each item. The present study also involved a small number of trial repetition (9 trials) per image, which was indeed much smaller than included in a typical ERP study. Again, the present study offered clear statistical results. It is yet to be seen whether this analysis approach could be extended to other types of stimuli or topics other than the uncanny valley (e.g., likability of commercial products or human faces), but it does provide a viable alternative to more conventional analysis techniques of EEG data. Finally, given the finding in the present study that the alpha asymmetry did not show an expected uncanny valley effect, a more targeted examination of the subcortical regions by using other imaging techniques (e.g., fMRI) would be needed to confirm the involvement of emotional reactions in the uncanny valley effect. Such examinations are particularly important to understand possible impact of uncanny appearance of robots on users' emotions and, by a natural extension, on their well-beings.

5. Concluding remarks

The present study found the N400 to produce non-linear trajectories that resembled that observed in self-rated likability scores against the human-likeness of robot faces, which is consistent with the account that the uncanny valley originates from the expectancy violation of human-like faces. The LPP and theta oscillations also peaked at or near the point of the human-likeness where the likability score hit the bottom of the uncanny valley. Although these electrophysiological components may reflect psychological responses to robot faces that contribute to uncanny feelings (such as categorical ambiguity), the outcomes have left the question as to why these electrophysiological responses did not recover near human faces as observed in the self-rated likability scores or N400. We, thus, suggest that the expectancy violation is the main source of the uncanny valley effect. As the present study is concerned only with neural correlates and does not establish a causal relationship between neural activities and the uncanny valley, future studies should aim to establish causal relationships between the uncanny valley effect and the neural processes underlying these neural correlates.

Note

1. Although Mustafa and Magnor (2016) examined the P200, we excluded this component because of its temporal overlap with the N170. One reviewer of an earlier draft suggested visual mismatch negativity (vMM), which is also used in an oddball paradigm (e.g., Kimura et al., 2009), but we excluded it from our list

of candidates because of its temporal and spatial overlap with the N170. For a similar reason, we also excluded the P300 from our list of candidate ERP components because of its temporal overlap with the N400.

Acknowledgements

The experimental program and analysis scripts are available for reanalysis purposes from the Open Science Framework project page (https://osf.io/pf3j7/). The raw EEG data are available from the Harvard Dataverse (https://doi.org/10.7910/DVN/ZTCWPG). The authors thank Ethan Richards, Fabien Coomber, and I-Cheng Liu for their assistance in data collection, and Keisuke Fukuda, Heinrich Liesefeld, and Eva Gutierrez-Sigut for valuable discussions on the data.

Author contributions

CRediT: **Motonori Yamaguchi**: Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing; **Maemi J. Bautista**: Formal analysis, Writing – review & editing; **Ian Daly**: Formal analysis, Methodology, Writing – original draft, Writing – review & editing.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

The project was supported partly by a grant from ESSEXlab, Univeristy of Essex. The authors declare that they had no conflict of interest with respect to their authorship or the publication of this article. For the purpose of open access, the author has applied a Creative Commons Attribution (CC-BY) license to any Author Accepted Manuscript version arising.

ORCID

Motonori Yamaguchi http://orcid.org/0000-0002-8405-9741
Maemi J. Bautista http://orcid.org/0009-0009-5427-5783
Ian Daly http://orcid.org/0000-0001-5489-0393

References

Ahern, G. L., & Schwartz, G. E. (1985). Differential lateralization for positive and negative emotion in the human brain: EEG spectral analysis. *Neuropsychologia*, *23*(6), 745–755. https://doi.org/10.1016/0028-3932(85)90081-8

Bagdasarian, M. T., Hilsmann, A., Eisert, P., Curio, G., Müller, K.-R., Wiegand, T., & Bosse, S. (2020). EEG-based assessment of perceived realness in stylized face images. *Proceedings of the 2020 Twelfth International Conference on Quality of Multimedia Experience*.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

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*(3), 759–771. https://doi.org/10.1016/j.chb.2012.11.021

Cavanagh, J. F., & Shackman, A. J. (2015). Frontal midline theta reflects anxiety and cognitive control: Meta-analytic evidence. *Journal of Physiology, Paris*, *109*(1–3), 3–15. https://doi.org/10.1016/j.jphysparis.2014.04.003

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

Davidson, R. J. (1992). Emotion and affective style: Hemispheric substrates. *Psychological Science*, *3*(1), 39–43. https://doi.org/10.1111/j.1467-9280.1992.tb00254.x

Davidson, R. J., Schwartz, G. E., Saron, C., Bennett, J., & Goleman, D. J. (1979). Frontal versus parietal EEG asymmetry during positive and negative affect. *Psychophysiology*, *16*(2), 202–203. https://doi.org/10.1016/j.jneumeth.2003.10.009

Delorme, A., & Makeig, S. (2004). EEGLAB: An open-source toolbox for analysis of single-trial EEG dynamics. *Journal of Neuroscience Methods*, *134*(1), 9–21. https://doi.org/10.1016/j.jneumeth.2003.10.009

Delorme, A., Sejnowski, T., & Makeig, S. (2007). Enhanced detection of artifacts in EEG data using higher-order statistics and independent component analysis. *NeuroImage*, *34*(4), 1443–1449. https://doi.org/10.1016/j.neuroimage.2006.11.004

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. 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), 1–33. https://doi.org/10.1145/347074

Dubal, S., Foucher, A., Jouvent, R., & Nadel, J. (2011). Human brain spots emotion in non humanoid robots. *Social Cognitive and Affective Neuroscience*, *6*(1), 90–97. https://doi.org/10.1093/scan/nsq019

Gajewski, P. D., & Stoerig, P. (2011). N170—An index of categorical face perception? An ERP study of human, nonhuman primate, and dog faces. *Journal of Psychophysiology*, *25*(4), 174–179. https://doi.org/10.1027/0269-8803/a000057

Gray, K., & Wegner, D. M. (2012). Feeling robots and human zombies: Mind perception and the uncanny valley. *Cognition*, *125*(1), 125–130. https://doi.org/10.1016/j.cognition.2012.06.007

Hanson, D. (2006). Exploring the aesthetic range from humanoid robots. Proceedings of the ICCS/CogSci-2006 Long Symposium: Toward Social Mechanisms of Android Science (pp. 16–20).

Hinojosa, J. A., Mercado, F., & Carretié, L. (2015). N170 sensitivity to facial expression: A meta-analysis. *Neuroscience and Biobehavioral Reviews*, *55*, 498–509. https://doi.org/10.1016/j.neubiorev.2015.06.002

Karakaş, S. (2020). A review of theta oscillation and its functional correlates. *International Journal of Psychophysiology*, *157*, 82–99. https://doi.org/10.1016/j.ijpsycho.2020.04.008

Kassambara, A. (2022). *ggpubr: 'ggplot2' based publication ready plots*. R package version 0.5.0. https://CRAN.R-project.org/package=ggpubr

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

Kimura, M., Katayama, J., Ohira, H., & Schröger, E. (2009). Visual mismatch negativity: New evidence from the equiprobable paradigm. *Psychophysiology*, *46*(2), 402–409. https://doi.org/10.1111/j.1469-8986.2008.00767.x

Lopez-Calderon, J., & Luck, S. J. (2014). ERPLAB: An open-source toolbox for the analysis of event-related potentials. *Frontiers in Human Neuroscience*, *8*, 213. https://doi.org/10.3389/fnhum.2014.00213

MacDorman, K. F. (2024). Does mind perception explain the uncanny valley? A meta-regression analysis and (de)humanization experiment. *Computers in Human Behavior: Artificial Humans*, *2*(1), 100065. https://doi.org/10.1016/j.chbah.2024.100065

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., 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*(3), 695–710. https://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*(3), 297–337. https://doi.org/10.1075/is.7.3.03mac

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. https://doi.org/10.1016/j.cognition.2015.09.008

Mathur, M. B., Reichling, D. B., Lunardini, F., Geminiani, A., Antonietti, A., Ruijten, P. A. M., Levitan, C. A., Nave, G., Manfredi, D., Bessette-Symons, B., Szuts, A., & Aczel, B. (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

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 and uncanny valley. *i-Perception*, *2*(1), 10–12. https://doi.org/10.1068/i0415

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

Mori, M. (1970). Uncanny Valley. *Energy*, 33–35.

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

Mustafa, M., & Magnor, M. (2016). *EEG based analysis of the perception of computer-generated faces. Proceedings of the* [Paper presentation]. 13th European Conference on Visual Media Production (CVMP 2016). https://doi.org/10.1145/2998559.2998563

Mustafa, M., Guthe, S., Tauscher, J. P., Goesele, M., & Magnor, M. (2017). *How human am I? EEG-based evaluation of animated virtual characters* [Paper presentation]. Proceedings of the Conference on Human Factors in Computing Systems (Vol. *5*, pp. 5098–5108). https://doi.org/10.1145/3025453.3026043

Pion-Tonachini, L., Kreutz-Delgado, K., & Makeig, S. (2019). ICLabel: An automated electroencephalographic independent component classifier, dataset, and website. *NeuroImage*, *198*, 181–197. https://doi.org/10.1016/j.neuroimage.2019.05.026

R Core Team (2021). R*: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/

Reichardt, J. (1978). *Robots: Fact, fiction and prediction*. Thames & Hudson.

Sammler, D., Grigutsch, M., Fritz, T., & Koelsch, S. (2007). Music and emotion: Electrophysiological correlates of the processing of pleasant and unpleasant music. *Psychophysiology*, *44*(2), 293–304. https://doi.org/10.1111/j.1469-8986.2007.00497.x

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*(1), 45003. https://doi.org/10.1038/srep45003

Seyama, J., & Nagayama, R. (2007). The uncanny valley: Effect of realism on the impression of artificial human faces. *Presence: Teleoperators and Virtual Environments*, *16*(4), 337–351. https://doi.org/10.1162/pres.16.4.337

Syrov, N., Bredikhin, D., Yakovlev, L., Miroshnikov, A., & Kaplan, A. (2022). Mu-desynchronization, N400 and corticospinal excitability during observation of natural and anatomically unnatural finger movements. *Frontiers in Human Neuroscience*, *16*, 973229. https://doi.org/10.3389/fnhum.2022.973229

Tinwell, A., Nabi, D. A., & Charlton, J. P. (2013). Perception of psychopathy and the Uncanny Valley in virtual characters. *Computers in Human Behavior*, *29*(4), 1617–1625. https://doi.org/10.1016/j.chb.2013.01.008

Tu, Y.-C., Chien, S.-E., & Yeh, S.-L. (2020). Age-related differences in the uncanny valley effect. *Gerontology*, *66*(4), 382–392. https://doi.org/10.1159/000507812

Urgen, B. A., Kutas, M., & Saygin, A. (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

Urgen, B., Plank, M., Ishiguro, H., Poizner, H., & Saygin, A. P. (2013). EEG theta and Mu oscillations during perception of human and robot actions. *Frontiers in Neurorobotics*, *7*, 19. https://doi.org/10.3389/fnbot.2013.00019

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

Vidaurre, C., Sander, T. H., & Schlögl, A. (2011). BioSig: The free and open source software library for biomedical signal processing. *Computational Intelligence and Neuroscience*, *2011*, 935364–935312. https://doi.org/10.1155/2011/935364

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

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L. D., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T. L., Miller, E., Bache, S. M., Müller, K., Ooms, J., Robinson, D., Seidel, D. P., Spinu, V., … Yutani, H. (2019). Welcome to the tidyverse. *Journal of Open Source Software*, *4*(43), 1686. https://doi.org/10.21105/joss.01686

Yamada, Y., Kawabe, T., & Ihaya, K. (2013). Categorization difficulty is associated with negative evaluation in the "uncanny valley" phenomenon. *Japanese Psychological Research*, *55*(1), 20–32. https://doi.org/10.1111/j.1468-5884.2012.00538.x

Yamaguchi, M. (2025). Item-level implicit affective measures reveal the uncanny valley of robot faces. *International Journal of Human-Computer Studies*, *196*, 103443. https://doi.org/10.1016/j.ijhcs.2024.103443

About the authors

Motonori Yamaguchi is an Associate Professor in the Department of Psychology at the University of Essex. He obtained a PhD in Cognitive Psychology from Purdue University, and his research interests include cognitive control, skill acquisition, cognitive modeling, and human-robot interaction.

Maemi J. Bautista is a PhD student in psychology at Lancaster University. Her research examines how deepfake technology perpetuates gender-based violence, focusing on the impact of sexual digital forgeries on victims' embodiment of self.

Ian Daly is an Associate Professor in neural engineering and artificial intelligence at the University of Essex. He has over 10 years experience in neural engineering, affective computing, and physiological signal analysis. He obtained a PhD degree in Cybernetics from the University of Reading.