Review of neuroimaging and psychophysiological measures of the Uncanny Valley Effect

Ciechanowski L.[a,b,1], Kovbasiuk A.[b,1], Dworczyk M.[b], Kundziołka J.[b], .Lewandowska P.[b], Ozimek A.[b], Pawlak A.[b], Stróż A.[c], Szymanowska J.[a], Trybuś M.[b], Przegalińska A.[a]

[a] *Department of Management, Kozminski University, Jagiellonska 57/59 Warsaw, Poland, (+48) 22 519 22 89.*
[b] *Department of Psychology, University of Social Sciences and Humanities, Chodakowska 19/31, Warsaw, Poland, (+48) 22 517 96 00.*
[c] *Department of Neuroinformatics, University of Warsaw, Krakowskie Przedmieście 26/28, Warsaw, Poland, (+48) 22 55 20 000.*

Abstract

In the article, we are discussing the Uncanny Valley Effect (henceforth UVE or UV effect) in the light of its objective measures. In other words, we have carried out a critical literature review of the UV experiments employing psychophysiological and neuroimaging methods. We have analysed 19 papers from the Google Scholar and EBSCO databases, using the 'uncanny valley', ‘psychophysiology’ and ‘neuroimaging’ keywords and selected all papers which were experimental and investigated UVE, and employed one or more of the neuroimaging or psychophysiological tools. In the paper we first describe the UV hypotheses that are mentioned in the literature, with an emphasis on review and meta-analytic articles. Then we critically evaluate the selected papers and present them in a context of their hypotheses, used methodology, and outcomes. In the end, we show the shortcomings of the research carried out so far, and present our suggestions of pushing the domain forward.

Keywords

uncanny valley effect, neuroimaging, psychophysiology, human-computer interaction, affective robotics, literature review

[1] L. Ciechanowski and A. Kovbasiuk have equally contributed to the article.
Corresponding author: L. Ciechanowski, e-mail: lciechanowski@kozminski.edu.pl

Introduction

Automated inanimate objects or even human-like beings are present in human civilisation almost since the beginning of our written accounts. Ancient Greek mythology as presented in *Iliad* and *Odyssey* by Homer was filled with automata - self-operating machines - such as self-driving chariots and carts, 'programmable' robots like a bronze giant called Talos, and others [1]. Other ancient cultures are not different in this respect. For instance, robot-like beings were said to guard Buddhist relics, while in Chinese writings from the third century BC Lie Zi describes the engineer Yan Shi, a mechanical engineer, presenting a human-like mechanical creature to the king Mu of Zhou [2].

The above-mentioned historical examples prove that Human-Robot, or in general - Human-Machine Interaction, was a widespread dream, present in people’s imagination long before the production of human-like (ro)bots was technologically possible. Presently, we are witnessing a plethora of academic, commercial, and amateurish chatbots and robots being developed. Therefore, this is the time when we are bound to ask what is the quality of interaction between human and artificial beings, and thus, what kind of bots should we produce to make them more pleasing and interactable for people.

Up to date, researchers have developed a couple of measures to study the Human-Machine Interaction. One of the most famous is the Imitation Game, known under the label of the Turing Test (TT) [3]. A successful passing of the test would mean that a machine became human enough to fool people into thinking that it is in fact a human. There are however various methodological reasons why this test is not a proper one. Due to the limited space we will only mention a few of them. First of all, the outcome of the test depends strongly on the 'quality' of the human judge and the human opponent, and not on the quality of the bot itself. Second, the criteria for passing the TT is dubious - why convincing only a small (or bigger, depending on the version of assumed TT) percentage of human judges is enough to be announced as a human-like bot? Even though TT is still used by e.g. chatbot inventors, there are voices expressing their doubts and calling for abandoning it altogether [4,5]. In that situation, we would need another tool to assess the quality of Human-Machine Interaction. In response to the need, the Uncanny Valley Effect (or Hypothesis) has been proposed.

The UVE was put forward by a Japanese roboticist Masahiro Mori almost half a century ago in the *Energy* journal [6]. The effect comes down to a nonlinear relationship between machines’ perceived human likeness and familiarity with them (or their likability or emotion evoking). UVE graph representation is supposed to be nonlinear, since - as stipulated by Mori - the relation between human likeness and emotions is such. That is, when we increase the similarity of artificial beings to human, we may observe a simultaneous increase in the domain of positive emotions or familiarity with those beings. This increase on the chart is supposed to reach a peak at a certain point and then fall in the 'trench', or 'valley' - so for example a zombie, or a Japanese Kabuki doll are quite high on the human likeness axis but at the same time they usually cause the emotions of eeriness or fear and disgust (see Figure 1).

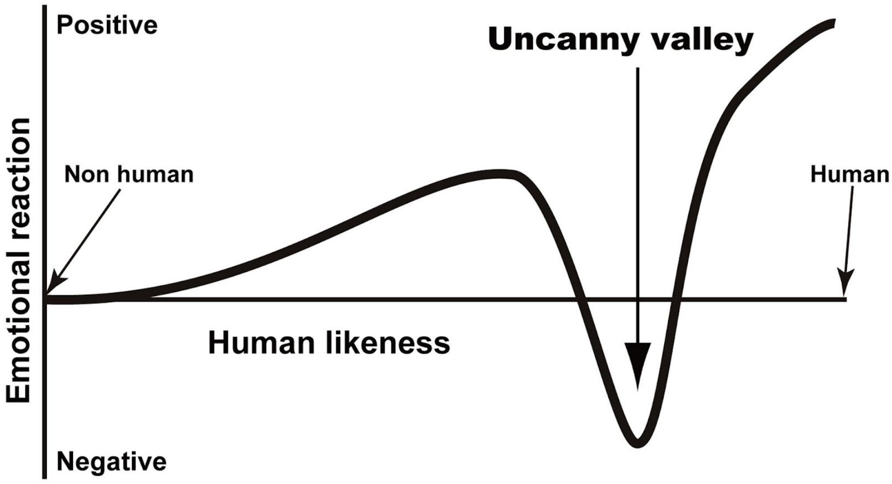

Figure 1: A visual representation of the proposed Uncanny Valley Effect. Adapted from [6].

The stipulated UVE is still debated and researched. There are studies providing support to it, for adults [7–9], children [10], infants [11], and primates [12]. At the same time, there is also experimental data questioning it, and showing that it may not be a valley, but a cliff [13], slope [14] inverted UV [15], uncanny wall [16], or failed to prove UV for movement [17,18]. Moreover, some studies using similar stimuli (morphed images), provided opposing findings [19,20]. However, one of the most detailed and meticulous studies of shape of the UV has proven, that it is indeed a valley-shaped model [21]. Mathur and colleagues have used a plethora of 182 face stimuli - both robotic, human, and border cases - and modeled the results with a polynomial regression, which allows to discern the UV shape. This is an unusual practice, but we believe that only this way is a correct one in investigating the UVE.

Considering the conflicting results in the domain, the need to develop further experimental paradigms is still urgent. In this paper, we are analysing the articles that have used more objective measures than just subjective reports or questionnaires. We are of the belief that developing such methods will provide better results and understanding of the UVE.

In the following subsection we are going to discuss the various types of the UV hypotheses and briefly consider the causes of its existence.

Types of Uncanny Valley hypotheses

The most recent reviews of the UVE [22–25] attempt to enlist and categorise the UV hypotheses. According to them, we can list the following hypotheses and mechanisms:

1. The Naive Hypotheses,
2. The Morbidity Hypothesis / Mortality salience hypothesis,
3. The Pathogen avoidance hypothesis or disease or threat avoidance and mate selection
4. The Movement Hypothesis
5. Categorisation Ambiguity/Difficulty or the category conflict hypothesis

    a. The atypical feature hypothesis
6. Perceptual Mismatch, the Atypical Feature Hypothesis or Conflict Induced Hypothesis
    a. Perceptual Narrowing, Differentiation and Learning
7. Violation of expectations (predictive coding & Bayesian approach)
8. The Dehumanization Hypothesis
9. FunCanny Valley
10. Fluency Amplification Model (contrary to UV)
11. Evolutionary/Developmental Hypothesis
12. The Social Functionalist Approach

In the following passages, we are going to describe these UV hypotheses and mechanisms. In each point, we are providing a brief description of the given hypothesis and then discuss the evolutionary, psychological etc., reasons for its formulation.

1. The Naive Hypotheses

The Naive Hypothesis, according to Kätsyri et al. [22], is the supposition that any experimental manipulation of human likeness would lead to UV in the domain of familiarity or emotional valence, because 'almost humanlike' beings are always more eerie and negative than 'somewhat humanlike' or 'fully humanlike' beings.

2. The Morbidity Hypothesis / Mortality Salience Hypothesis / Terror Management Theory

The Morbidity Hypothesis claims that dead, or undead characters, such as zombies, evoke negative familiarity or emotions than any other type of character. This can be caused by the appearance of some artificial beings which may visually resemble dead individuals, while at the same time being mobile and/or capable of humanlike communication. This may in turn remind us of our own mortality and cause anxiety [20,26].

3. The Pathogen Avoidance Hypothesis, or Disease or Threat Avoidance and Mate Selection

This set of hypotheses focuses on a seeming health risk that humans purportedly recognize in artificial beings. We tend to view robots as carrying transmissible diseases due to their observable imperfections [27]. It is possible that the emotion of disgust - which is naturally connected to our fear of catching diseases - and the individual differences in its intensity, may be a mediator of the UVE [28].

4. The Movement Hypothesis

The Movement Hypothesis states that almost humanlike movement leads to more negative emotions than 'somewhat humanlike' or 'fully humanlike'. Mori himself originally stipulated that motion might deepen the UV [6], and this hypothesis has been proven in some studies [29], but there were also other presenting data against it [17].

5. Categorisation Ambiguity/Difficulty or the Category Conflict Hypothesis

In general and broad terms, UV is caused by an ambiguity of almost-human beings categorisation, or, in other words, by the difficulty in categorising those beings as either humans or robots [30,31]. This type of hypotheses is sometimes rendered as Category Identification Difficulty.

a. Category Identification Difficulty

This hypothesis states that since it is difficult for us to discriminate between characters located near the threshold of human - artificial spectrum, these characters will evoke more negative emotions than the characters from the extremes of the spectrum (human and the most 'robotic' beings) [32]. In other words, the characters on the bottom of the UV (compare Fig. 1) are mostly problematic in terms of their perceptual discrimination. Some research shows that discerning androids from simpler robots and from humans is the most neuronally demanding venture [33].

6. Perceptual Mismatch, the Atypical Feature Hypothesis or Conflict Induced Hypothesis

This family of hypotheses states that the phenomena causing the UV are incongruencies between various sensory cues. For instance, an overtly exaggerated size of human eyes presented as experimental stimuli would probably evoke negative emotions or increase perceived eeriness [31,34].

a. Perceptual Narrowing

This is a more specific version of the Perceptual Mismatch hypothesis. When an artificial character is trying to imitate human behavior/physique, but lacks a human face, we may experience cognitive dissonance, since we are not able to discern feelings and emotions of such beings [35]. Some research suggest that facial features (eyes, mouth etc.) are most crucial in the process of conveying emotions, therefore it is the hardest to render them truthfully and to avoid the sense of eeriness when they are rendered poorly [36,37].

7. Violation of expectations (Bayesian estimation or predictive coding)

This hypothesis states that when we encounter almost humanlike individuals we may experience either a fulfilment or violation of our expectations, based on congruent or incongruent perceptual and cognitive cues [38]. When for example we observe a gait, we have certain expectations concerning each movement. When the consecutive movements are incongruent, these expectations are violated, and we may experience less familiarity with the phenomenon we are observing. The advantage of this theory is that our phenomenal experiences are accompanied with their neuronal counterparts, which may be modelled and experimentally investigated, for example in the form of Event Related Potentials in the electroencephalographic signal.

8. The Dehumanization Hypothesis

Wang and colleagues [23] claim that it is not the attribution of humanlike aspects to a robot that causes the UVE, but the opposite - namely, the dehumanization of a seemingly humanlike being. The hypothesis states that when we attribute humanlike aspects to an artificial individual, we tend to recognise its nonhuman features, thereby dehumanising it. And this dehumanisation process leads to the experience of eeriness or uncanniness.

9. FunCanny Valley

The FunCanny or Fun Uncanny Valley hypothesis was proposed in one paper [39]. This hypothesis states that the category of strangeness correlates most strongly with amusement and not with negative emotions. Strangeness in Mäkäräinen and colleagues' study [39] is proved to correlate positively with fear, disgust, surprise, and amusement, and negatively with sadness, but the correlation coefficient is the highest in the case of amusement.

10. Fluency Amplification Model

Cheetham and colleagues [40] claim that we may observe an effect opposing to the UVE when we take into consideration another factor: processing fluency. They base on the findings of Albrecht and Carbon [41], who coined the Fluency Amplification Model. The model states that increasing processing fluency - for example in threshold cases of discerning between simple robots, more human-like avatars, and humans - does not necessarily increase familiarity or likeability of more, but not perfectly, human-like faces (we describe the detailed methodology of Cheetham and colleagues' study below). On the contrary, increasing processing fluency can lead to increased negative affect in response to non-perfect avatars. This result is also supported by other study by Cheetham and colleagues [32]. The possible explanation of the said effect is that we do not trust the unfamiliar beings. The trust may increase when we realise that the being is nonthreatening. This explication goes along the lines of the effect well-known in psychology: the Mere Exposure Effect [42].

11. Evolutionary/Developmental Hypothesis

Studies of primates [12] and infants [11,43] suggest that the UV may reflect our biological adaptation to alien beings.

12. The Social Functionalist Approach

Olivera-La Rosa [25] proposed to interpret UV as a social or emotional response, stemming from visual properties and ideational qualities of observed artificial objects. This response would prevent us from social interactions with somehow threatening or abnormal beings, which is indicated by uncanny facial features of human-like these beings, supposedly manifesting the 'moral status' of the said entities. The perception of possible psychopathy and dehumanization of uncanny entities entails moral attributions.

Methods

Using the keywords 'uncanny valley', ‘psychophysiology’ and ‘neuroimaging’ in Google Scholar and EBSCO databases we found all articles investigating the UV. Since we were interested only in experimental investigations of the UVE using psychophysiological and neuroimaging methods, we have narrowed down the search to 19 papers, which were fulfilling these criteria.

Results

Eye movements in the UV effect

Herein we will discuss the usage of the eye-tracking systems in studies of various research subjects such as monkeys [12], infants [11,43,45,46] autistic children [47], and healthy adults [40,48,49].

UV effect in monkeys (evolutionary origins of the UV effect)

Researchers investigating the evolutionary origin of the UV, explored the group of *macaca fascicularis* (macaque monkeys) [12]. Subjects were exposed to images of their conspecifics with unrealistic and realistic artificial faces and real macaques faces with either dynamic or static facial expressions (such as 'coo’’, 'scream’’ and 'neutral’’). Researchers measured number of fixations and duration of fixations of five monkeys.

It was reported that monkeys’ duration of fixations was smaller for realistic computer generated faces in comparison to real monkeys and unrealistic faces. It is worth to mention that the effect was stronger for the dynamic in comparison to static faces. Results demonstrated that UV effect is also present in other species, hence its nature may be evolutionary.

UV effect in infants (developmental origins of the UV effect)

Evidence in favor of the UV effect in infants

Furthermore, gaze tracking was widely used in studies of infants in order to investigate the developmental origins of the UV. However, several objections can be identified to the first eye-tracking studies of the UV. First researchers investigating this phenomenon in infants [11,45] used the camera and independent observers’ decoding of gaze which increased the likelihood of obtaining spurious results in comparison to modern eye-tracking systems. Nevertheless, we decided to include the studies in our review because of the evidences they provide in the context of developmental sources of the UV. Eye-tracking was used as the tool of measuring infants preferences in the paradigm known as preferential-looking.

In the study of Lewkowicz and Ghazanfar [11], researchers used the dynamic facial stimuli of human, realistic and unrealistic avatar created by enlarging eyes to 150% in order to induce the uncanniness.

According to the researchers, 6 month old babies are expressing slightly more preferences to the faces with uncanny characteristics. In turn, infants as young as 1 year start to show preferences towards human in comparison to uncanny avatar. Researchers stated, that this effect exists due to the perceptual learning process meaning that the reason behind the UV is more advanced expertise with the conspecifics faces [11]. Facial stimuli are connected to the positive outcomes and, in turn, leads to special status of human faces for our species [11]. Other effects playing role in our expertise acquiring are differentiation [50] and narrowing processes [11,51] leading to the more detailed way of exploration of narrower stimuli characteristics by infants. Other results indicate that there were no differences in a proportion of total looking time among human face in comparison to avatar face without the enlarged eyes in the group of 12 month old infants. Researchers also compared the uncanny avatar with the avatar without uncanny characteristics predicting that the duration of fixations will be different for these two entities. It was shown that infants from 6 to 12 month old exhibited significant preferences towards the avatar without the enlarged eyes. Together, the results of this study confirm the existence of the UV phenomenon as early as in one year old infants and show that atypical facial characteristics might be used to induce uncanniness.

In follow-up studies of infants by Matsuda et al. [43] and Matsuda, Ishiguro and Hiraki (2015) [46] more advanced eye-tracking tools were used to investigate the UV effect. Matsuda and colleagues [43] concentrated on the familiarity and novelty dimensions of the UV measured by duration of fixation on stimuli comprising faces of infant's mother, stranger, as well as the morphed faces with 50% mother's and 50% stranger's face.

It seems that infants had similar preferences towards mothers and strangers faces but spent significantly less time looking at the ambiguous morphs. Developmental origin of the UV was also confirmed in this study showing lack of the UV effect for 7- and 8-month olds and emergence of the UV effect in 9-,10-, 11- and 12-month old babies. Similarly to the Lewkowicz and Ghazanfar [11] study, results were explained by the acquired perpetual expertise, improved ability to distinguish stimulus characteristics and narrowing processes [43]. Importantly, results were also interpreted in terms of conflict-induced uncanny valley account which states that stimuli which fails to satisfy the expectations may evoke the feeling of unease [43].

Evidence against the UV effect in infants

In the study of Shimada et al. [45] researchers manipulated the human-likeness of the stimuli using movies of the android Repliee R1, mechanical-looking humanoid robot Robovie and real human. Age groups investigated in this study were 12 month-, 18 month- and 24-month olds.

As the researchers suggest, the youngest age group expressed more interest towards android in comparison to mechanical-looking robot. Moreover, according to scientists, when children are 18 months old, their gaze pattern changes in direction of preferences for robot, whereas the 24-month-olds are again preferring to look at the android. However, all the reported results were statistically insignificant according to the international scientific standards ($p<.1$ vs $p<.05$). These ambiguous results may lead us to conclusion, that the UV effect is not observed in the infants as young as 24 month old.

In the other study by Matsuda, Ishiguro and Hiraki [46] human, human-like android Repliee Q2, and mechanical robot performing grasping action stimuli were used. The gaze patterns of infants from 6- to 14-month-olds were measured for three areas of interest (AOI) face, goal of action, and body.

Interestingly, no UV effect was reported in this study for all age groups. Infants' gaze was attracted the most towards robot which was explained by the novelty of this stimulus. No difference between human and android were found. Face area attracted the most attention of all infants regardless of the age group which was explained by the general importance and increased attention towards human-like faces [52]. Movement, which is one of the most important aspects of the UV according to [6], was also investigated in this study. It was shown that the movement discrimination is not well developed in infants as young as 1 year old. Additionally, older infants have better ability to predict human actions as shown by the larger looking times in the goal AOI.

UV effect in children from clinical group

It is difficult to underestimate the hope for the social robots in such areas as health care, psychotherapy and education. Scientists attempted to use them to help people with social phobia [53], assist children with autism and dementia patients [54]. Designers of such interventions should particularly take into account the existence of the UV effect. According to our best knowledge, only one study appeared investigating clinical groups in the context of the UV effect using eye-tracking measures.

Feng et al., [47] used the two-alternative forced choice task in their study of Autism Spectrum Disorder (ASD) children in comparison to the healthy children. Researchers measured the proportion of looking time at the eyes, nose, and mouth AOI of morphed pictures of a cartoon and human face with 100%, 125%, 150% enlarged eyes.

Study provided a thorough evidence against the UV effect for ASD children in comparison to typically developing children. Therefore, according to Feng et al. robots designed for ASD children should be created taking into account the goal of the intervention. In contrast, healthy subjects had slightly more preferences towards cartoon faces than human faces. This study demonstrates that healthy children as young as 5 to 7.4 years old exhibit UV effects. One more interesting finding was that ASD group fixated less on pupils and more on the peripheral areas in comparison to non-clinical group. However, the proportion of time spent on eyes was the same for the both groups leading to the conclusion that ASD children have difficulties with mismatch characteristics discrimination [47]. The lack of the UV effect in the ASD children was also explained by the deficits in the creation of the face prototype [47].

UV effect in healthy adults

In the next section we will discuss eye-tracking studies on healthy adults.

MacDorman et al. [48] investigated the social effect of eye contact breaking in two conditions: interaction with human or robot interlocutor which were asking either questions which participants can answer immediately ('know questions') or think questions which required participants to consider the answer ('think questions'). It was hypothesised that in the think condition participants would be breaking the eye contact more often. This effect was described in the 'social signal theory' which suggests that this eye movement occurs in order to inform the interlocutor about the thinking process [55]. In this study the eye gaze direction was measured based on the recorded video clips. In one experiment participants were told that the robot is autonomous, whereas in the second that it is controlled by the researcher.

MacDorman and colleagues [48] hypothesized that the eye contact breaking would occur more often for the avatar in the controlled condition than in autonomous, but findings were contrary to the predictions. However, the research was not methodologically sound. For the human-human interaction no information regarding the sample size was provided, whereas for the human-robot condition the sample size was too small. Moreover, the methods used in this study cannot be equal in precision to the advanced eye-tracking systems.

Next study by Cheetham et al. [56] used the human-morph continua with 13 images different along the dimension of human-likeness. This allowed researchers to explore variability along this dimension which was not present in other eye-tracking studies. Human and avatar images from the ends were considered categorically unambiguous while the avatar in the middle difficult to categorize. Importantly, researchers constructed the categorisation response function which had the sigmoid shape and found the category boundary which was defined as the position of the most ambiguous avatar. Research was aimed to investigate the number and duration of fixation on three AOI, eyes, nose, and mouth, for stimuli with various levels of human likeness and categorical ambiguity. It is worth to mention, that researchers were also expecting gender-dependent differences in eye-movement and RT patterns.

Cheetham and colleagues [56] have reported that the increase in the categorization ambiguity is connected to the more detailed processing of eyes and mouth regions for the ambiguous avatar faces in comparison to the nose area for the easy-to-categorize avatars. The same pattern was shown for the human and ambiguous avatars, but was not statistically significant. One of the most equivocal findings of the study was that RT for avatars was faster than humans, the effect explained by the 'race-feature hypothesis' [57]. These findings, according to the researchers, provide an empirical evidence in favour of the 'course-to-fine strategy' [56] which suggests two different processes underlying the categorisation of the ambiguous and unambiguous stimuli. First process is rapid and more general helping to categorise the avatar stimuli using more primary characteristics such as luminance and shading [56]. In contrast, process used for the ambiguous categorisation is slower and more detailed, therefore categorised by shift in subjects attention towards eyes and mouth [56]. Interestingly, females were faster to make category decisions at the category boundary. However, patterns of eye-movement for both genders were similar excluding one difference in the higher duration of fixation for the eye region in the female group explained by the increased attention towards socially important information [58].

Recent study by Reuten, van Dam & Naber [49] used the measure of pupillary response in the emotion recognition task. Stimuli used in this study were pictures of robots faces and human faces with different expressions (such as happy, neutral, sad, angry, and fearful).

Researchers obtained empirical support in favor of the UV hypothesis**.** Results demonstrated that pupil dilation was less robust for the facial expressions of the uncanny robot while the pattern of reaction was similar for the robot and human emotions. The eye-tracking results were consistent with self-reports in which humanlike robots received higher uncanniness ratings, participants experienced difficulties in their emotions recognition and they scored lower on imagined social interaction scale. Moreover, the results of this study suggest that humans' emotional reactions in human-robot interaction resemble those with the human counterpart, hypothesis known as 'media equation' [59].

Facial muscles activity in the UV effect

Neuroscience methods allowed researchers to explore unorthodox aspects of the UV effect. For example, study by Mäkäräinen et al. [39] has shown that strangeness, associated with UV, can be connected with increase in positive-valenced emotions in comparison to negative-valenced emotions contrary to what was previously considered by the large body of literature and empirical studies. This new hypothesis, named the FunUncanny Valley, was investigated by the self-assessed reports as well as *zygomaticus major and corrugator supercilii* muscles activity. Stimuli in this study were images of happy, angry and surprised facial expressions with varied levels of magnitude and realism.

Researchers have found that strangeness for positive facial expressions was positively correlated with *zygomaticus major* muscle activity and negatively with *corrugator supercilii* activity. However, there were no significant *corrugator supercilii* activity associated with strangeness for presented angry and surprised facial expressions. This was also confirmed by the self-reports showing that the connection between strangeness and amusement was much stronger than strangeness with fear and disgust.

Event related potentials in the UV effect

Study conducted by Schindler et al. [60] concentrated on such event-related components as N170, early posterior negativity (EPN), and late positive potential (LPP).

Male and female faces with six levels of realism and happy, angry and neutral facial expressions were presented to participants.

Importantly, the ‘U-shape’ with regard to the UV effect was reported in N170 component. The amplitude of this component was the strongest for the least realistic and most realistic faces in comparison to average faces and was interacting with the emotional expression. However, there were no statistics provided which will confirm the fact that there were differences between the real and cartoon faces in comparison to medium-stylized. For the most cartoonish faces the effect was stronger in right cuneus and lingual gyrus (occipital face area), whereas real faces led to enhanced amplitudes in inferior and middle occipital areas (fusiform face area) [60]. According to the researchers, fusiform-dependent holistic processing for realistic faces and feature-based processing in occipital face areas for neoteny faces resulted in the ‘U-shape’ modulation of the N170 component [60]. Happy expressions and angry faces were connected with similar N170 amplitudes for abstract stimuli, whereas for the realistic faces the effect was found only for angry expressions. The LPP component was better fitted with the linear shape of the model in which more realistic human faces were linked to the increase in LPP in superior occipital and superior parietal areas. This effect may be explained by the inferences about the identity of the realistic faces [60]. The EPN component was only modulated by the emotional expression being more pronounced on the right side for the angry faces. The main conclusion from this study is that humans process differently faces with various degrees of realism. In spite of the fact that both real and cartoonish faces influence the same stages of processing in the primarily visual areas, only realistic faces induce the necessary processing which can help to identify with the agent [60].

Predictive nature of the UV effect was explored in the study by Urgen et al.[38] in which the researchers were investigating the ‘predictive coding theory’ in the context of ‘appearance-motion incongruence’ [38]. As a stimuli the videos of human, android and mechanical robot were used in static and dynamic modes similarly to Saygin et al. [33]. Congruent stimuli included human with biological motion and robot with artificial motion. Incongruent stimuli was the human-like android with artificial motion. Humanoid robot Repliee Q2 was used for the avatar and robot condition without the surface elements in former.

This study revealed that N400 component was present in the brain activity at frontal areas in all experimental conditions. Human and robot agents elicited the same amplitude of the N400 response irrespective of the mode, whereas in android condition the amplitude of N400 was greater for dynamic in comparison to static mode which, according to the researchers, suggests violation of predictions [33]. The N400 component was generated in the widespread areas including prefrontal areas, superior and middle temporal areas, temporal-parietal junction at its pick in inferior parietal lobule [33]. This area also known as Brodmann area 40 (BA40) is the region in the brain which is known to integrate different sources of information in the perception of movements [61]. According to the researchers, their study confirmed the prediction violation theory as the mechanism responsible for the existence of the UV effect in appearance-motion incongruence conditions [38].

Functional Magnetic Resonance (fMRI)

Relationship between biological or artificial movement and appearance were also explored in the fMRI study of Chaminade, Hodgins & Kawato [62]. Researchers used videos of agents created from movements of actors wearing reflective markers (human-like movement) or key framed motions designed by an animator (artificial movement) [62]. The two-alternative forced-choice paradigm was used in which participants were asked to classify movement as biological or artificial. The behavioural measures derived from the data were the response bias which was defined as a tendency to perceive movement as biological and sensitivity which described the discriminability of the motions by humans' sensory systems [62].

Study reports that increase in the anthropomorphism is linked to the decrease in the bias to characterise the movement as natural. The response bias was found to be correlated with an increase of activity in mentalizing network in the left superior temporal sulcus (STS), temporoparietal junction (TPJ), right superior temporal gyrus (STG), anterior cingulate cortex (ACC) and precuneus (PreC) [62]. In turn, the decrease of activity was found in regions belonging to the mirror neuron network [63] in the right inferior frontal and intraparietal sulcus as well as right ventral premotor and posterior parietal cortices, which both contribute to controlling allocation of attention [64,65]. As researchers report, those findings confirm that anthropomorphism facilitates motor resonance [66], which elicit the knowledge of action while observing the action performed by other agent [62]. In contrast, the process underlying the biological movement was considered to be mentalizing which is a tendency to attribute the intentions to the agent [62]. Interestingly, there were no significant activations in the areas responsible for emotional processing, such as amygdala, insular and orbitofrontal cortices [67]. Hence, the study undermines the literal interpretation of the UV stating that anthropomorphic agents unsuccessfully trying to replicate human appearance and movement would elicit negative emotions [62]. Then the underlying mechanism of response bias is perceiving the action as atypical in the absence of the emotional response [62]. Importantly, there were no effect of characters on brain responses which may be due to the individual differences and experience of subjects. However, there was no measure of it included in the study.

Cheetham, Suter and Jäncke (2011) [68] published a study where they used fMRI in order to investigate the effect of categorical perception [69]. Categorical perception states that underlying mechanism of UV effect is ambiguity in object classification to human or non-human category at the categorical boundary irrespective of differences in human-like appearance [70]. The stimuli constituted 32 human-avatar continua of male face with M0 (avatar endpoint) to M12 (human endpoint) different along the dimension of human-likeness (similarly to eye-tracking experiment of Cheetham et al., 2013 [56]). The first behavioural experiment allowed researchers to fit the classification response function in order to indicate the avatar–human category boundary and to select the morphs to present in the following behavioural and fMRI studies (M0, M4, M8, and M12 were selected).

Behavioural results has shown that there were faster responses for morphs at the endpoints (M0 and M12) than at category boundary (M4 and M8 - no difference between them). Interestingly, it was also found that there is a variation in human-likeness within the human category which implies a need of cautiousness with treating human as the reference point in the research design and necessity of non-human–human category boundary definition in UV effect research. The perceptual discrimination task in a same-different paradigm was conducted where participants had to answer whether the presented stimuli from the pair were both the same or different in appearance. Trials were divided into avatar and human depending on what object was presented first. There were three different conditions: within (the faces of a pair were from the same category) same (identical) and between (opposite sides of the category boundary). Faces drawn from within a category were more difficult to distinguish than those from the category boundary, which confirmed the categorical perception effect.

Subsequent fMRI study in the pair-repetition paradigm with prime and target investigated neural mechanisms linked to two distinct human-likeness aspects one of which is physical similarity as defined by Mori and the other one categorical perception which was confirmed in previous behavioural studies. In fact, the different brain areas were found to process differences in physical features in comparison to category change across the boundary. Physical change in human trials was connected with activity in the right and left mid-fusiform areas, whereas in avatar trials with right mid-fusiform area usually linked to physical and texture similarity of faces [68]. Consequently, a change of category of prime and target for avatar targets was linked to the dorsal and ventral striatum (putamen and head of caudate), thalamus, and red nucleus responses whereas for human target to hippocampus, entorhinal, and perirhinal areas, mid-insula, and amygdala responses [68]. According to the researchers, this findings may indicate that avatar and human stimuli belong to different categorisation problems which require different strategies to solve them [68] .

Study which we discuss next explored perception of human-human and human-robot dyadic interactions [71]. The stimuli in this experiment were realistic images demonstrating social interactions of human with either human or robot. This study explored wide range of characteristics of robot counterpart underlying the UV effect such as robots' eeriness, believability, intelligence and capability of emotions [71]. Subjects performed interaction categorization task in the fMRI scanner. Researchers concentrated on two networks: person perception network (PPN) which includes the occipital face area (OFA), extrastriate body area (EBA), fusiform face area (FFA), fusiform body area (FBA) and the posterior superior temporal sulcus (pSTS) [72] and the mentalizing network (MTN) including ventral and dorsal medial prefrontal cortex (VMPFC, DMPFC), anterior temporal lobe (aTL), temporal–parietal junction (TPJ) and the precuneus (PrC) [73], which were previously reported to be engaged in the processing of social interactions. PPN was reported to be engaged in processing of appearance, whereas the MTN was linked to the extraction of agent's intentions, beliefs and feelings linked to the behavior [71]. Study appeared as a response to the ongoing debate in literature regarding the engagement of these areas in perception of human-robot interactions and how it is linked to the UV phenomenon.

Researchers report that interaction with robot in comparison to human was rated as ‘eerier’ and less ‘believable’ [71]. Moreover, robot partners were considered as less ‘capable of emotions’ and less ‘intelligent’ [71]. Robots were more often perceived as performing actions of helping humans. Eeriness was negatively correlated with emotional capacity and intelligence, however only for humans’ interaction [71]. According to the behavioural results, there is no link between eeriness and mind attribution in the human-robot condition which allows to conclude that there are distinct strategies and factors in the perception of human-human and human-robot interactions [71].

fMRI results suggested that MTN was involved in processing in both conditions [71]. Human-human interaction was linked to increase activation in the right FFA, bilateral pSTS of the PPN as well as left TPJ, area of MTN responsible for ‘situation-specific mental state attributions’ [71]. Whereas human-robot interactions involved the increase of the activity in the middle occipital gyrus and the inferior temporal gyrus (areas outside the PPN) as well as PrC and VMPFC areas in the MTN later considered to be responsible for ‘for script-based social reasoning’ [71]. Increase in the activity in PrC was linked to the enhanced ratings on ‘believability’ dimension and right amygdala, right insula and the left STS were linked to higher evaluated robot’s emotional capacity [71]. Higher eeriness was connected to the greater activity in VMPFC in human-robot condition, which was suggested to be the neural correlate of the UV effect [71].

Combination of methods

Evidence in favor of the UV effect

Combination of methods can be extremely helpful taking into account the ambiguous findings in the UV domain.

Recent study by Ciechanowski et al. [74] provided important evidence in favour of the UV effect. Researchers used EMG, RSP, ECG and EDA measures in two groups of subjects: one interacting with simple text and the second one with the animated chatbot.

It was found that group interacting with animated avatar was characterised by the more intensive arousal (significant differences in corrugator supercilii muscle activity, EDA and HR) leading them to conclusion that complex avatar was in general perceived as more uncanny. This was also confirmed by the self-reports showing that subjects felt more negative emotions towards realistic avatar chatbot. Moreover, simple chatbot was perceived as having less negative traits, significantly 'less weird' and less inhuman than more human-like avatar.

Next study which we will discuss was conducted by Saygin et al. [33] in the similar design as the recent EEG study of Urgen et al. [38]. The videos of human, android and mechanical robot performing actions were used. Congruent stimuli included human with biological motion and robot with artificial motion. Incongruent stimuli was the human-like android with artificial motion. Humanoid robot Repliee Q2 was used for the avatar and robot condition without the surface elements in the former. fMRI repetition suppression measures in combination with eye-tracking were used. Repetition suppression reflects processing of the stimuli rather than task demands [75]. Study investigated the response of action perception system (APS) to appearance and motion in congruent and incongruent stimuli. The APS, which overlaps a mirror-neuron system (MNS) [66], consists of temporal, parietal and frontal regions (in particular lateral temporal, inferior frontal or ventral premotor and anterior intraparietal cortex) and is linked to the motor resonance theory of MNS [66].

None of the eye-tracking measures were significant with regard to three conditions. All stimuli induced repetition suppression in lateral temporal cortex [33]. Congruent human and robot stimuli were connected with similar repetition suppression whereas android condition lead to the repetition suppression in the left and right anterior intraparietal sulcus (aIPS) as well as posterior and superior parietal region (sIPS) in the left hemisphere [33]. The aIPS is the area in brain which integrates visual and motor information in the 'action perception system' [76] and according to the researchers plays the main role in prediction generation [33]. However, they suggested the need of the further investigation using EEG or MEG in order to better understand this process. In turn, suppression for human and avatar condition was greater in left lateral temporal cortex (extrastriate body area), which was linked to 'form-based processing of action' [77]. Consequently, the research revealed that robot condition induced greater response in visual cortex bilaterally linked to processing of 'low-level visual differences' [33]. Dissimilar suppression patterns in brain activity for android condition were explained by the 'predictive coding framework' according to which stimuli which violates predictions will be linked to greater brain activity [78]. Researchers claimed that process which drives the UV effect is the violation of prediction.

Evidence against the UV effect

In contrast to study of Saygin et al. [33], study by Cheetham et al. [40] undermined these crucial aspects of the UV effect. Similarly to a study by Saygin et al., researchers explored one of the main UV predictions stating that ambiguous objects, which are difficult to categorise into human or non-human category, will evoke negative feelings in comparison to explicitly non-human or human objects [40]. Researches concentrated on such dimensions as valence, arousal and familiarity of categorically ambiguous faces in comparison to human and non-human faces and measured *corrugator supercilii muscle* activations and Late Positive Potential (LPP) [40]. The creation of human-morph continua was similar to Cheetham et al. [56].

It seems that ambiguous stimuli are not unequivocally leading to feeling of strangeness contrary to what was previously considered. There were no differences in *corrugator supercilii* reaction between ambiguous avatar, unambiguous avatar and human stimuli. Furthermore, arousal and negative valence indicated by LPP and self-reports were generally increasing with the distance from the human end of the continua. According to the researchers this effect is linked to cautiousness towards any new stimuli [79]. Analysis of reaction time (RT) data in this study has shown that participants had shorter RT for avatar faces due to their novelty in comparison to human faces. Effects were described in the framework of 'Fluency Amplification Model' [41] which states that 'higher processing fluency of negative stimuli (shorter RT for avatar stimuli) can lead to higher negative evaluation' in contrast to the predictions proposed by the UV hypothesis [40].

Discussion

A plethora of research was conducted trying to explore the equivocal uncanny valley effect. Due to the lack of consistency and significant limitations of the behavioural research, recent studies applied neuroscience methods in order to dig deeper into the underlying correlates of the UV effect in humans' and monkeys' nervous systems.

Methods measuring central nervous system activity used in the UV effect research were electroencephalography (EEG) and event-related potentials (ERPs) as well as functional magnetic resonance imaging (fMRI).

ERPs reflect synchronized electrical activity of populations of cortical pyramidal cells recorded from the scalp in response to the particular event [80]. Components investigated in terms of the UV effect were N170, N400, early posterior negativity (EPN), and late positive potential (LPP). N170 component in previous studies was linked to face processing with larger amplitude for human faces processing in comparison to other face-like stimuli [60]. The N400 is considered to be the biomarker of meaningful linguistic and visual stimuli processing with the greater amplitude for stimuli violating the predictions [81]. In turn, EPN and LPP ERPs were found to be more pronounced for emotional facial expressions and connected to higher attractiveness of facial stimuli [60]. The difference between the two is that the EPN engages the attention on the initial stages of processing, whereas LPP engages episodic memory on the higher stages of processing [60]. Interestingly, LPP is also linked to mentalizing process and it was found that its amplitude is higher in response to human in comparison to puppet faces, as well as realistic computer generated avatars with biographical information in comparison to avatars without identity [60]. Due to its excellent temporal resolution, EEG provides researchers with important insights into the time course of the response. However, its main disadvantage includes difficulties with the source localization due to the poor spatial resolution.

On the other hand, fMRI allows to collect data with better spatial precision but has disadvantages in terms of temporal resolution. fMRI also allows to access the deep subcortical areas which is difficult to obtain with the EEG method [80]. The rationale of Blood Oxygen Level Dependent (BOLD) method is that different brain states are linked to regional changes in the brain's oxygen utilization, the process which relies on the ratio between Oxyhemoglobin and Deoxyhemoglobin in the blood across the brain areas [80]. In turn, in the study of Saygin et al. [33] researchers used repetition suppression method to investigate the neural response to processing of repeated stimuli, which, according to Xu et al. [75], is reduced as the effect of adaptation. This, according to Saygin et al. [33], allowed to draw conclusions regarding the differences in processing of the stimuli in comparison to other low-level visual differences which may influence the BOLD response.

The investigators of the UV effect mainly concentrated on the three brain networks: mentalizing network (MTN), mirror neuron system (MNS) and person perception network

(PPN). Mentalizing network includes such brain areas as ventral medial prefrontal cortex (VMPFC), dorsal medial prefrontal cortex (DMPFC), temporal–parietal junction (TPJ), precuneus (PrC) and anterior temporal lobe (aTL) [73]. This network has been found to play a role in processing of interlocutor's intentions, beliefs and feelings [82]. In turn, person perception network (PPN) includes the occipital face area (OFA), extrastriate body area (EBA), fusiform face area (FFA), fusiform body area (FBA) and the posterior superior temporal sulcus (pSTS) [72]. As it was previously reported, these areas are mainly engaged in processing of agent's appearance [82]. Both PPN and MTN are engaged into the processing of social interactions [71]. The mirror neuron system recruits regions in the inferior parietal areas, the dorsal premotor cortex and the medial frontal cortex (MFC), in particular the pre-supplementary motor (pre-SMA) and anterior cingulate cortex (for more detailed review see: [83]). Researchers also claim that neurons with such properties are located in the ventrolateral prefrontal cortex (VLPF), primary motor cortex and send projections to basal ganglia (BG) [83]. When subject observes agent's performance, the motor resonance mechanism [66] induces the knowledge of action. The action perceptions system overlaps the mirror neuron network in such areas as inferior frontal, lateral temporal, ventral premotor and anterior intraparietal [33].

Other psychophysiological methods which found their application in the preceding UV research are eye-tracking (ET), facial electromyography (fEMG), electrodermal activity or galvanic skin reaction (EDA/GSR), respiratory rate recording (RSP), and electrocardiography (ECG).

Eye-tracking (ET) was the method used most often in research of the UV effect. However, the fact that gaze was recorded using a camera (and not a professional ET device) and that independent observers' were decoding the gaze manually increased the likelihood of obtaining spurious results in the first UV studies [11,45,48]. In contrast, modern state-of-art eye-tracking devices allow to investigate eye movements with notably better precision.

In general, eye-tracking technique is based on measuring the relation between the central point of pupil and reflection of the light from infrared spectrum from the cornea [84]. The most popular metrics measured in the UV studies were number of fixations and duration of fixations. Fixation can be defined as 'maintenance of the point of gaze at a single location in space for a specified duration' [84]. Some studies were also concentrated on times of looking at particular areas of interest (AOI), for example face, goal of action, and body AOI was in focus of the infants study of Matsuda, Ishiguro and Hiraki [46], which has shown the importance of the human-like facial stimuli for humans, but did not confirm the existence of the UV effect. In turn, Feng et al. [47] provided evidence that autistic children are concentrating less on eyes and did not exhibit UV effect in comparison to a typically developing group. Cheetham et al. [56] have reported that the increase in categorization ambiguity is connected to the more detailed processing of eyes and mouth regions for ambiguous avatar faces in comparison to the nose area of the easy-to-categorize avatars.

Pupillometry metrics which reflects the pupillary constrictions and dilation in response to stimuli is considered to be susceptive to various emotions [80]. Pupil dilation has been investigated in the study of Reuten, van Dam & Naber [49] which allowed researchers to observe the reaction to human in comparison to robotic facial stimuli, in the emotion recognition task. Results of this study confirmed that humans experience difficulties in the robot emotions recognition task, which has been accompanied by the less robust pupil dilation in response to an uncanny robot in comparison to human and robot stimuli.

Eye-tracking studies provided conflicting evidence regarding the existence of the UV effect in various research groups. On the one hand, it confirmed the Evolutionary Hypothesis of the UV in the group of macaque monkeys [12]. In contrast, more equivocal findings were revealed in case of the Developmental Hypothesis. On the other hand, some studies confirmed this effect in the group of infants with various age [11,43], but there were also some critical voices [45,46]. Cheetham et al. [56] study on adults revealed that the increase in the categorization conflict is connected to the more detailed processing of eyes and mouth regions in case of the avatar faces falling in the categorical boundary in comparison to the nose area for the easy-to-categorize avatars. However, the effect for the human and for ambiguous avatars did not reach significance. Importantly, eye-tracking was used in order to study the clinical group of ASD children [47]. However, more research is needed in order to thoroughly examine this effect in patients with various disorders which will allow to create more effective robots for health care, psychotherapy and education.

Electrodermal activity (EDA) measure is sensitive to increases in skin conductance (decrease in the skin resistance) which is measured by two electrodes on the surface of skin [80]. It reflects the autonomic nervous system arousal under the influence of various states or processes. However, it was used only in one study by Ciechanowski et al. [74] in which researchers report that there were notable differences in the reaction of EDA between the group interacting with the simple and complex avatar (significantly higher arousal for the latter). EDA is a promising measure which is informative in investigation of significance, intensity and attention towards the stimulus [80]. However, its is impossible to derive decisive conclusions due to the scarcity of the scientific evidence in favor of the increased arousal in response to the UV effect. Future reports should include EDA to investigate its validity as a measure of the UV effect.

Respiratory rate recording (RSP) known as an index of the lung functions is altering under the influence of emotions, physical effort, stress and various dysfunctions. Different systems in our organisms rely on the proper functioning of the respiration which is known to influence the air for production of words as well as playing an important role in cognitive and muscle activity. In turn, electrocardiography (ECG) allows to measure functioning of the cardiac system. Both measures were also applied in one study of Ciechanowski et al. [74], where the heart rate (HR) index was found to be significantly higher for the uncanny stimuli.

Electromyography allows researchers to explore functioning of the skeletomotor system using the surface electrodes. Distinct patterns of muscles activations can discriminate between stimuli as well as individual differences. Two most widely used muscle activities measured in the context of the UV effect are *corrugator supercilii* and *zygomaticus major* due to their inevitable role in the generation of the emotional reactions. *Corrugator supercilii*, a muscle which is responsible for frowning, in the previous reports was linked to the negative-valenced emotions [85]. In turn, *zygomaticus major* muscle was found to be responsible for a generation of smile [85]. In the study by Ciechanowski et al. [74], the activity of the *corrugator supercilii* was facilitated by the uncanny avatar indicating the negative response to this stimuli.

However, the discussion regarding the role of emotions in the UV effect is still ongoing. Olivera-La Rosa [25] trying to define the UV effect stated that 'At a basic (core affect) level, the conceptualization of the [Uncanny Valley Effect] as an unpleasant feeling seems indisputable […]'. However, an EMG study [39] has shown that this hunch may be ungrounded. It revealed that the feeling of strangeness in response to the uncanny avatars, which was previously considered a reaction to the negative stimuli, is more closely connected to the amusement than fear or disgust. Although, there is still scarce body of literature regarding the emotions of fear or disgust in the UV effect, even in spite of the fact that these emotions are core in some UV hypotheses (for example Pathogen avoidance hypothesis). MacDorman and Entezari [28] also suggest that individual sensitivity to disgust may play a role in the UV effect, even though more investigation should be done in this context. Besides, there are some significant gaps in the EMG studies which bypassed such matters as empathy and facial mimicry in response to the uncanny stimuli. Due to the issues in the operationalization of this term, larger patterns of muscle activity in combination with autonomic system measures should be pursued.

Moreover, it is important to investigate individual differences and demographical characteristics which can possibly influence the reaction to the uncanniness. Studies taking such differences into account are scarce. STAI questionnaire measuring anxiety was used in the study of [40]. However, no significant results were mentioned in this study, regarding the relations between anxiety and uncanniness. Other study took gender of participants into account [56]. Researchers observed that gender may influence the speed of making decisions regarding the ambiguous stimuli near the categorical boundary, with females being faster than males in this task. However, in this study only one difference between two groups was found, namely, longer fixation duration for the face area of interest in the group of females. Studies of infants, in turn, took into account the age of participants. More studies should be carried out investigating the possible moderation and mediation effects of individual differences and demography in context of the UV.

Findings discussed in the present review are also important in the context of the widespread usage of the computer generated stimuli in human studies. According to some researchers [34], behavioral and psychophysiological metrics suggest that humans react to artificial entities similarly as to their conspecifics. However, results presented in the present review suggest that researchers should be cautious and take the uncanniness as a confound variable into the account, when trying to directly transfer findings from the laboratory to real-life situations.

Important directions of future UV studies can be developed with the use of computer models. An example of such study is Chattopadhyay and MacDorman [36] which, with the help of Bayesian model of categorical perception, revealed the important role of perceptual narrowing in the UV effect.

Evidence regarding the neural correlates of the UV effect and mechanisms underlying its emergence and further development are still scarce. More studies using psychophysiological tools should be done in attempt to operationalize this equivocal term and explore individual differences which may predict the existence and strength of the phenomenon. Tremendous scientific influence of Masahiro Mori intuitions and previous studies are beyond doubt. However, more replication attempts should be done in order to understand the intricate nuances of the UV effect.

References

1. Mayor A. Gods and Robots: Myths, Machines, and Ancient Dreams of Technology. Princeton University Press; 2018.
2. Needham J. Science and Civilisation in China, Vol. 2, History of Scientific Thought. Cambridge University Press; 1991.
3. Turing AM. Computing machinery and intelligence (1950). The Essential Turing: The Ideas that Gave Birth to the. books.google.com; 2004; Available: https://www.google.com/books?hl=en&lr=&id=dSUTDAAAQBAJ&oi=fnd&pg=PA433&dq=Turing+Computing+Machinery+and++Intelligence&ots=IdE3UMvz7-&sig=W9-oyp5GTsFFOgnpT1wGv6Bunc0
4. Hayes P, Ford K. Turing test considered harmful. IJCAI . researchgate.net; 1995; Available: https://www.researchgate.net/profile/Kenneth_Ford/publication/220813820_Turing_Test_Considered_Harmful/links/09e4150d1dc67df32c000000.pdf
5. Moor J. The Turing Test: The Elusive Standard of Artificial Intelligence. Springer Science & Business Media; 2003.
6. Mori M. Bukimi no tani [the uncanny valley]. Energy. 1970;7: 33–35.
7. Mathur MB, Reichling DB. Navigating a social world with robot partners: A quantitative cartography of the Uncanny Valley. Cognition. 2016;146: 22–32.
8. Burleigh TJ, Schoenherr JR, Lacroix GL. Does the uncanny valley exist? An empirical test of the relationship between eeriness and the human likeness of digitally created faces. Comput Human Behav. 2013;29: 759–771.
9. Skjuve M, Haugstveit IM, Følstad A, Brandtzaeg PB. Help! Is my chatbot falling into the uncanny valley? An empirical study of user experience in human-chatbot interaction. 2019; doi:10.17011/ht/urn.201902201607
10. Tinwell A, Sloan RJS. Children's perception of uncanny human-like virtual characters. Comput Human Behav. Elsevier; 2014;36: 286–296.
11. Lewkowicz DJ, Ghazanfar AA. The development of the uncanny valley in infants. Dev Psychobiol. 2012;54: 124–132.
12. Steckenfinger SA, Ghazanfar AA. Monkey visual behavior falls into the uncanny valley. Proc Natl Acad Sci U S A. 2009;106: 18362–18366.
13. Bartneck C, Kanda T, Ishiguro H, Hagita N. Is The Uncanny Valley An Uncanny Cliff? [Internet]. RO-MAN 2007 - The 16th IEEE International Symposium on Robot and Human Interactive Communication. 2007. doi:10.1109/roman.2007.4415111
14. Kätsyri J, de Gelder B. Uncanny Slope instead of an Uncanny Valley: Testing the Uncanny Valley Hypothesis in Painted, Computer-Rendered, and Human Faces [Internet]. doi:10.31234/osf.io/u7gq9
15. Hanson D. Expanding the aesthetic possibilities for humanoid robots. IEEE-RAS international conference on humanoid robots. Citeseer; 2005. pp. 24–31.

16. Tinwell A, Grimshaw M, Williams A. The uncanny wall. International journal of arts and technology. Inderscience Publishers Ltd; 2011;4: 326–341.
17. Piwek L, McKay LS, Pollick FE. Empirical evaluation of the uncanny valley hypothesis fails to confirm the predicted effect of motion. Cognition. 2014;130: 271–277.
18. Thompson JC, Trafton JG, McKnight P. The perception of humanness from the movements of synthetic agents. Perception. 2011;40: 695–704.
19. Seyama J 'ichiro, Nagayama RS. The Uncanny Valley: Effect of Realism on the Impression of Artificial Human Faces. Presence: Teleoperators and Virtual Environments. MIT Press; 2007;16: 337–351.
20. MacDorman KF, Ishiguro H. The uncanny advantage of using androids in cognitive science research. Interact Stud. 2006;7: 297–337.
21. Mathur MB, Reichling D, Lunardini F, Geminiani A, Antonietti A, Ruijten P, et al. Uncanny but not confusing: Multisite study of perceptual category confusion in the Uncanny Valley [Internet]. Open Science Framework. 2019. doi:10.31219/osf.io/89sf4
22. Kätsyri J, Förger K, Mäkäräinen M, Takala T. A review of empirical evidence on different uncanny valley hypotheses: support for perceptual mismatch as one road to the valley of eeriness. Front Psychol. 2015;6: 390.
23. Wang S, Lilienfeld SO, Rochat P. The Uncanny Valley: Existence and Explanations. Rev Gen Psychol. SAGE Publications; 2015;19: 393–407.
24. Lay S, Brace N, Pike G, Pollick F. Circling Around the Uncanny Valley: Design Principles for Research Into the Relation Between Human Likeness and Eeriness. Iperception. 2016;7: 2041669516681309.
25. Olivera-La Rosa A. Wrong outside, wrong inside: A social functionalist approach to the uncanny feeling. New Ideas Psychol. 2018;50: 38–47.
26. Ho C, MacDorman KF, Pramono ZAD. Human emotion and the uncanny valley: A GLM, MDS, and Isomap analysis of robot video ratings. 2008 3rd ACM/IEEE International Conference on Human-Robot Interaction (HRI). 2008. pp. 169–176.
27. MacDorman KF, Green RD, Ho C-C, Koch CT. Too real for comfort? Uncanny responses to computer generated faces. Comput Human Behav. 2009;25: 695–710.
28. MacDorman KF, Entezari SO. Individual differences predict sensitivity to the uncanny valley. Interact Stud. John Benjamins; 2015;16: 141–172.
29. White G, McKay L, Pollick F. Motion and the uncanny valley. J Vis. The Association for Research in Vision and Ophthalmology; 2007;7: 477–477.
30. Ramey CH. The uncanny valley of similarities concerning abortion, baldness, heaps of sand, and humanlike robots. Proceedings of views of the uncanny valley workshop: IEEE-RAS international conference on humanoid robots. 2005. pp. 8–13.
31. Pollick FE. In Search of the Uncanny Valley. User Centric Media. Springer Berlin Heidelberg; 2010. pp. 69–78.
32. Cheetham M, Suter P, Jancke L. Perceptual discrimination difficulty and familiarity in the uncanny valley: more like a "Happy Valley." Front Psychol. Frontiers; 2014;5: 1219.
33. Saygin AP, Chaminade T, Ishiguro H, Driver J, Frith C. The thing that should not be: predictive coding and the uncanny valley in perceiving human and humanoid robot actions. Soc Cogn Affect Neurosci. 2012;7: 413–422.
34. de Borst AW, de Gelder B. Is it the real deal? Perception of virtual characters versus humans: an affective cognitive neuroscience perspective. Front Psychol. 2015;6: 576.
35. MacDorman KF, Chattopadhyay D. Reducing consistency in human realism increases the uncanny valley effect; increasing category uncertainty does not. Cognition. 2016;146: 190–205.
36. Chattopadhyay D, MacDorman KF. Familiar faces rendered strange: Why inconsistent realism drives characters into the uncanny valley. J Vis. 2016;16: 7.
37. Watt MC, Maitland RA, Gallagher CE. A case of the "heeby jeebies": An examination of intuitive judgements of "creepiness." Canadian Journal of Behavioural Science/Revue

canadienne des sciences du comportement. Educational Publishing Foundation; 2017;49: 58.
38. Urgen BA, Kutas M, Saygin AP. Uncanny valley as a window into predictive processing in the social brain. Neuropsychologia. 2018;114: 181–185.
39. Mäkäräinen M, Kätsyri J, Förger K, Takala T. The funcanny valley: a study of positive emotional reactions to strangeness. ResearchGate. 2015. doi:10.1145/2818187.2818292
40. Cheetham M, Wu L, Pauli P, Jancke L. Arousal, valence, and the uncanny valley: psychophysiological and self-report findings. Front Psychol. 2015;6: 981.
41. Albrecht S, Carbon C-C. The Fluency Amplification Model: fluent stimuli show more intense but not evidently more positive evaluations. Acta Psychol . 2014;148: 195–203.
42. Zajonc RB. Attitudinal effects of mere exposure. J Pers Soc Psychol. American Psychological Association; 1968;9: 1.
43. Matsuda Y-T, Okamoto Y, Ida M, Okanoya K, Myowa-Yamakoshi M. Infants prefer the faces of strangers or mothers to morphed faces: an uncanny valley between social novelty and familiarity. Biol Lett. 2012;8: 725–728.
44. Tinwell A, Grimshaw-Aagaard MN, Williams A. Uncanny Speech. 2011; doi:10.4018/978-1-61692-828-5.ch011
45. Shimada M, Minato T, Itakura S, Ishiguro H. Uncanny Valley of Androids and Its Lateral Inhibition Hypothesis. RO-MAN 2007 - The 16th IEEE International Symposium on Robot and Human Interactive Communication. 2007. pp. 374–379.
46. Matsuda G, Ishiguro H, Hiraki K. Infant discrimination of humanoid robots. Front Psychol. 2015;6: 1397.
47. Feng S, Wang X, Wang Q, Fang J, Wu Y, Yi L, et al. The uncanny valley effect in typically developing children and its absence in children with autism spectrum disorders. PLoS One. 2018;13: e0206343.
48. MacDorman KF, Minato T, Shimada M. Assessing human likeness by eye contact in an android testbed. Proceedings of the. psy.herts.ac.uk; 2005; Available: http://www.psy.herts.ac.uk/pub/SJCowley/docs/humanlikeness.pdf
49. Reuten A, van Dam M, Naber M. Pupillary Responses to Robotic and Human Emotions: The Uncanny Valley and Media Equation Confirmed. Front Psychol. 2018;9: 774.
50. Gottlieb G. Experiential canalization of behavioral development: Theory. Dev Psychol. psycnet.apa.org; 1991; Available: https://psycnet.apa.org/journals/dev/27/1/4/
51. Werker JF, Tees RC. Speech perception as a window for understanding plasticity and commitment in language systems of the brain. Dev Psychobiol. 2005;46: 233–251.
52. Csibra G, Gergely G. Natural pedagogy. Trends Cogn Sci. 2009;13: 148–153.
53. Klinger E, Bouchard S, Légeron P, Roy S, Lauer F, Chemin I, et al. Virtual reality therapy versus cognitive behavior therapy for social phobia: a preliminary controlled study. Cyberpsychol Behav. 2005;8: 76–88.
54. Cao H, Van de Perre G, Kennedy J, Senft E, Esteban PG, De Beir A, et al. A personalized and platform-independent behavior control system for social robots in therapy: development and applications. IEEE Transactions on Cognitive and Developmental Systems. 2018; 1–1.
55. McCarthy A. Eye movements as social signals during thinking: Age differences. for Research in Child Development, Tampa, FL, 2003. ci.nii.ac.jp; 2003; Available: https://ci.nii.ac.jp/naid/10018821190/
56. Cheetham M, Pavlovic I, Jordan N, Suter P, Jancke L. Category Processing and the human likeness dimension of the Uncanny Valley Hypothesis: Eye-Tracking Data. Front Psychol. 2013;4: 108.
57. Levin DT. Race as a visual feature: using visual search and perceptual discrimination tasks to understand face categories and the cross-race recognition deficit. J Exp Psychol Gen. 2000;129: 559–574.
58. Saether L, Van Belle W, Laeng B, Brennen T, Øvervoll M. Anchoring gaze when categorizing faces' sex: evidence from eye-tracking data. Vision Res. 2009;49: 2870–2880.

59. Nass C, Moon Y. Machines and Mindlessness: Social Responses to Computers. J Social Isssues. 2000;56: 81–103.
60. Schindler S, Zell E, Botsch M, Kissler J. Differential effects of face-realism and emotion on event-related brain potentials and their implications for the uncanny valley theory. Sci Rep. 2017;7: 45003.
61. Amoruso L, Gelormini C, Aboitiz F, Alvarez González M, Manes F, Cardona JF, et al. N400 ERPs for actions: building meaning in context. Front Hum Neurosci. 2013;7: 57.
62. Chaminade T, Hodgins J, Kawato M. Anthropomorphism influences perception of computer-animated characters' actions [Internet]. Social Cognitive and Affective Neuroscience. 2007. pp. 206–216. doi:10.1093/scan/nsm017
63. Rizzolatti G, Craighero L. The mirror-neuron system. Annu Rev Neurosci. 2004;27: 169–192.
64. Büchel C, Josephs O, Rees G, Turner R, Frith CD, Friston KJ. The functional anatomy of attention to visual motion. A functional MRI study. Brain. 1998;121 (Pt 7): 1281–1294.
65. Rees G, Frith CD, Lavie N. Modulating irrelevant motion perception by varying attentional load in an unrelated task. Science. 1997;278: 1616–1619.
66. Rizzolatti G, Fogassi L, Gallese V. Neurophysiological mechanisms underlying the understanding and imitation of action. Nat Rev Neurosci. 2001;2: 661–670.
67. Davidson RJ, Irwin W. The functional neuroanatomy of emotion and affective style. Trends Cogn Sci. 1999;3: 11–21.
68. Cheetham M, Suter P, Jäncke L. The human likeness dimension of the "uncanny valley hypothesis": behavioral and functional MRI findings. Front Hum Neurosci. 2011;5: 126.
69. Harnad S. Psychophysical and cognitive aspects of categorical perception: A critical overview. Harnad S, editor. Categorical Perception: The Groundwork of Cognition. Cambridge University Press; 1987; 1–52.
70. Pastore RE. Categorical perception: Some psychophysical models. psycnet.apa.org; 1987; Available: https://psycnet.apa.org/record/1987-98256-001
71. Wang Y, Quadflieg S. In our own image? Emotional and neural processing differences when observing human-human vs human-robot interactions. Soc Cogn Affect Neurosci. 2015;10: 1515–1524.
72. Weiner KS, Grill-Spector K. Sparsely-distributed organization of face and limb activations in human ventral temporal cortex. Neuroimage. 2010;52: 1559–1573.
73. Iacoboni M, Lieberman MD, Knowlton BJ, Molnar-Szakacs I, Moritz M, Throop CJ, et al. Watching social interactions produces dorsomedial prefrontal and medial parietal BOLD fMRI signal increases compared to a resting baseline. Neuroimage. 2004;21: 1167–1173.
74. Ciechanowski L, Przegalinska A, Magnuski M, Gloor P. In the shades of the uncanny valley: An experimental study of human–chatbot interaction. Future Gener Comput Syst. 2019;92: 539–548.
75. Xu Y, Turk-Browne NB, Chun MM. Dissociating task performance from fMRI repetition attenuation in ventral visual cortex. J Neurosci. 2007;27: 5981–5985.
76. Petrides M, Pandya DN. Association fiber pathways to the frontal cortex from the superior temporal region in the rhesus monkey. J Comp Neurol. 1988;273: 52–66.
77. Lange J, Lappe M. The role of spatial and temporal information in biological motion perception [Internet]. Advances in Cognitive Psychology. 2007. pp. 419–428. doi:10.2478/v10053-008-0006-3
78. Rao RP, Ballard DH. Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. Nat Neurosci. 1999;2: 79–87.
79. Zajonc RB. Emotions. In: Gilbert DT, editor. The handbook of social psychology, Vols. New York, NY, US: McGraw-Hill, x; 1085. pp. 1–2.
80. Cacioppo JT, Tassinary LG, Berntson G, Berntson GG, Wager TD, Hernandez L, et al. Handbook of Psychophysiology edited by John T. Cacioppo. Cambridge University Press; 2007.
81. Kutas M, Federmeier KD. Thirty Years and Counting: Finding Meaning in the N400 Component of the Event-Related Brain Potential (ERP) [Internet]. Annual Review of Psychology. 2011. pp.

621–647. doi:10.1146/annurev.psych.093008.131123
82. Gobbini MI, Koralek AC, Bryan RE, Montgomery KJ, Haxby JV. Two takes on the social brain: a comparison of theory of mind tasks. J Cogn Neurosci. 2007;19: 1803–1814.
83. Bonini L. The Extended Mirror Neuron Network: Anatomy, Origin, and Functions. Neuroscientist. 2017;23: 56–67.
84. Gredebäck G, Johnson S, von Hofsten C. Eye Tracking in Infancy Research [Internet]. Developmental Neuropsychology. 2009. pp. 1–19. doi:10.1080/87565640903325758
85. Larsen JT, Norris CJ, Cacioppo JT. Effects of positive and negative affect on electromyographic activity over zygomaticus major and corrugator supercilii [Internet]. Psychophysiology. 2003. pp. 776–785. doi:10.1111/1469-8986.00078