

Investigating the Uncanny Valley for Prosthetic Hands

**Ellen Poliakoff¹, Sophie O’Kane¹, Olivia Carefoot¹,
Peter Kyberd² and Emma Gowen¹**

1 University of Manchester, Manchester, UK

2 Faculty of Engineering & Science, University of Greenwich, UK

Abstract

Background: In 1970, Mori hypothesised the existence of an 'uncanny valley', whereby stimuli falling short of being fully human are found to be creepy or eerie.

Objectives: To investigate how eerie people find different prosthetic hands and whether perceptions of eeriness can be accounted for by categorical ambiguity.

Study Design: Students participated in computerised experiments during which photographic images of hands were presented.

Methods: We compared photographs of prosthetic hands pre-selected as more (H+) or less human-like (H-), as well as mechanical and real hands. Participants rated the hands for eeriness and human-likeness, as well as performing a speeded classification (human/non-human) and location judgment (control) task.

Results: The H- prosthetic hands were rated as more eerie than the H+ prosthetic, mechanical and real hands, and this was unaffected by hand orientation. Participants were significantly slower to categorise the H+ prosthetic hands compared to the H- prosthetic and real hands, which was not due to generally slower responses to the H+ prosthetic hands (control task).

Conclusions: People find prosthetic hands to be eerie, most consistently for less human-like prosthetic hands. This effect is not driven by ambiguity about whether to categorise the prosthetic hand as human or artificial.

Clinical Relevance

More obviously artificial, less realistic, prosthetic hands consistently generate a sense of eeriness, while more realistic prosthetic hands avoid the uncanny valley, at least on initial viewing. Thus, greater realism in prosthetic design may not always incur a cost, although the role of movement and cutaneous input requires further investigation.

Background

In 1970, the roboticist Mori¹ hypothesised that as artificial stimuli become more human-like our affinity towards them increases, but for levels of human-likeness just below that of a real human, there is a sudden decline in affinity. This effect is often accompanied by feelings of eeriness, repulsion or unsettlement. This was termed the “uncanny valley”, to describe the associated negative affect and analogise the shape of its graphical representation. In his original account, Mori positioned prosthetic hands within the uncanny valley. While users may prefer their prosthesis to be as human-like as possible in order to blend in², Mori’s account suggests that higher level of human-likeness brings the risk that observers experience a sense of eeriness upon discovering that the limb is a prosthesis.³ Therefore, it is of relevance to users and designers of prosthetic limbs to understand the cause and characteristics of the uncanny valley in relation to prostheses.

Although Mori conceived of the uncanny valley more than 40 years ago, empirical investigations have only been carried out more recently, with the majority exploring the uncanny valley for faces (an early example⁴). The first empirical demonstration that participants rate prosthetic hands as more eerie than mechanical and real hands was reported in 2013.⁵ Moreover, when only prosthetic hands were analysed, those rated as less human-like were rated as more eerie. Since the sample of hands used was opportunistic, the first aim of the current work was to more systematically explore responses to prosthetic hands selected for higher and lower human-likeness, using a new stimulus set matched for hand posture. Our second aim was to examine the effect of viewing orientation by presenting the hands from a first person (egocentric) or a third person (allocentric) perspective. The area of the brain involved in visually perceiving body parts (extrastriate body area), responds differentially to body parts presented from the first versus third person perspective.^{6, 7}

Therefore, we hypothesised that participants might rate hands as more eerie when viewed from the first person perspective.

Various accounts of the uncanny valley have been proposed⁸ and two possible explanations are particularly relevant to prosthetic hands. The *mismatch hypothesis* suggests that features or cues that are inconsistent produce a negative response.⁹ Indeed Mori hypothesised that a prosthetic hand would evoke eeriness when touched, since its temperature and hardness would contrast with the visual appearance of a human-like hand. The *categorical ambiguity hypothesis* suggests that difficulty in categorising the stimulus as either human or non-human produces negative affect.⁸ This is supported by evidence using photographic images morphed as blends between human and the artificial showing that people dislike faces that they are slower to categorise.¹⁰ Therefore, the third aim of the current study was to investigate whether the eeriness evoked by prosthetic hands is linked to categorical ambiguity. If this is the case, we would expect participants to be slowest to categorise the hands (human versus artificial) that they rate as most eerie. Since reaction times can also be slowed down by negative affect itself,^{11, 12} we used a control task which involved judging the location of the hand to investigate the general influence of viewing the hands on speeded responses. We hypothesised that responses would be slower for the hands rated as most eerie.

Methods

Participants

All participants were right handed student volunteers. Forty participants (32 female; mean age 20.9 years, SD = 1.28) were recruited for Experiment 1 and 50 (45 female; mean age 19.6, SD = 2.16) for Experiment 2. The study had University of Manchester Research Ethics Committee approval. Participants gave written consent having read a participant information sheet and

were compensated with course credit. Ten additional right-handed participants (8 female; mean age 20.5, SD = 1.84) were used for stimulus characterisation (see below).

Materials

Twelve photographic images were used (Figure 1A) with 3 exemplars in each category: real human, mechanical, prosthetic realistic (H+) and prosthetic unrealistic (H-). These were selected from 21 images using human-likeness ratings from naïve participants (using a 9-point scale, see Procedure). The selected human (range 8.6-8.9) and mechanical (1.1-1.4) hands were the exemplars with the highest and lowest human-likeness ratings and the ratings were used to separate the H+ (5.4-5.9) and H- (3.6-3.9) categories. Each image had a resolution of approximately 650 × 500 pixels and depicted the back of the hand, cropped at the wrist, presented on a black background (Figure 1A). In Experiment 1, stimuli were presented from a first person (wrist downwards) or a third person (wrist upwards) perspective by vertically flipping the image and as a left or right hand by horizontally flipping the image. Presentation of stimuli and recording responses was conducted using E-Prime software. Stimuli were viewed on a 17-inch computer screen with a viewing distance of approximately 60cm (Experiment 1) and 30cm (Experiment 2).

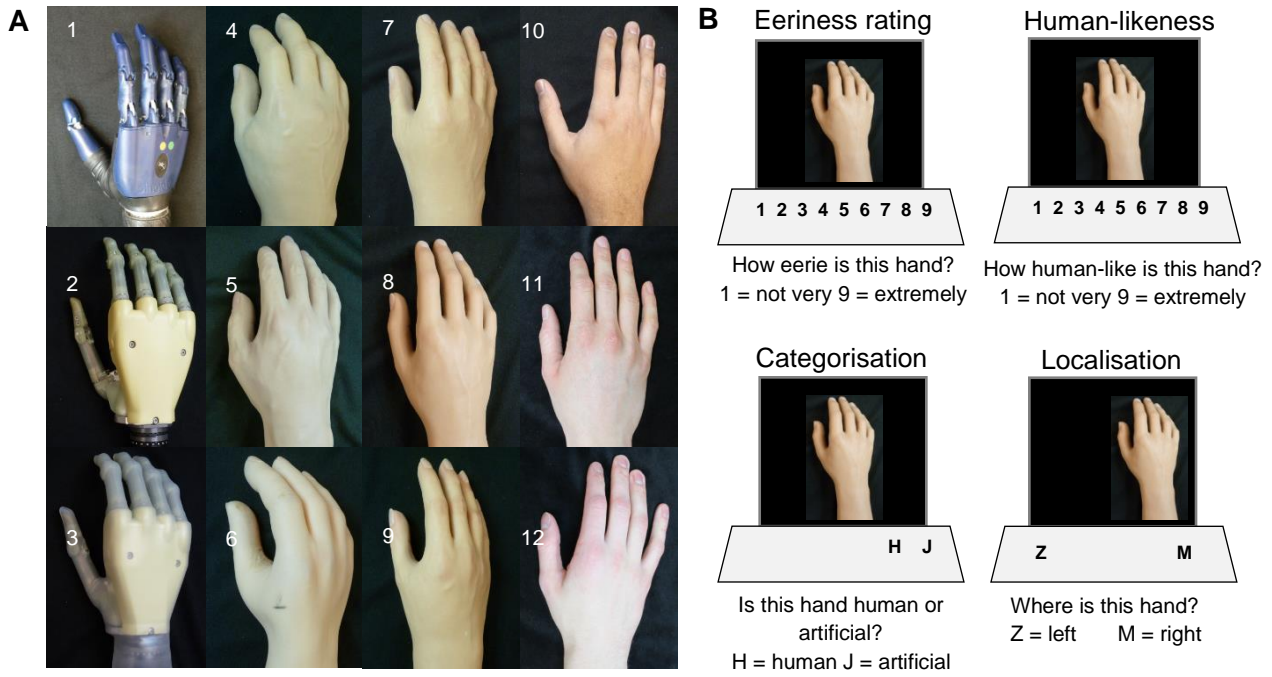


Figure 1 The hand stimuli used in the study: 3 exemplars of mechanical hands (1-3), less human-like (H-) prosthetic hands (4-6), more human-like (H+) prosthetic hands (7-9) and real hands (10-12; Panel A). The four different tasks used in the study (Panel B).

Procedure

For the stimulus characterisation, participants viewed a series of photographic images of real and prosthetic hands. The hand stimuli were presented as right hands and from an first person perspective. A definition of human-likeness ('having human form or attributes') was given to participants. Participants rated the stimuli for human-likeness on a 9-point scale (Figure 1B). Each image was presented individually on a computer screen, where it remained until participants had rated the image. The next image was then presented. Each image was preceded by a fixation slide, which comprised of a small cross in the centre of the screen, visible for 2000 milliseconds. To avoid order effects the images were randomly ordered.

In Experiment 1, participants judged the eeriness of the presented hands (Figure 1B), which was defined as 'mysterious, strange, or unexpected as to send a chill up the spine'. Each image was preceded by a cross presented in the centre of the screen for 2000ms and a reminder of the question was presented alongside every image. The image remained onscreen until the participant provided their rating on a 9 point scale, after which the next trial commenced. Each of the 12 hand exemplars was presented as a left and right hand (horizontally flipped) and from a first person and third person perspective (vertically flipped). All 48 trial types were presented once in a random order before being presented for a second time in a random order (96 trials in total). Participants were instructed to rate each image individually and to avoid rating them in comparison to those previously viewed.

In Experiment 2, participants carried out 4 tasks (see Figure 1B for overview of tasks). They first rated the eeriness of the 12 hands as before (except that the fixation cross was replaced by a 1000ms blank screen), viewing each exemplar twice as a right hand from a first person perspective (24 trials). They then carried out the categorisation and localisation tasks; half of the participants completed the categorisation task first and half completed the tasks in the opposite order. In the categorisation task, they were asked to categorise each hand as

either human or artificial, as quickly and accurately as possible by pressing 'H' for human or 'J' for artificial with their right hand. Each hand exemplar was categorised 12 times (144 trials), with the order randomised. Following the participant's response, the screen turned blank for 1000ms before the next trial. In the localisation task, each hand exemplar was presented 6 times to the left of the screen and 6 times to the right of the screen (144 trials). Participants were asked to indicate the spatial location of each hand image as quickly and accurately as possible, by pressing 'Z' or 'M' with their left and right hand, respectively. To ensure that participants attended to the stimuli,¹³ participants were given an additional task. They were instructed to respond by pressing the spacebar if they saw a small purple square (16x16 pixels) in the centre of the hand image. The square was presented in 12 additional trials. Finally, as a manipulation check, participants carried out a human-likeness rating task, viewing each of the 12 hand exemplars twice (24 trials).

Analysis

In Experiment 1, the ratings of eeriness for left and right images were averaged prior to analysis. In Experiment 2, reaction times were removed if they were extremely low (<150ms) or high (>2000ms for categorisation, 3.4% removed; >1000ms or incorrect response for localisation, 2.1% removed following the removal of one participant with a very high error rate, 46.5%); incorrect responses were included in the averages for the categorisation task. For the localisation task, responses were averaged across left and right hands. Repeated measures ANOVAs were used with Greenhouse Geisser corrections, which were followed up using repeated contrasts.

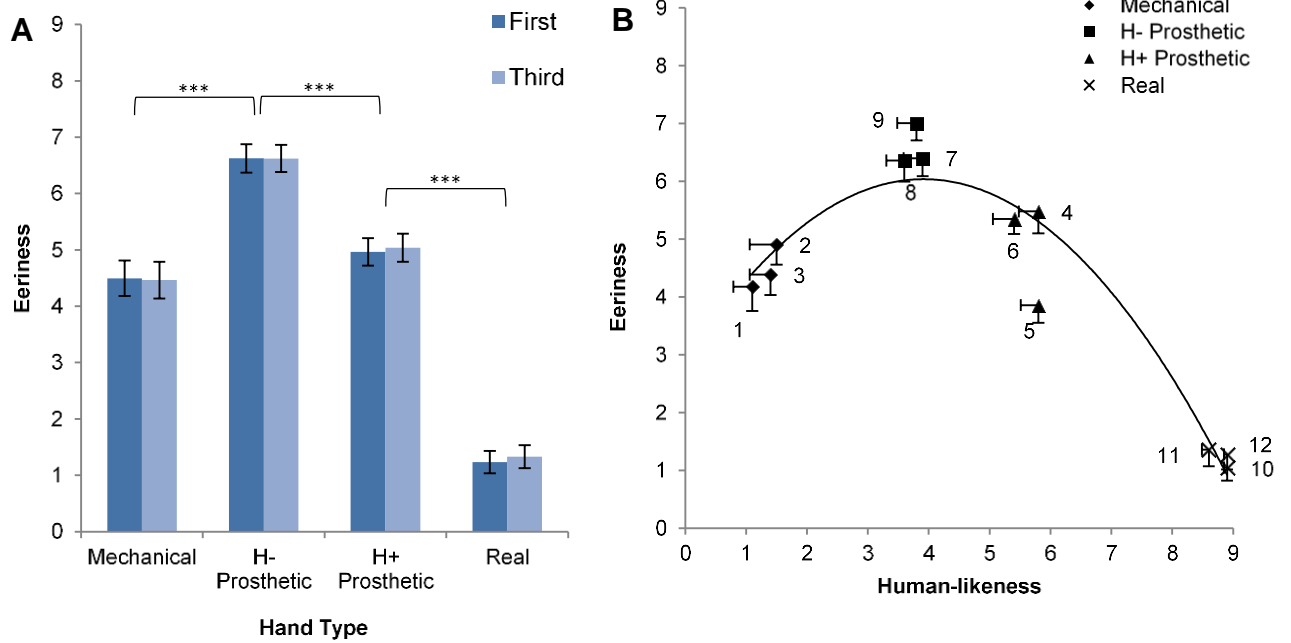


Figure 2 Mean (\pm SEM) eeriness ratings for each type of hand presented from the first and third person perspective (panel A). Mean eeriness (\pm SEM) ratings plotted against mean human-likeness ratings for individual hand exemplars; the numbers refer to the images in Figure 1 (panel B). *** denotes significant at $p < .001$.

Results

Experiment 1

A Hand Type (mechanical, prosthetic H-, prosthetic H+, Real) x Orientation (first-person, third-person) repeated measures ANOVA was used to explore eeriness ratings (see Figure 2A). There was a significant main effect of Hand Type ($F(1.7, 65) = 86.02, p < .001, \text{partial } \eta^2 = .688$). Prosthetic H- hands were rated as significantly more eerie than both mechanical ($F(1,39) = 20.9, p < .001, \text{partial } \eta^2 = .349$) and prosthetic H+ hands ($F(1,39) = 118.8, p < .0005, \text{partial } \eta^2 = .753$). Prosthetic H+ hands were, in turn, rated as significantly more eerie than real hands ($F(1,39) = 214.0, p < .001, \text{partial } \eta^2 = .846$). Note that eeriness generally reduced with increased human-likeness, with the exception of the mechanical hands which had intermediate ratings of eeriness, but were lowest in human-likeness. Neither the main effect of Orientation ($F(1, 39) = .85, p = .361, \text{partial } \eta^2 = .021$) nor the interaction between Hand Type and Orientation ($F(2.8, 108) = .51, p = .674, \text{partial } \eta^2 = .013$) were significant.

The relationship between eeriness and human-likeness was further explored across the different hand exemplars in the right egocentric orientation (Figure 2B). Although there was a significant negative linear correlation between ratings of eeriness from Experiment 1 and human-likeness from the stimulus characterisation phase ($r^2(12) = .47, p = .015$), the relationship was better captured by a significant quadratic relationship ($r^2 = .92, p < .001$).

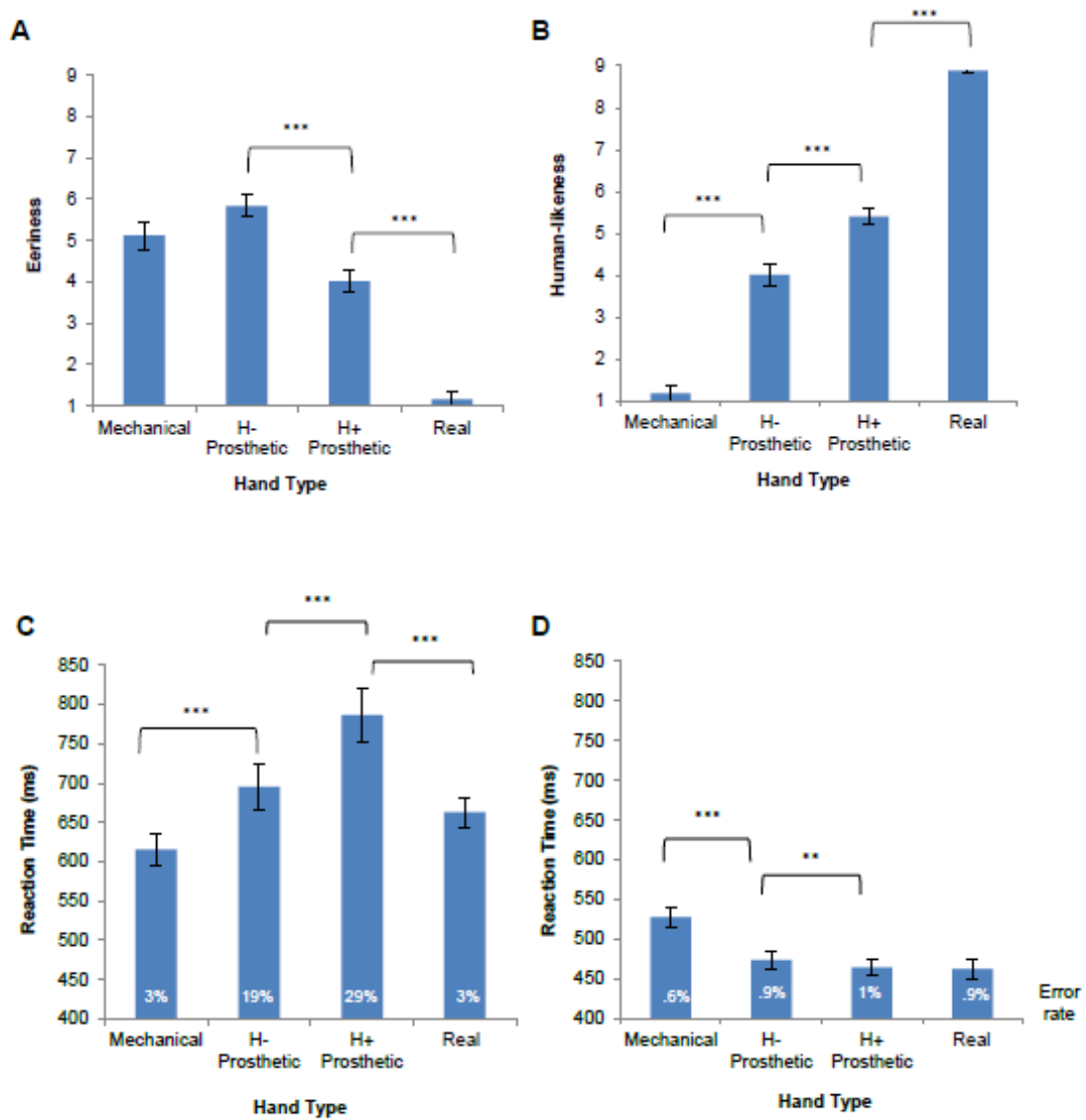


Figure 3 Descriptive statistics for the results of the 4 tasks across each hand type. Mean eeriness rating (panel A), mean human-likeness rating (panel B), mean reaction time (ms) to categorise hands as human or artificial (panel C) and mean reaction time (ms) to categorise by spatial location (panel D). Mean error rates are presented in panels C and D. Note that ** denotes significant at $p < .01$; *** significant at $p < .001$.

Experiment 2

Hand type significantly affected eeriness ratings ($F(2.0,99) = 86.18, p < .001$ partial $\eta^2 = .64$; Figure 3A). Prosthetic H- hands were rated as more eerie than prosthetic H+ hands, $F(1,49) = 145.82, p < .001$, partial $\eta^2 = .74$) which in turn were rated more highly than human hands ($F(1,49) = 136.21, p < .001$, partial $\eta^2 = .75$). Mechanical hands received lower eeriness ratings than prosthetic H- hands and this difference was borderline significant ($F(1,49) = 3.97, p = .052$, partial $\eta^2 = .08$).

Hand type significantly affected human-likeness ratings ($F(2.0,98) = 369.49, p < .001$, partial $\eta^2 = .88$; Figure 3B). Significant differences were found between all hand-types in ascending order: prosthetic H- hands were rated as significantly more human-like than mechanical ($F(1,49) = 85.42, p < .001$, partial $\eta^2 = .64$), prosthetic H+ hands were rated as significantly more human-like than prosthetic H- ($F(1,49) = 86.89, p < .001$, partial $\eta^2 = .64$) and human hands were rated as significantly more human-like than prosthetic H+ ($F(1,49) = 331.69, p < .001$, partial $\eta^2 = .87$).

Hand type significantly affected categorization reaction times ($F(1.9,94.6) = 23.96, p < .001, \eta^2 = .33$; Figure 3C). Categorization of prosthetic H- hands was slower than for mechanical hands ($F(1,49) = 11.11, p = .002$, partial $\eta^2 = .19$, but significantly faster than for prosthetic H+ hands ($F(1,49) = 33.94, p < .001$, partial $\eta^2 = .41$), while categorisation of human hands was significantly faster than prosthetic H+ hands ($F(1,49) = 31.56, p < .001$, partial $\eta^2 = .39$). Error rates were not analysed statistically, but the error rate was highest for the prosthetic H+ hands (Figure 3C).

Hand type significantly affected localisation reaction times ($F(2.6,122.7) = 102.25, p < .001$, partial $\eta^2 = .68$; Figure 3D). Reaction time for prosthetic H- hands was significantly decreased relative to mechanical ($F(1,48) = 122.08, p < .001$, partial $\eta^2 = .72$) and significantly

higher than for prosthetic H+ hands ($F(1,48) = 7.73, p = .008, \text{partial } \eta^2 = .14$). The difference in reaction time between H+ prosthetic and human hands was not statistically significant ($F(1,48) = 0.32, p = .573, \text{partial } \eta^2 = .01$). Error rate was very similar across hand types. Participants performed the additional task (responding to the purple square) with a good level of accuracy (89.1%) showing that they were attending to the images.

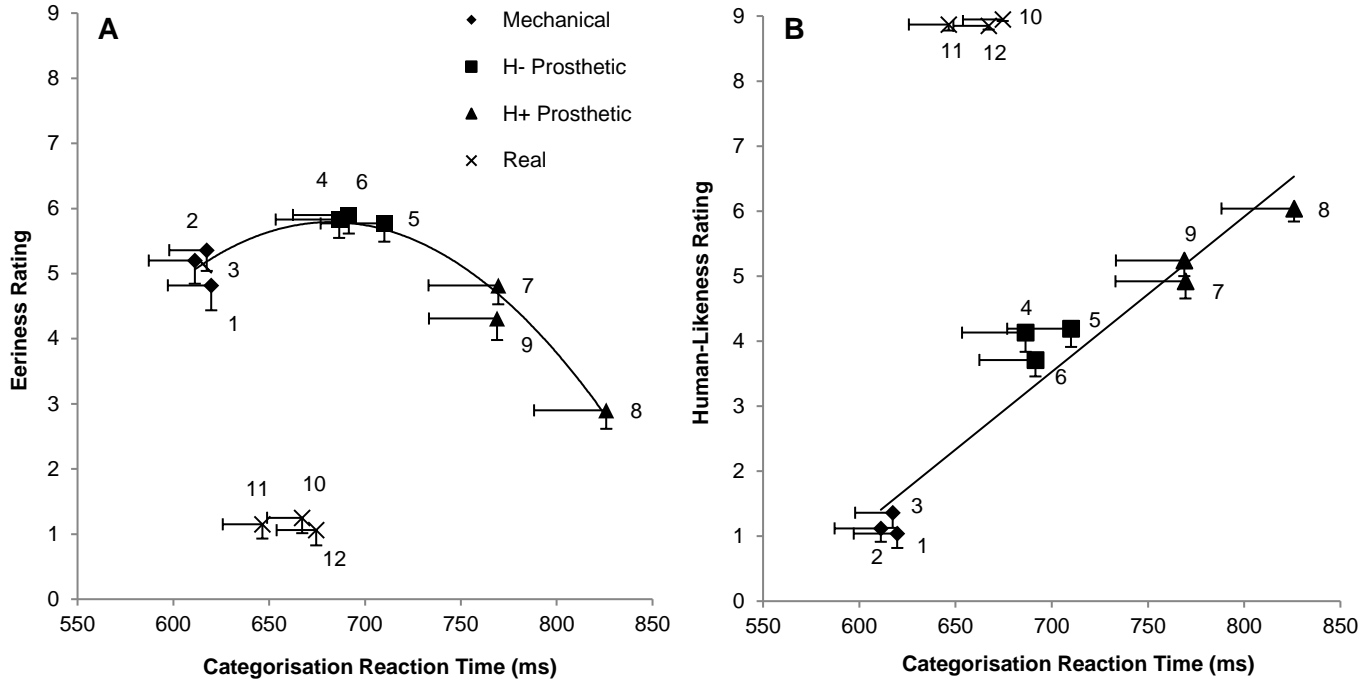


Figure 4 Mean (-SEM) eeriness ratings plotted against mean (-SEM) reaction time to categorise each individual hand image as human or artificial (msec) across the 3 artificial hand-types and real human hands. With the exception of the real hands, categorisation time shows a curvilinear relationship with eeriness ratings (panel A) and a linear relationship with human-likeness (panel B); the more human-like artificial hands are slower to categorise as human or artificial. The numbers refer to the images in Figure 1.

The relationship between categorisation times and eeriness ratings was further explored for individual hand exemplars and showed a curvilinear relationship, when the real hands were excluded (Figure 4A). H- prosthetic hands were rated as most eerie, whilst falling in the midrange of the reaction time data. As seen in the analyses above, participants were slowest to categorise the H+ prosthetic hands. Mechanical hand types were found to generate the fastest responses overall and generated intermediate levels of eeriness. Note that aside from the positioning of the real hands, Figure 4A is very similar to the relationship between eeriness and human-likeness ratings obtained in Experiment 1 (Figure 2B). Therefore, we examined the relationship between human-likeness ratings and categorisation RT (Figure 4B). The artificial hands exhibited a linear relationship ascending from mechanical to realistic prosthetic ($r(7) = .92, p = .001$); participants were slower to categorise the artificial hands that they rated as more human-like.

Discussion

Across two experiments, static images of less human-like prosthetic hands were consistently rated as eerie relative to more human-like prosthetic, real human and mechanical hands. This replicates previous findings,⁵ with hand stimuli pre-selected for their level of human-likeness. We also extend this work by firstly demonstrating that ratings of eeriness (Experiment 1) were not influenced by orientation (first versus third-person), fitting with findings that higher level processing may be unaffected by perspective¹⁴. Second, we showed that eeriness could not be accounted for by categorical ambiguity (Experiment 2). Participants were slowest to categorise the H+ prosthetic hands as artificial versus human and people were slower to categorise the more human-like hand exemplars. That people are slower to categorise more human-like artificial hands is unsurprising, but what is interesting is that this pattern contrasts with the ratings of eeriness. That is, the H+ prosthetic hands were slowest to categorise, but

rated less eerie than the H- hands; indeed hand 8 produced the slowest categorisation time, but the lowest eeriness rating. This indicates that categorical ambiguity cannot be driving the uncanny phenomenon for prosthetic hands. The categorisation hypothesis has received only limited support in the wider literature and previous findings may have been driven by a dislike of image artefacts produced by morphing.¹⁵ The current findings provide further evidence against this hypothesis, using a different type of stimulus, without the potentially complicating effects of facial expressions. Nevertheless, categorisation could still play a role, in that the H- hands which were rated as more eerie were more easily categorised as artificial.¹⁵

The control (localisation) task (Experiment 2) indicated that speeded responses were slightly slower for the H- than the H+ prosthetic hands in general, which is likely to be due to a general slowing of responses to negative stimuli. This fits with a study which found that reaction times were slower for fish stimuli rated as more eerie¹². An unexpected finding was that responses were significantly slower for the mechanical hands compared to all other stimuli. It is likely that this was generated by the additional task (finding the purple square) designed to ensure that participants attended to the detail of the image, as well as its location. Given the fine detail and coloured patches on the mechanical hands, this task may have disproportionately slowed responses for this category of hand.

The current findings show that prosthetic hands do evoke eeriness in the observer, which is generally consistent with the notion of the uncanny valley. However, rather than focussing on the valley, the evidence for the which is scarce and problematic to obtain^{15, 16}, it may be more appropriate to consider the underpinnings of the “uncanny phenomenon” (feelings of uncanniness)¹⁶ for prosthetic hands. A strong contender is the aforementioned perceptual mismatch hypothesis, which would suggest that key features such as unnatural skin colour or the absence of veins or fingernails may have driven the high eeriness ratings for the H- hands. Indeed, recent evidence showed that inconsistency in realism, rather than

categorical ambiguity, produced uncanny responses to animal and human faces.¹⁷ The dehumanisation hypothesis should also be explored; this posits that the uncanny phenomenon arises when an artificial entity which has previously been anthropomorphised is dehumanised.¹⁶

Mori originally suggested that prosthesis designers should avoid aiming for realism and opt for appealing materials such as wood¹. The fact that our more realistic prosthetic hands appeared less eerie than the less realistic hands argues against this (see also³). Nevertheless, future research should address limitations of the current work by exploring the role of movement and cutaneous input elicited by moving three-dimensional hands; indeed, Mori¹ hypothesised that movement would accentuate the uncanny valley (although see¹⁸) and that the feel of a prosthetic hand could be surprising and unsettling.¹ It is also possible that orientation could affect perceptions of eeriness if the hands were presented closer to the participant and/or as if the hand could be attached to the participant's own body (peripersonal space).¹⁸ Furthermore, the hand stimuli in the current study were not matched for luminance, contrast and colour; colour in particular merits further investigation since it contributes to eeriness in artificial faces.⁸ Finally, familiarity¹⁹ with prosthetic hands and individual differences should be investigated; in the current study it is likely that some participants mistook the more human-like prosthetic hands for real hands.

Conclusion

Viewing static prosthetic hands that are less human-like generates a feeling of eeriness for the observer. This uncanny phenomenon was not related to difficulty in categorising the hands as human versus non-human and was not modulated by the orientation of the viewed hand. It is likely that inconsistent features in the less human-like hands played a role in generating the phenomenon.

References

1. Mori M. The uncanny valley [Bukimi no tani] (K. F. MacDorman & Norri Kageki, Trans.). *IEEE Robotics and Automation*. 2012; 19: 98-100.
2. Murray CD. Being like everybody else: the personal meanings of being a prosthesis user. *Disability and Rehabilitation*. 2009; 31: 573-81.
3. Cabibihan JJ, Carrozza MC, Dario P, Pattofatto S, Jomâa M and Benallal A. The uncanny valley and the search for human skin-like materials for a prosthetic fingertip. *Humanoid Robots*. 2006; 1: 474-7.
4. MacDorman KF and Ishiguro H. The uncanny advantage of using androids in cognitive and social science research. *Interaction Studies: Social Behaviour and Communication in Biological and Artificial Systems*. 2006; 7: 297–337.
5. Poliakoff E, Beach N, Best R, Howard T and Gowen E. Can looking at a hand make your skin crawl? Peering into the uncanny valley for hands. *Perception*. 2013; 42: 998-1000.
6. Saxe R, Jamal N and Powell L. My body or yours? The effect of visual perspective on cortical body representations. *Cerebral Cortex*. 2006; 16: 178-82.
7. Chan AW-Y, Peelen MV and Downing PE. The effect of viewpoint on body representation in the extrastriate body area. *Neuroreport*. 2004; 15: 2407-10.
8. MacDorman KF, Green RD, Ho CC and Koch CT. Too real for comfort? Uncanny responses to computer generated faces. *Computers in Human Behavior*. 2009; 25: 695-710.
9. Brenton H, Gillies M, Ballin D and Chatting D. The uncanny valley: Does it exist? *Proceedings of conference of human computer interaction, workshop on human animated character interaction*. 2005.

10. Yamada Y, Kawabe T and Ihaya K. Categorization difficulty is associated with negative evaluation in the “uncanny valley” phenomenon. *Japanese Psychological Research*. 2013; 55: 20-32.
11. Poliakoff E, Latif A, Maehr AM, et al. An unsuccessful attempt to demonstrate attentional orienting within the purely emotional domain. *Emotion*. 2016; 16: 6-10.
12. Takahashi K, Fukuda H, Samejima K, Watanabe K and Ueda K. Impact of stimulus uncanniness on speeded response. *Frontiers in Psychology*. 2015; 6: 662.
13. Brown RJ, Danquah AN, Miles E, Holmes E and Poliakoff E. Attention to the body in non-clinical somatoform dissociators depends on emotional state. *Journal of Psychosomatic Research*. 2010; 69: 241-8.
14. Oosterhof NN, Tipper SP and Downing PE. Viewpoint (in)dependence of action representations: An MVPA study. *Journal of Cognitive Neuroscience*. 2012; 24: 975-89.
15. Kätsyri J, Förger K, Mäkräinen M and Takala T. 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*. 2015; 6: Article 390.
16. Wang S, Lilienfeld SO and Rochat P. The Uncanny Valley: Existence and Explanations. *Review of General Psychology*. 2015; 19: 393–407.
17. MacDorman KF and Chattopadhyay D. Reducing consistency in human realism increases the uncanny valley effect; increasing category uncertainty does not. *Cognition*. 2016; 146: 190-205.
18. Lloyd D. Spatial limits on referred touch to an alien limb may reflect boundaries of visuo-tactile peripersonal space surrounding the hand. *Brain and Cognition*. 2007; 64: 104-9.