Video Article

Perceptual and Category Processing of the Uncanny Valley Hypothesis' Dimension of Human Likeness: Some Methodological Issues

Marcus Cheetham[1], Lutz Jancke[1]

[1]Department of Neuropsychology, University of Zurich

Correspondence to: Marcus Cheetham at m.cheetham@psychologie.uzh.ch

Abstract

Mori's Uncanny Valley Hypothesis[1,2] proposes that the perception of humanlike characters such as robots and, by extension, avatars (computer-generated characters) can evoke negative or positive affect (valence) depending on the object's degree of visual and behavioral realism along a *dimension of human likeness* (*DHL*) (**Figure 1**). But studies of affective valence of subjective responses to variously realistic non-human characters have produced inconsistent findings [3, 4, 5, 6]. One of a number of reasons for this is that human likeness is not perceived as the hypothesis assumes. While the DHL can be defined following Mori's description as a smooth linear change in the degree of physical humanlike similarity, subjective perception of objects along the DHL can be understood in terms of the psychological effects of categorical perception (CP) [7]. Further behavioral and neuroimaging investigations of category processing and CP along the DHL and of the potential influence of the dimension's underlying category structure on affective experience are needed. This protocol therefore focuses on the DHL and allows examination of CP. Based on the protocol presented in the video as an example, issues surrounding the methodology in the protocol and the use in "uncanny" research of stimuli drawn from morph continua to represent the DHL are discussed in the article that accompanies the video. The use of neuroimaging and morph stimuli to represent the DHL in order to disentangle brain regions neurally responsive to physical human-like similarity from those responsive to category change and category processing is briefly illustrated.

Video Link

The video component of this article can be found at https://www.jove.com/video/4375/

Protocol
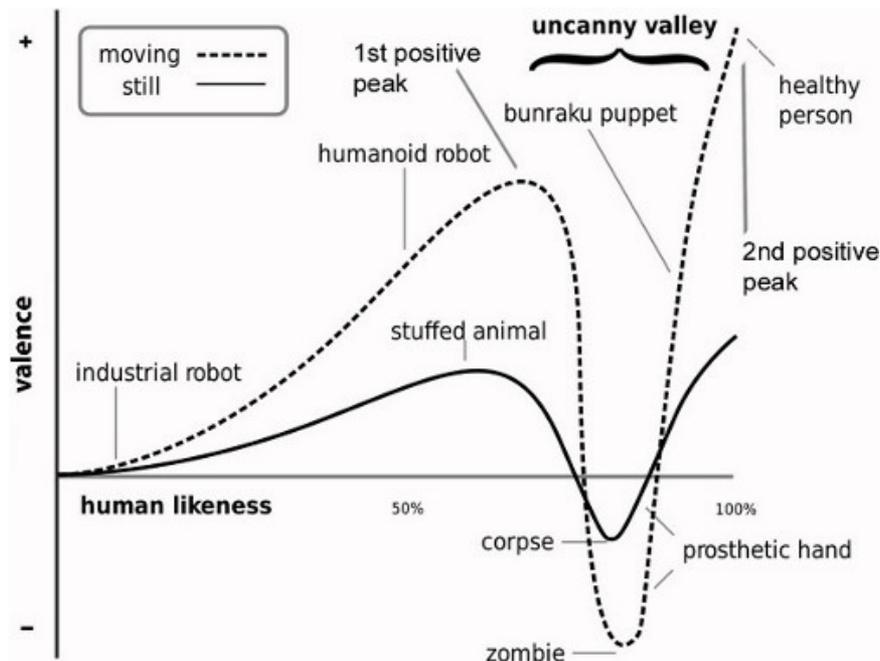
Figure 1. Illustration of the non-linear relationship between the experience of negative and positive affect (valence) and perceived human likeness. The otherwise positive relationship shows a sharp negative peak (*i.e.* uncanny valley) at the level of realism between the first and second positives peaks of the depicted curve at which subtle differences in the appearance and behavior of a highly realistic yet discernibly unnatural humanlike object is suggested to elicit a sense of strangeness and personal discomfort (*i.e.* an uncanny feeling). Illustration adapted from [2].

We used different groups of participants for each of the following tasks.

1. Forced Choice Classification Task

1.1 Stimuli

1. Use avatar and human images as parent faces (*i.e.* continua endpoints) in the morphing procedure to produce linear morph continua to represent the DHL. We created 32 human-avatar continua using 32 images of human and avatar faces, respectively. Generate avatars using the modeling suite Poser 7 (Smith Micro Software, www.smithmicro.com), though other software is available. We generated these morph continua using Funmorpher (Zealsoft Inc., Eden Prairie, MN), but other morphing software may be used.
2. Using the morph software, place control points on the corresponding features of the parent faces. For each face, we placed 20 points on the mouth, 18 points on each eye, 20 points on the nose, and 8 points on each eyebrow. We thus used around 100 control points. Try to keep the number of control points constant, but add further points to eliminate any artefacts in the final morphs of the continua.
3. Ensure potential confounds are not introduced into the morphing procedure. For example, we used as endpoints of each continuum images of unknown indistinctive male faces with neutral expression, direct gaze and no other salient features such as facial hair or jewelery, and the endpoint images were closely matched for age, configural cues and general facial geometry.
4. Use photo editing software to crop external features by using for example a black overlay in elliptic form; we used Adobe; Photoshop; CS3 (www.adobe.com). Before morphing, adjust the position of the images to ensure alignment between the endpoint images of configural cues, and adjust contrast levels, overall brightness and skin tone of each pair of endpoint stimuli of each continuum to match.
5. Each morph of a DHL continuum represents a difference in physical humanlike similarity at predefined increments. We generated 13 different morphed images and labeled these M0 to M12, that is, the two endpoints and 11 intermediate morphs (**Figure 2B**).

1.2 Stimulus presentation and instructions

1. Use a *two-alternative forced choice classification task* to determine which of these morphs are clearly categorized as avatars and as human and to define the position of the category boundary [8].
2. Present trials beginning with a fixation point for 500 msec (participants are required to maintain fixation) followed by a morph image for 750 msec. We used Presentation; software (Version 14.1, www.neurobs.com) for stimulus presentation in all tasks in this protocol, but other stimulus presentation platforms may be used.
3. Instruct the participant to identify the presented morph stimulus as either an avatar or human as quickly and precisely as possible by pressing one of two response keys.

1.3 Data Analysis

Summarize the avatar-human classification data using polynomial regression to describe the shape of the response function. Determine this by fitting logistic function models to the response data of each participant and continuum. First, analyze the individual continua across participants to ensure best fit of logistic functions. Then, test against zero in a one-sample t-test for a step-like shape in the avatar-human category response function across all continua using the parameter estimates derived from the logistic function of each continuum, averaged across participants. Estimate the position of the category boundary along each continuum by submitting the parameter estimates of the logistic function of each continuum to a logit transformation [9]. We performed all analyses for the forced choice classification and perceptual discrimination tasks using SPSS Version 16 (www.ibm.com/software/analytics/spss).

Response time (RT) data may also be analysed. In the present analysis, differences in response times depending on morph position are entered in a one factorial ANOVA, with 13 morph positions, using the mean RT of each individual across all continua as dependent variable.
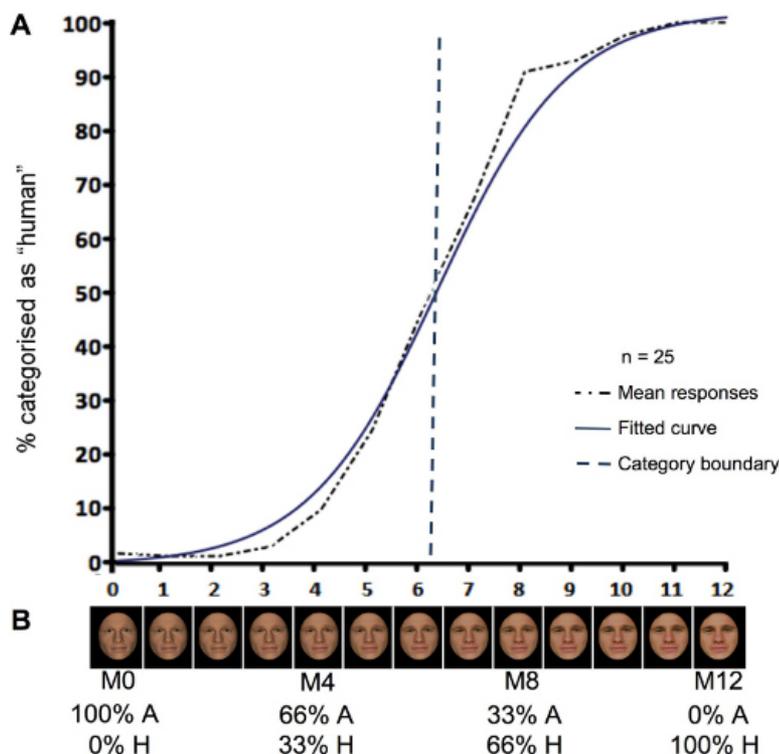
Figure 2. Results from the forced choice categorisation task (**A**) and an example of a morph continuum (**B**). In panel **B**, the relative degree of linear physical transition along the 13 morph-continuum between the avatar and human endpoints is shown as a percentage. M0 and M4 were identified as avatars and M8 and M12 as human in the forced choice classification task, as shown in panel **A**.

2. Perceptual Discrimination Task

2.1 Stimuli

1. For this version of the *same-different perceptual discrimination task* [10], select from each morph continuum two morphs categorized in the preceding classification task as avatars (*e.g.* M0 and M4) and two as human (*e.g.* M8 and M12). To control for physical differences between morphs, select morphs that represent equivalent increments of physical change along each continuum. We used increments of 33.33% (*i.e.* M0, M4, M8, M12) (**Figure 2B**).
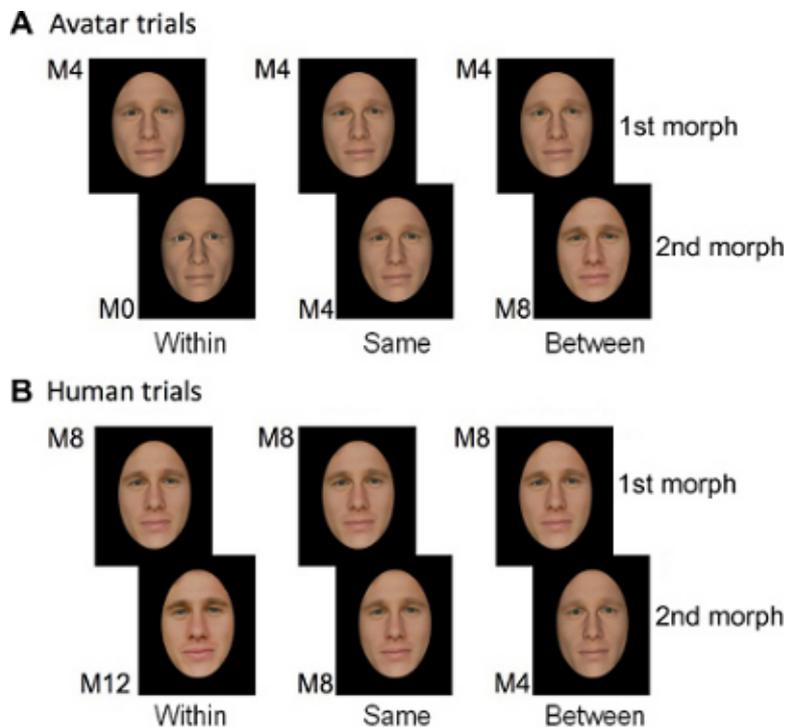
Figure 3. Stimulus conditions for the "same-different" perceptual discrimination task (N = 20). Morphs are selected to form pairs. The morphs of a pair are drawn from within the same category ("within"), are identical ("same"), or they show a change in category between them ("between"). The morphs M0, M4, and M8 are used for avatar trials (**A**) and M4, M8, and M12 for human trials (**B**). Note that the first morph of a morph pair in avatar trials is always M4 and in human trials M8 and that avatar and human trials are based on morphs drawn from different continua.

2. Sort the selected morphs into pairs according to the three experimental morph-pair conditions (**Figure 3**): "same" (the morphs of a pair are identical, representing no physical or category change), "within" (the morphs of a pair are drawn from within a category), and "between" (the morphs of a pair represent different categories).

3. To investigate discrimination performance between the morphs of the morph pairs in relation to the avatar category (these morph pairs are thus termed "avatar trials") ensure that the first morph of each morph pair in the three conditions is always M4 (from the avatar category) (**Figure 3A**). This results in morph pairs M4 - M4 for the "same", M4 - M0 for the "within" and M4 - M8 for the "between" conditions. The same procedure can be applied for morph pairs in relation to the human category (thus termed "human trials"), ensuring that the first morph is always M8: "same" (M8 - M8), "within" (M8 - M12), and "between" (M8 - M4) (**Figure 3B**).

4. Always ensure that both morphs of a morph pair are drawn from the same continuum in which they were originally morphed. Pseudo-randomize presentation of morph pairs so that no pairs from within the same continuum are shown in close sequence. Presentation of avatar or human trials from a given continuum is random but counterbalanced across all participants to ensure that each participant views either avatar or human trials from any given continuum but not both, and that an equal number of avatar or human trials are viewed.

2.2 Presentation and instructions

1. Present a fixation cross for 500 msec (participants are required to maintain fixation) followed by each face of a face pair for 500 msec with an inter stimulus interval (ISI) of 300 msec between the faces of a pair. We also used an ISI of 75 msec to verify whether different durations of ISI would differentially influence discrimination performance. Present a variable inter-trial interval between trials of morph pairs: we used an interval averaging 2,500 msec.

2. Instruct participants to view each trial comprising a morph pair, the morphs being presented successively in the trial, and to indicate by button press as quickly and precisely as possible whether the faces of each face pair are the 'same' or 'different' in appearance.

2.3 Data Analysis

Discrimination accuracy is analysed for face pairs that cross the category boundary compared with face pairs from the same side of the boundary. For this, the 'different' responses (indicating that both faces of a pair are of different physical appearance) are computed as proportions of the total number of morph face pairs and subjected to a 2 X 3 factorial ANOVA, with 3 "face-pair trial types" (within, between, same) and 2 "ISI" conditions (75 msec, 300 msec). Greenhouse-Geisser adjustment is used when the assumption of sphericity is violated. The data for avatar trials and human trials are treated separately in analysis.

Individual accuracy scores may also be determined using the *A'* statistic [47,79] (for Signal Detection Theory, see, *e.g.* [45, 46, 47]). *A'* provides a measure of discrimination sensitivity that is independent of response bias. It varies between 0.5 (chance) and 1 (perfect discrimination). Various software packages may be used to compute A' and other measures of discrimination sensitivity (and bias) [46, 47, 48 49, 50]. We analysed discrimination sensitivity using a 2 X 2 repeated measures ANOVA, with 2 "face-pair trial types" (within, between) and "ISI" conditions (75 msec, 300 msec), with separate analyses for avatar trials and human trials, and A' as the dependent variable. Response bias is not often generally

reported, but see [38]. For response bias, we used the β''_D statistic [47] as the dependent variable in a separate analysis using otherwise the same 2 X 2 ANOVA design.

RT data may also be analysed for "different", "same" and "between" responses. In this example, we compare the "different", "same" and "between" conditions for avatar and human trails in one analysis to gain a summary view of RT across all conditions. For this, we performed a 3 X 2 X 2 ANOVA with the factors "face-pair trial types" (different, same, between), "category" (avatar, human) and "ISI" (75 msec, 300 msec), using the mean RT of correct responses of each individual across all continua as the dependent variable.

3. fMRI Task

3.1 Stimuli

The stimulus conditions, *i.e.* the morph stimuli for the face pairs in the within, same and between conditions in the avatar and human trials, are the same as described in the preceding perceptual discrimination task.

3.2 Presentation and instructions

1. Use a target monitoring task to examine implicit processing of physical and category-related change along the DHL while maintaining participants' attention to the stimuli of interest.
2. Instruct participants to press a response button upon detection of a rare target. We presented 15% of all morph pairs as targets, the faces being shown upside-down. Use as targets one of four possible morphs (M0, M4, M8, or M12) selected at random from a morph continuum not used for stimulus presentation otherwise. Ensure that the target morph is presented as the first or second morph of a morph pair to avoid differential attention during monitoring for targets to the first or second morph of the morph pairs.
3. Each scanning session consists of two experimental runs of stimulus presentation counterbalanced in order across participants. The break between runs allows the participants a brief rest. Participants fixate a cross at the beginning of each run to establish a steady state in the MR signal.

3.3 Preparing the Subject for the Scan

1. All participants provide written informed consent before the experimental protocol is conducted. The protocol, all procedures and consent forms are approved by the local Ethics Committee. Avoid confounds in the lateralization of brain activations by scanning right-handed participants. Control for the potential impact of previous experience with avatars.
2. Before scanning, participants are familiarized with the laboratory, informed about the scanning procedures, given clear instructions as to the target monitoring task, total scanning time and how to alert staff if required.
3. For scanning, the participant lies supine on the scanning table. Head cushions are used to ensure comfort and minimize head movement during scanning. Participants are given earplugs and headphones to reduce the impact of scanner noise and to enable communication with the experimenter.
4. The participants' right hand is positioned over the response panel for the target monitoring task. The left hand is placed next to the emergency stop button should the participant want to stop scanning.
5. The visual stimuli may be presented on a projection screen placed in front or at the rear of the MRI scanner. We used an MRI-compatible head-mounted display ("VisuaStim - Digital", Resonance Technology Inc.). This has the advantage of excluding from sight all visual input other than the intended stimuli.
6. Before commencing data collection, ensure that stimulus presentation, the response panel and the emergency stop button are operating properly.

3.4 Data Recording and Scanning Parameters

We acquired structural and functional images of the entire brain using a 3-T whole-body MR unit (Philips Medical Systems, Best, The Netherlands). Structural images were registered using a T1-weighted 3D, spoiled gradient echo pulse sequence (180 slices, TR = 20 msec, TE = 2.3 msec, flip angle = 20°, FOV = 220 mm × 220 mm × 135 mm, matrix size = 224 × 187, voxel size = 0.98 mm × 1.18 mm × 0.75 mm, resliced to 0.86 mm × 0.86 mm × 0.75 mm). Functional images were acquired from 225 whole-head scans per run using a single-shot echo planar sequence (repetition time, TR = 2.6 sec; echo time, TE = 35 msec; field of view = 220 mm × 220 mm × 132 mm; flip angle = 78°; matrix size = 80 × 80; voxel size = 2.75 mm × 2.75 mm × 4 mm, resliced to 1.72 mm × 1.72 mm × 4 mm).

3.5 Data Analysis

1. We used MATLAB 2006b (Mathworks Inc., Natick, MA, USA) and the SPM5 software package (http://fil.ion.ucl.ac.uk/spm) for preprocessing and MRI data analysis. Preprocessing typically entails alignment of images to the first recorded volume, motion correction, normalization into standard stereotactical space, and smoothing (*e.g.* 6 mm³ Kernel).
2. The fMRI data analysis makes use of a phenomenon referred to as repetition suppression (RS) ([11, 13, 14], for reviews, see [15, 16]). Considered in the context of the DHL, the morphs of a morph pair are presented in rapid succession. Repetition in the second morph of the stimulus or stimulus attributes presented in the first morph results in a decrease in activation (*i.e.* RS) in brain region's sensitive to that specific stimulus or its attributes (*e.g.* physical or category-related attributes). In this protocol, repetition of the stimulus or stimulus attributes between the first and second morph is manipulated in the "within", "between", and "same" conditions in terms of similarity or dissimilarity of physical and category-related attributes of the DHL. By contrasting these conditions, the fMRI data analysis identifies brain regions engaged in processing a particular stimulus or physical or category-related stimulus attribute on the basis of the extent of relative differences in signal decrease following stimulus repetition [17, 18, 19, 20].
3. Identify brain regions responsive to physical and to category-related change along the DHL using the following contrast of the stimulus conditions (within, between, and same). These contrasts are defined in terms of the morph used as the second face in the three face-pair conditions (note that the first morph is the same in the avatar and human trials, respectively). To detect sensitivity to physical change for the avatar trials, use the contrast M0 plus M8 > M4, and use M12 plus M4 > M8 for the human trials. To detect brain regions selectively

responsive to category change across the boundary in the direction avatar to human (*i.e.* avatar trials), use the contrast M8 > M4 plus M0. For the direction human to avatar, use the contrast M4 > M8 plus M12.

4. For individual level analyses, the fMRI responses of each subject to the second morph of each morph pair in each of the six morph pair conditions (*i.e.* within, same, and between for avatar and human trials) may be used to contrast brain activity between these conditions. These individual contrasts are then entered into group level analyses for inferential purposes.

Representative Results

1. Forced choice classification task

Analysis of the response data of N = 25 participants was already reported in [7]. This confirmed that the slope of the fitted regression curve of each individual continuum and across all continua has a logistic profile (**Figure 2A**). This slope reflects a sigmoid step-like function consistent with the presence along the DHL of a categorical component in the responses of the participants to the morph faces of the continua. The slope of the curve is thus characterized by lower and upper asymptotes of avatar or human categorization responses which approach 100% for avatars and 100% for humans. In contrast, the estimate of the mean category boundary value derived from the fitted logistic curve and the ordinate midpoint between the lower and upper asymptotes of the categorization responses indicates that the maximum uncertainty of 50% in categorization judgments is associated with the morph M6.

Analysis of RT data was reported also in [7]. The RT analysis of all morphs (see **Figure 4**) showed shortest RTs for the avatar and human ends of the continua, increasing RT with greater morph distance from the avatar and human ends of the continua, and longest RTs at M6 at which there is maximum uncertainty in the category decision responses, as can be seen in **Figure 2B**. To verify the latter finding more clearly, the mean RT values at M6 can be compared with the mean RT values at all other morph positions. A one-way RM-ANOVA analysis with morph position (two levels: M6 versus all other morphs) and RT as dependent variable collapsed across continua showed that RT for M6 (M = 1.42, SD = 0.26) differed highly significantly from RT for the other morph positions (M = 0.99, SD = 0.46), $F(1,24)$ = 62.04, $p < 0.001$.

Taken together, the categorization response data confirm that the first criterion for the presence of CP is fulfilled, namely that there is a category boundary (for all criteria, see *e.g.* [11]), and the response times for the category decisions are consistent with the response data in that they show longer response times with increasing categorization uncertainty.
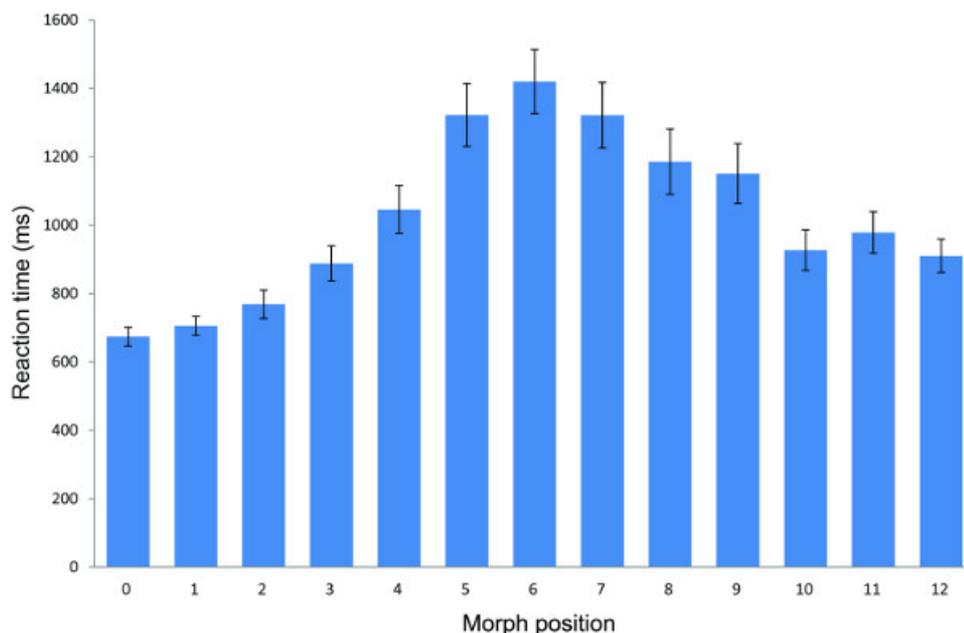
Figure 4. Reaction time results of the forced choice categorization task, showing longest mean response latency for categorization judgments for stimuli at morph position M6 at which categorization ambiguity is greatest. Error bars show ±1 standard error.

2. Perceptual discrimination task

The data analyses of N = 20 participants was already reported in [7]. Using as an example the data for avatar trials from that study (**Figure 5**), the analysis showed enhanced discrimination accuracy for face pairs that cross the category boundary in the between condition compared with attenuated discrimination accuracy for face pairs in the within condition. This is consistent with CP. The data show also that there is a significant difference in discrimination accuracy within the category in that there is greater discrimination accuracy for face pairs in the within condition than in the same condition. The variation in ISI of 75 and 300 msec differentially affected participants' responses, but not in the human trials.
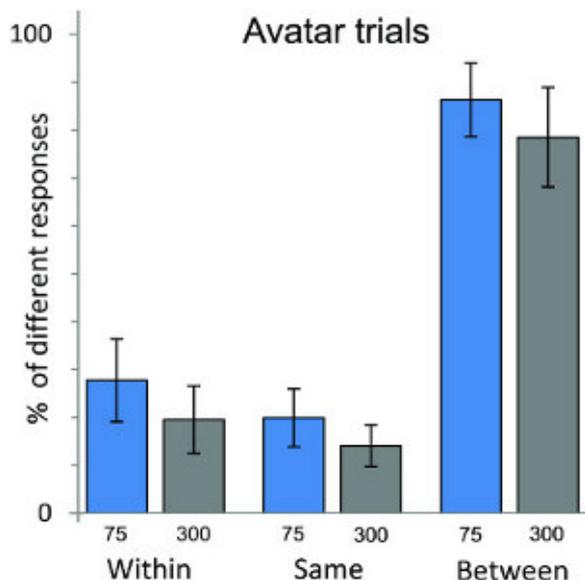
Figure 5. Results of the "same-different" perceptual discrimination task for avatar trials. Participants (N = 20). judged whether the morphs of a morph pair were the same or different in physical appearance. Controlling for relative distance of morphs along the continua, results show better discrimination accuracy for face pairs that crossed the category boundary (that was determined in the forced choice classification task) than for pairs drawn from the same (*i.e.* avatar or human) side of the boundary, thus demonstrating categorical perception along the continua of human likeness. The impact of a shorter and longer ISI of 75 msec and 300 msec was also tested and found to influence discrimination performance for avatar trials only. Error bars show ±1 standard error.

Using the *A'* statistic as a measure of discrimination performance independent of response bias, there was in the avatar trials a significant main effect on discrimination sensitivity of face-pair trial types (*i.e.* (within and between), $F(2,38) = 107.11$, $p < 0.001$, with greater discrimination sensitivity for cross-category (*A'* = 0.89, *SD* = 0.07) than for within-category pairs (*A'* = 0.55, *SD* = 0.17) (**Figure 6**). Similarly, there was significantly greater discrimination sensitivity for cross-category (*A'* = 0.94, *SD* = 0.1) than for within-category pairs (*A'* = 0.56, *SD* = 0.22) in the human trails, $F(2,38) = 107.11$, $p < 0.001$. There was no effect of face-pair trial types on ISI. Using the β''_D statistic as a measure of response bias, there was a significant main effect on bias of face-pair trial types [$F(2,38) = 70.53$, $p < 0.001$], with participants showing a strong tendency to judge within-category pairs as different (β''_D = 0.81, SD = 0.23) compared with the response to cross-category pairs (β''_D = -0.18, SD = 0.59). This is consistent with the idea that participants tend to favor "different" decisions in this particular task when the same-different decision is more difficult for within-category pairs.
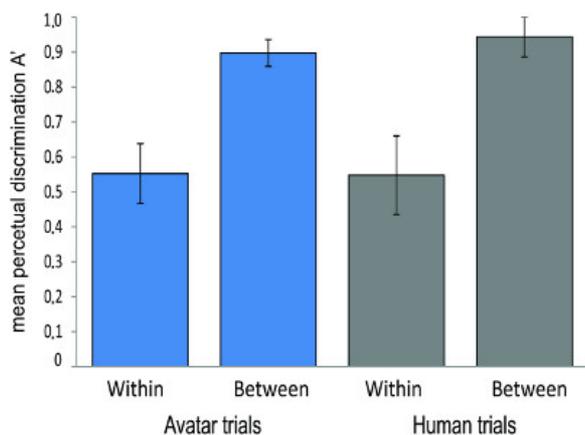
Figure 6. Using the A' statistic as a measure of discrimination performance independent of response bias (N = 20), discrimination sensitivity was greater for cross-category than for within-category pairs in both avatar and human trials. Error bars show ±1 standard error.

The analysis of RT data showed no differences between avatar and human trials and between short and long ISI. There was as expected a main significant effect for RT between the three stimulus pair conditions (see **Figure 7**), $F(2,38) = 34.55$, $p < 0.001$. Pre-planned tests of within-subject contrasts showed that RT for cross-category faces (*i.e.* 'between' face-pair trial type) were significantly faster (*M* = 0.79, *SE* = 0.05) than RT for face pairs from within a category ('within' trial type) (*M* = 1.26, *SE* = 0.09) [$F(1,19) = 60.09$, $p < 0.001$] and face pairs in the same face pair condition (*M* = 0.88, *SE* = 0.08), $F(1,19) = 43.1$, $p < 0.001$.
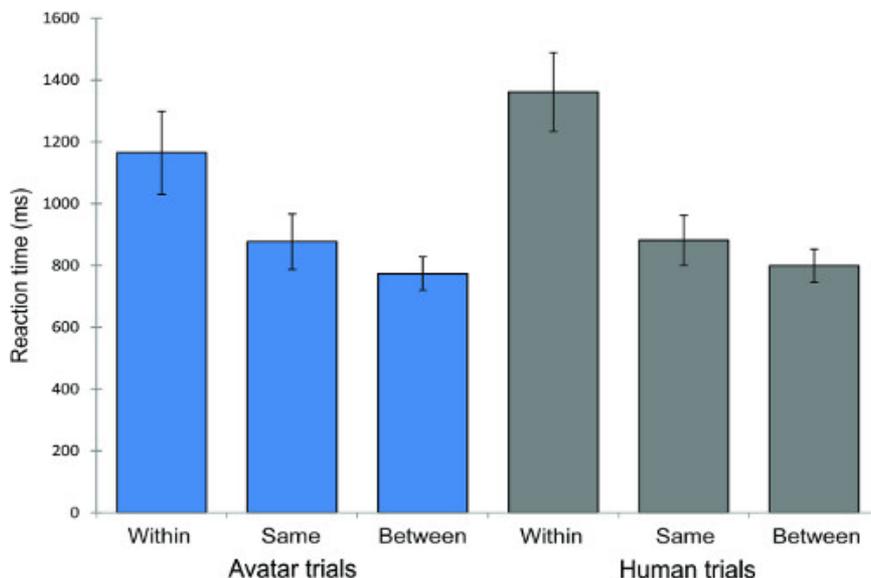
Figure 7. Reaction time (RT) results of the "same-different" perceptual discrimination task for avatar and human trials (N = 20). The graph shows that RT for stimulus pairs that cross the category boundary (*i.e.* in the between condition) were shorter than the RT for faces from within a category. Error bars show ±1 standard error.

The categorization response data thus confirm the second criterion for the presence of CP in that there is better discrimination accuracy for pairs that cross the category boundary than for equidistant pairs drawn from within a category. This demonstrates that there is a so-called discrimination boundary with enhanced sensitivity for the physical stimulus features close to the category boundary. The RT data support this in showing shorter response latencies for cross-category compared with with-category face pairs.

This particular perceptual discrimination task does not define the specific point of the discrimination boundary along the DHL. A much smaller morph distance between pairs of presented morphs could be used to resolve this. Here we show an example using a traditional ABX discrimination task [12, 13]. ABX discrimination entails sequential presentation of different face stimuli (*e.g.* Morph A and Morph B) followed by a second presentation of either A or B as the target stimulus X. After viewing images A, B and X, participants are required to indicate whether A or B is identical to X. In this example, a 2-step discrimination procedure between morphs (*i.e.* 1-3, 2-4, 3-5, etc.) is presented (**Figure 8B**). Analyses are described in [8]. For the purpose of illustration, the ABX discrimination task was performed on 24 participants using 4 morph continua, each with 11 morphs, using endpoint stimuli drawn from the study of Cheetham *et al.* [7]. Following the ABX discrimination task, a forced choice categorisation task was performed with the same participants. This sequence of task presentation is thought to minimize the influence of explicit category decision making on the ABX discrimination task. **Figure 8B** indicates clearly that there is a peak in perceptual discrimination sensitivity at the morph position predicted by and aligned with the category boundary (see **Figure 8A**). Using the 2-step distance between morphs, the peak in discrimination performance can be clearly identified in the interval between morph pair M5-M7. See [8] for findings using the ABX paradigm and morph stimuli drawn from dimensions of human likeness with monkey, cow and human faces as the endpoints of the continua.
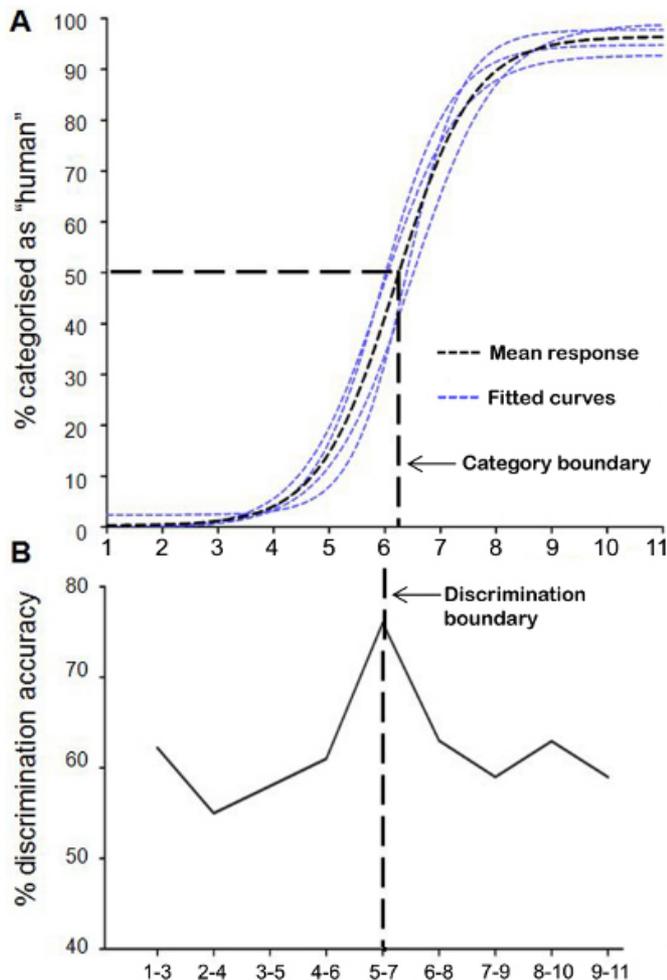
Figure 8. Representative results of the ABX perceptual discrimination and forced choice categorization tasks. The 2-step discrimination procedure (*i.e.* 1-3, 2-4, 3-5, etc.) in the ABX perceptual discrimination task in panel **B** shows that the peak in perceptual discrimination sensitivity is predicted by the category boundary determined in the forced choice categorization task shown in panel **A**. Panel **A** shows the logistic profile of the fitted regression curves of the four continua. Maximum uncertainty of 50% in categorization judgments of morphed faces as human is associated with morph M6.

The same-different discrimination task confirms that the third criterion for the presence of CP in showing that the discrimination boundary is aligned with the category boundary. In other words, the position of the category boundary predicts the position of the discrimination boundary.

The fourth criterion, which is not always applied in studies of CP [13, 14] is that discrimination is at chance within the categories. The data of the illustrative example using the ABX design would suggest that discrimination is slightly above chance for those morphs located between the continua endpoints and the category boundary.

3. fMRI task

4.3.1 Sensitivity to physical change

By comparing the conditions in which there is a physical change between the first and second morph with the condition in which there is no such change, a brain region in the fusiform gyrus (**Figure 9A**) is shown to be sensitive to the presentation of fine-grained change along the DHL in the physical appearance of face morphs in the avatar trials. A similar result for human trials is not shown in the figure. This region has been referred to as the fusiform face area because of its role as part of the visual system in processing facial information. Together with the human trials, this finding is consistent with the reported response of fusiform areas to differences in facial physical attributes [23], facial geometry [16, 21, 24] and facial texture [21].

4.3.2 Sensitivity to category change

Figure 9B shows, using the example of avatar trials, brain regions sensitive to category change along the DHL. This was achieved by comparing the conditions in which there is a category change between the first and second morph with the condition in which there is no such change. The imaging data show that category change in avatar trials (*i.e.* a change from avatar-to-human direction along the DHL) revealed responsiveness of the hippocampus, amygdala, and insula. The role of these regions needs to be interpreted in the context of the paradigm used and categorization and has already been described [7]. Generally, the amygdala is responsive to faces, affective valence, novelty, and uncertainty [55, 56, 57, 58, 59]. The amygdala is suggested to influence processing of other brain regions involved in categorization depending on the

affective meaning of a situation [60]. The insula is consistently reported in association with category processing and processing under conditions of uncertainty [61, 62, 63]. In the context of the paradigm used, this region might contribute to enhancing attentional resources for categorization processing [63]. The specific region of activation could also be associated with signaling the presence of uncertainty, threat, or potential threat [64, 65]. The hippocampus is involved in visual categorization and perceptual learning [66]. The category change in human trials (*i.e.* a change in the human-to-avatar direction along the DHL) revealed that the putamen, head of caudate, and thalamus, are responsive to this condition. Generally, these regions are associated with learning stimulus-category associations, signaling category membership, decision uncertainty during categorization, switching between potential category rules used to establish category membership and adjustment of the represented categorical boundary in order to minimize errors [67, 68, 69, 70].

Interpretation of these results at a broad level and within the context of the experimental paradigm used suggests that avatar and human faces represent different categorization problems depending on the degree of previous categorization experience with a given category (*e.g.* [25]); the participants are expert in human face processing but were especially selected on the basis that they report no explicit knowledge of previous experience with avatar faces (*e.g.* in video games, movies, second life) and, as confirmed at debriefing, had never previously seen faces of the kind we presented.
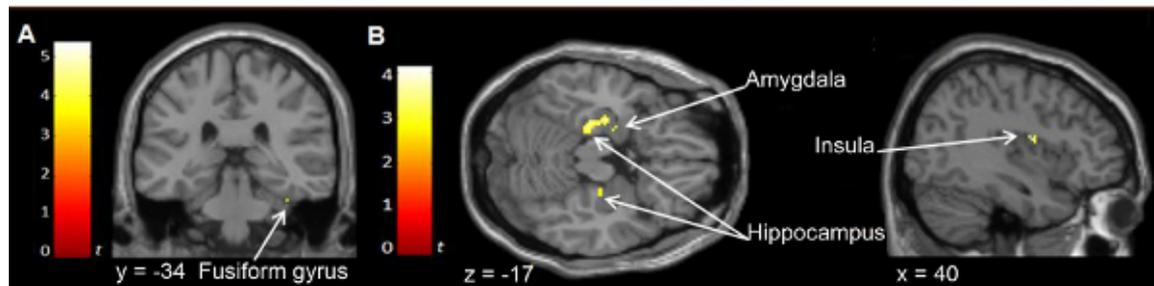
Figure 9. Neural correlates of physical and of category change along the DHL in avatar trials. The activation maps are superimposed on the coronal (**A**), transversal (**B**) and sagittal (**C**) views of a single subject. The color bars signify the gradient of *t* values of the activation maps (*p* < 0.005, 20 contiguous voxels).

Discussion

The core prediction of the uncanny valley hypothesis is that positively or negatively valenced experience can be evoked as a function of perceived human likeness [77] (for an informative overview, see [78]). Careful examination of how human likeness is actually perceived is in itself therefore an important research undertaking. Similarly important is how the DHL is represented in experiments of uncanny experience. This protocol focuses therefore on the DHL. One approach is to represent human likeness using morph continua, as already implemented in "uncanny" research [5, 6, 26, 27, 28]. The advantage of morph continua is that their use permits experimentally controlled differences in humanlike appearance to be brought into relationship with behavioural measures of subjective perception and experience (*e.g.* category decisions, uncanny feelings) and with underlying neural processes [7]. This fine-grained approach is particularly important because the uncanny valley hypothesis does not predict the actual degree of human likeness at which the transition between positively valenced and uncanny experience should occur [78]. If Mori's conjectures are correct, the findings relating to category processing along the DHL [7] would suggest that uncanny experience is most likely to occur at the category boundary where perceptual decision ambiguity is greatest. This has still to be tested.

To be able to interpret the investigated relationship between the DHL, as represented using morph continua, and other variables of interest, a single morph continuum rather than two or even three different juxtaposed continua should be used [5, 28]. The juxtaposed continua fail to represent and, in effect, alter Mori's concept of human likeness by introducing discontinuities to the DHL. This could affect performance in a perceptual discrimination task, because the point of the discontinuity and that of any disparities resulting from the morphing procedure might be used as a reliable but experimentally unintended point of reference for guiding perceptual discrimination (see, [29]). Within each morph continuum all morphs should be carefully controlled so that equivalent increments of physical change are represented along the entire continuum [5, 28]. This is especially important in this protocol, because experimental control of morph distance along the continua enables examination of whether the sensory information relating to linear differences in physical human-like similarity along the DHL is cognitively represented in a linear or nonlinear way. Nonlinearity is reflected in the step-like function in the slope of the categorization responses (**Figures 2A** and **5A**) and in differences in perceptual sensitivity to stimulus attributes along the DHL (see **Figures 4** and **5B**). This protocol uses faces as endpoints without applying any further experimental manipulations. Further studies of CP and human likeness could examine for example how specific features such as eye realism compared with the realism of other facial features or manipulations of facial geometry compared with facial texture (cf. [30,38]) differentially influence category processing along the DHL.

The morphing procedure enables smooth blending together of corresponding features of the continuum's endpoints such as facial configural cues. Difficulty in morphing facial information like upper facial features and hair profile [26] can potentially bias participant's responses by drawing attention to disparities in the alignment of features during the morphing procedure. This bias is likely to be systematic in that morphing disparities are related to the morph distance from the continua endpoints, the disparities being greatest at the midpoint of morph continua. For our morph continua, the midpoint of the continua corresponds with the category boundary around which there is greatest perceptual sensitivity. Reanalysis of data from one of our pilot studies (a forced choice categorisation task) compared continua in which the eye region was either well or poorly morphed (poor morphing resulted in a very slight inconsistency in the alignment of eye texture between morphs). The reanalysis confirmed a systematic bias in the categorisation decision responses of the poorly morphed continua such that poor morphing effectively caused a relative shift of the category boundary toward the human end of the dimension. This was presumably because the morphing disparity was perceived as a "nonhuman-defining" feature.

A response bias might result also from using continua generated on the basis of endpoint stimuli in which non-facial information such as head attire and facial jewelery are only present in one endpoint stimulus [27]. In this case, facial images could be cropped so that participants attend to the stimulus information of research interest rather than to other salient features presented in an image. A systematic response bias can result also from using an image as a continuum endpoint in which nonhuman attributes are presented together with human attributes, even though this image is intended to represent the human end of the DHL [6]. In this case, any relationship between human likeness and variables such as subjective measures of uncanny experience are not interpretable in terms of Mori's conception of the DHL and of the hypothesized uncanny valley.

CP can occur along dimensions other than human likeness [31, 10, 22, 32, 33, 34, 35], and category-relevant information can be automatically processed upon exposure to others [36]. In this protocol, care should be taken therefore to control for the effects of visual cues indicating differences along the DHL in terms of other category-relevant dimensions on participants' responses concerning human likeness. These cues might for example relate to ethnicity, gender, facial distinctiveness, familiarity and identity, and facial expression (cf. [5, 26, 27, 28]). The present protocol seeks to minimize perception of biological motion between face morphs presented in rapid succession in the perceptual discrimination task and fMRI study by closely matching the facial geometry and configuration of facial features of images used as continuum endpoints. This approach (together with the relative position along the continua of morphs used in the stimulus conditions) helps also to minimize any perception of different identities between morphs of a continuum.

The *forced choice classification task* determines which morphs of a continuum are clearly categorized as an avatar and as human in order to select morphs for use in the perceptual discrimination task and the fMRI study. We selected the four morphs M0, M4, M8 and M12 from each of the continua (**Figures 2B** and **2C**). In addition to controlling for the degree of physical change along the DHL, the choice of M4 and M8 is based on the following theoretical consideration. Mori described perceptual uncertainty (and associated uncanny experience) as occurring at levels of realism that correspond to the region along the DHL between the two positive peaks in the slope of the valence-human likeness relationship (see **Figure 1**). At these peaks, objects are regarded as either nonhuman or human. In reframing his considerations in terms of the framework of category processing, these peaks may be seen as reflecting degrees of human likeness at which correctly classified category instances (*i.e.* nonhuman and human) straddle the category boundary. But Mori did not specify how efficient this classification (*i.e.* perceptual certainty) must be at these peaks, though the identification of objects at each peak is clearly considered to be relatively efficient and effortless. For this reason, the two morph positions along the continua considered as defining the transition between the two categories and as reflecting the two positive peaks was determined using a more conservative criterion than often otherwise used in CP research (*e.g.* 66%, as in [32, 34]). Thus, morph M4 was identified on average as an avatar in more than 85% of trials and morph M8 as a human in more than 85% of trials. Please note that this criterion applies to both morphs M4 and M8 of any one continuum. Using this approach, this choice of morphs seeks to capture a sense of category change along the DHL between nonhuman and human objects in accordance with both an understanding of CP and Mori's description of the hypothesis.

This protocol uses a variant of the *same-different perceptual discrimination task* [10] to examine CP. The advantage of this task is that participants do not need a description as to what specific similarities and differences must be identified. It is sufficient that they simply identify stimuli as being the same or different. In addition, participants do not need to know the category labels. Labels might be used as a strategy to discriminate between stimuli when the memory load required by a discrimination task such as the ABX task increases [42]. The same-different task has the advantage that the memory load is comparatively low and that the task encourages direct comparison of stimuli. To reduce the potential influence of labelling, discrimination tasks are normally presented before the forced choice decision task [40]. The present protocol is based on two different participant groups for the discrimination and forced choice decision tasks [7, 41]. This is because the forced choice task is used to select stimuli for the discrimination task. Should however the same participants be tested in both tasks, the protocol should be modified so that the discrimination task is conducted before the forced choice decision task.

A fixed discrimination design is applied in the same-different discrimination task of this protocol (for roving designs, see *e.g.* [39]). This means that M4 and M8 are always shown as the first stimulus of each stimulus pair in the "same", "within" and "between" conditions of the avatar and human trials, respectively. This protocol includes the experimental constraint that each participant views only the morph stimuli of either avatar or human trials from a given continuum but not both. Using the avatar trials as an example, this means that the first stimulus of each stimulus pair is always M4, that the second stimuli in the "within" (*i.e.* M1) and "between" (*i.e.* M8) conditions are presented equally often for a given continuum, and that no further stimuli are drawn for human trials from that particular continuum. This approach aims to avoid selectively inducing stronger representation of and facilitating therefore discrimination of the cross-category faces of a given continuum. To exclude or, for purposes of comparison, to investigate any possible effect on cross-category representation and discrimination of presenting the described avatar and human trials in one experimental block, a design could be implemented in which the described avatar and human trials are presented in separate blocks (with blocks counterbalanced in order across participants).

The present same-different discrimination task has a ratio of same-to-different trials of 1:2. This ratio might induce a response bias in favour of "different" decisions (though other factors can also influence this bias [44, 51]). Measures derived from Signal Detection Theory (SDT) are often used to disentangle response bias (β or c) for selecting one response over another from the participant's sensitivity (A' or d') in discriminating sensory stimuli (for an overview see, [44]). As d' can vary with response bias due to violation of SDT assumptions [52], we used the nonparametric measure of sensitivity A' [53]. For response bias we used β''_D [47]. Alternatively c has been recommended by [43, 44], partly because it is independent of change in d' [54]. Overall, the present results indicate greater perceptual sensitivity for morph stimuli straddling the category boundary than for within-category stimuli.

The selection of morphs for the discrimination task in this protocol means that the task requires discrimination between morphs that are four steps apart along the continua (*i.e.* a four-step discrimination, see **Figure 2B**). But this four-step degree of dissimilarity between morphs is too large to allow better specification of the actual morph position at which discrimination is most enhanced (*i.e.* the discrimination boundary) (**Figure 5B**). An important criterion for CP (for the other criteria, see *e.g.* [11]) is that there is alignment between the category boundary in the forced choice task and the discrimination boundary in the discrimination task. In other words, the morph position of the category boundary should predict the morph position of the discrimination boundary. One approach to verifying the specific point of alignment would be to use a discrimination task in which the morph distance between pairs of morphs is reduced. For the purpose of illustration, **Figure 5B** shows results of pilot data using, as a possible alternative to the same-different discrimination task, a traditional ABX discrimination task [12, 13]. The figure indicates clearly that there is a peak in perceptual discrimination sensitivity at the morph position predicted by the category boundary. Such results in a study with a larger

number of participants and application of SDT in analyses would further verify the finding of effects of CP along the DHL. The actual choice of stimuli for the continua endpoints, the number of morphs generated in a continuum, and the size of the step in morphs to be discriminated will strongly influence the cognitive demands placed on the participant and his or her ability to discriminate morphs along the continua.

One classical criterion of CP is that the position of the category boundary predicts the position of the peak in actual discrimination performance (*i.e.* the discrimination boundary) [80]. This is arguably the most important criterion of CP [81]. Conclusive testing of this prediction requires an experimental design in which all morph pairs that together represent the entire length of the morph continuum are presented in the discrimination task in order to determine the actual position of the peak. In [38], discrimination performance was examined on the basis of only certain segments of the morph continua. This could mean that the true position of the actual peak in performance may have been missed, this in turn rendering it difficult to conclusively verify CP. It should be noted that even the early CP study of Lieberman *et al.* [82] failed to meet the studies own stringent criterion that predicted and actual peak in discrimination performance converge, and that other researchers have not applied this criterion stringently (*e.g.* [11], see also [80]). Determining the actual position of peak performance is nevertheless critical, even if a more liberal interpretation of this criterion is applied. Examining the entire length of the morph continuum also has the advantage of enabling inspection of the data as to whether there is a peak in performance at a point contrary to expectation due for example to an artefact resulting from the morphing procedure.

In addition to the responses, the response time (RT) data in the forced choice classification task is useful as an indicator of difficulty in cognitive processing of stimulus information and of the competing response tendencies to categorize a stimulus as "avatar" or "human" [70, 71]. RT should thus be longest for categorization judgments of stimuli positioned at or nearest to the category boundary. **Figure 4** shows that this is the case. Taken together, the shape of the response function and the RT data for category judgments show that assignment of a stimulus to a discrete category is subject to large differences in processing difficulty. To assess RT, this protocol instructs participants to respond during categorization as quickly and accurately as possible. Given the potential impact of a speed-accuracy trade-off on responses [72, 73], we examined and found in pilot testing that the shape and position of the avatar-human category response function is very robust, being unaffected by instructions to identify the presented morph stimulus either as quickly and accurately as possible or simply as accurately as possible. This would suggest that participants generally use a decision strategy weighted for accuracy, though this suggestion could be tested more thoroughly. In keeping with Mori's hypothesis that difficulty in distinguishing a humanlike object from the human image might evoke negatively valenced experience, it would be interesting to establish whether longer RT for humanlike stimuli is associated with measures of negative affect. RT data was also collected and analyzed for the same-different discrimination task. RT has been used to support response data [80]. In contrast to the ABX task, the same-different task provides a clear time point for RT measurement. The RT of correct responses should be shorter for between- than for within-pairs [74], though the interpretation of RT data can be complicated for same-different judgments because RT can be influenced by a number of factors in this task [75, 76]. The RT data are however consistent with the idea that less difficult cross-category decisions are made more quickly than within-category decisions (see **Figure 7**).

It should be pointed out that Mori's hypothesis does not consider the possibility that physical features might actually vary along the DHL within the human category (**Figure 2**) [7]. This is the reason why the second positive peak in the hypothesis' original valence-human likeness relationship is located at the human end of the DHL (**Figure 1**). The emphasis on the nonhuman aspect of the DHL has been influential in studies guided by the hypothesis, including studies that have not used morph continua [4, 37], while other studies have used a single human face to represent the human aspect of the DHL [3]. Such studies have sought to examine uncanny experience, with unclear results. The findings relating to CP suggest that these studies might not have presented the stimuli needed to evoke implicit or explicit processes of perceptual decision making and processes of conflict resolution in response to category ambiguity along the DHL.

This protocol illustrates an example of how morphs drawn from continua representing the DHL can be used to identify, with fMRI and using the effect of repetition suppression, brain regions sensitive to change in physical humanlike similarity and to change in category-related information. The effectiveness of the fMRI design is influenced strongly by careful generation and selection of the morph stimuli. The forced choice and perceptual discrimination tasks were thus used to ensure comparability between continua in the shape of the avatar-human classification curves (*i.e.* slope of the response function) and in discrimination performance. The advantage of this fMRI design is that it allows the stimulus conditions described by Mori (*i.e.* passive observation of novel non-human objects that are subtly different in physical appearance from that of their human counterpart) to be simulated within the constraints of fMRI methodology, using stimuli selected according to the hypothesis' definition of human likeness, and investigation of effects of category processing while controlling for effects of physical change along the DHL. The fMRI paradigm is not designed to examine uncanny experience, but it could be adapted to investigate affective experience associated for example with the category boundary itself. This would be an important step toward examining in the brain the effects of category processing and category ambiguity in association with affective experience for stimuli drawn from the DHL.

Disclosures

No conflicts of interest declared.

Acknowledgements

References

1. Mori, M. Bukimi no tani [The uncanny valley]. *Energy.* **7**, 33-35 (1970).
2. MacDorman, K. Androids as an experimental apparatus: Why is there an uncanny valley and can we exploit it? *Toward Social Mechanisms of Android Science: A CogSci 2005 Workshop.*, 106-118 (2005).
3. Tinwell, A., Grimshaw, M., & Williams, A. The uncanny wall. *International Journal of Arts and Technology.* **4**, 326-341 (2011).
4. Tinwell, A. *Uncanny as usability obstacle. HCI.* **12**, 622-631 (2009).

5. MacDorman, K.F. & Ishiguro, H. The uncanny advantage of using androids in cognitive and social science research. *Interaction Studies.* **7**, 297-337 (2006).
6. Hanson, D., Olney, A., Prillman, S., Mathews, E., Zielke, M., Hammons, D., Fernandez, R., & Stephanou, H.E. Upending the uncanny valley. *20th National Conf. on Artificial Intelligence and the 17th Innovative Applications of Artificial Intelligence Conf.*, 1728-1729 (2005).
7. Cheetham, M., Suter, P., & Jancke, L. The human likeness dimension of the "uncanny valley hypothesis": behavioral and functional MRI findings. *Front Hum. Neurosci.* **5**, 126 (2011).
8. Campbell, R., Pascalis, O., Coleman, M., Wallace, S.B., & Benson, P.J. Are faces of different species perceived categorically by human observers? *Proc. Biol. Sci.* **264**, 1429-1434 (1997).
9. Finney, D.J. *Probit analysis.*, (1964).
10. Angeli, A., Davidoff, J., & Valentine, T. Face familiarity, distinctiveness, and categorical perception. *Q.J. Exp. Psychol. (Hove).* **61**, 690-707 (2008).
11. Studdert-Kennedy, M., Liberman, A.M., Harris, K.S., & Cooper, F.S. Theoretical notes. Motor theory of speech perception: a reply to Lane's critical review. *Psychol. Rev.* **77**, 234-249 (1970).
12. Liberman, A.M., Hariris, K.S., Hoffman, H.S., & Griffith, B.C. The discrimination of speech sounds within and across phoneme boundaries. *J. Exp. Psychol.* **54**, 358-368 (1957).
13. Harnad, S.R. Introduction: Psychological and cognitive aspects of categorical perception: A critical overview. In: *Categorical perception: The groundwork of cognition.*, Harnad, S.R., ed., Cambridge University Press, New York, Haruno & Kawato, 1-25 (1987).
14. Repp, B.H. Categorical perception: Issues, methods, findings. *Speech Lang. Adv. Basic Res. Pract.* **10**, 243-335 (1984).
15. Grill-Spector, K., Henson, R., & Martin, A. Repetition and the brain: neural models of stimulus-specific effects. *Trends. Cogn. Sci.* **10**, 14-23 (2006).
16. Jiang, X., Rosen, E., Zeffiro, T., Vanmeter, J., Blanz, V., & Riesenhuber, M. Evaluation of a shape-based model of human face discrimination using FMRI and behavioral techniques. *Neuron.* **50**, 159-172 (2006).
17. Henson, R.N. Neuroimaging studies of priming. *Prog. Neurobiol.* **70**, 53-81 (2003).
18. Grill-Spector, K. & Malach, R. fMR-adaptation: a tool for studying the functional properties of human cortical neurons. *Acta. Psychol. (Amst).* **107**, 293-321 (2001).
19. Jiang, X., Bradley, E., Rini, R.A., Zeffiro, T., Vanmeter, J., & Riesenhuber, M. Categorization training results in shape- and category-selective human neural plasticity. *Neuron.* **53**, 891-903 (2007).
20. van der Linden, M., van Turennout, M., & Indefrey, P. Formation of category representations in superior temporal sulcus. *J. Cogn. Neurosci.* **22**, 1270-1282 (2010).
21. Jiang, F., Dricot, L., Blanz, V., Goebel, R., & Rossion, B. Neural correlates of shape and surface reflectance information in individual faces. *Neuroscience.* **163**, 1078-1091 (2009).
22. Rotshtein, P., Henson, R.N., Treves, A., Driver, J., & Dolan, R.J. Morphing Marilyn into Maggie dissociates physical and identity face representations in the brain. *Nat. Neurosci.* **8**, 107-113 (2005).
23. Xu, X., Yue, X., Lescroart, M.D., Biederman, I., & Kim, J.G. Adaptation in the fusiform face area (FFA): image or person? *Vision Res.* **49**, 2800-2807 (2009).
24. Liu, X., Steinmetz, N.A., Farley, A.B., Smith, C.D., & Joseph, J.E. Mid-fusiform activation during object discrimination reflects the process of differentiating structural descriptions. *J. Cogn. Neurosci.* **20**, 1711-1726 (2008).
25. Poldrack, R.A., Prabhakaran, V., Seger, C.A., & Gabrieli, J.D. Striatal activation during acquisition of a cognitive skill. *Neuropsychology.* **13**, 564-574 (1999).
26. Ho, C.C., MacDorman, K.F., & Pramono, Z.A.D. Human emotion and the uncanny valley: A GLM, MDS, and isomap analysis of robot video ratings. *In ACM/IEEE international conference on human-robot interaction.*, 169-176 (2008).
27. Seyama, J. & Nagayama, R.S. The Uncanny Valley: Effect of Realism on the Impression of Artificial Human Faces. *Presence: Teleoperators and Virtual Environments.* **16**, 337-351 (2007).
28. Hanson, D. Exploring the aesthetic range for humanoid robots. *Paper presented at the CogSci Workshop: Toward Social Mechanisms of Android Science.*, (2006).
29. Pastore, R.E. Categorical perception: Some psychophysical models. *In Categorical perception: The groundwork of cognition.*, Harnad, S., ed., New York, 29-52 (1987).
30. MacDorman, K.F., Green, R.D., Ho, C.C., & Koch, C.T. Too real for comfort? Uncanny responses to computer generated faces. *Computers in Human Behavior.* **25**, 695-710 (2009).
31. Levin, D.T. 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.* **129**, 559-574 (2000).
32. Beale, J.M. & Keil, F.C. Categorical effects in the perception of faces. *Cognition.* **57**, 217-239 (1995).
33. Calder, A.J., Young, A.W., Perrett, D.I., Etcoff, N.L., & Rowland, D. Categorical perception of morphed facial expressions. *Visual Cognition.* **3**, 81-117 (1996).
34. Etcoff, N.L. & Magee, J.J. Categorical perception of facial expressions. *Cognition.* **44**, 227-240 (1992).
35. Campanella, S., Quinet, P., Bruyer, R., Crommelinck, M., & Guerit, J.M. Categorical perception of happiness and fear facial expressions: an ERP study. *J. Cogn. Neurosci.* **14**, 210-227 (2002).
36. Macrae, C.N. & Bodenhausen, G.V. Social cognition: thinking categorically about others. *Annu. Rev. Psychol.* **51**, 93-120 (2000).
37. Schneider, E., Wang, Y., & Yang S. Exploring the Uncanny Valley with Japanese Video Game Characters. *Conference proceedings of DIGRA 2007.*, SMU Press, (2007).
38. Looser, C.E. & Wheatley, T. The tipping point of animacy: How, when, and where we perceive life in a face. *Psychological Science.* **21**, 1854-1862 (2010).
39. Macmillan, N.A., Goldberg, R.F., & Braida, L.D. Resolution for Speech Sounds: Basic sensitivity and context memory on vowel and consonant continua. *Journal of the Acoustical Society of America.* **84**, 1262-1280 (1988).
40. Newell, F.N., & Bulthoff, H.H. Cognition Categorical perception of familiar objects. *Cognition.* **85**, 113-143 (2002).
41. Rotshtein, P., Henson, R.N., Treves, A., Driver, J., & Dolan, R.J. Morphing Marilyn into Maggie dissociates physical and identity face representations in the brain. *Nat. Neurosci.* **8**, 107-113 (2005).
42. Massaro, D.W. & Cohen, M.M. Categorical or continuous speech perception: A new test. *Speech Communication.* **2**, 15-35 (1983).
43. Snodgrass, J.G. & Corwin, J. Pragmatics of measuring recognition memory: Applications to dementia and amnesia. *Journal of Experimental Psychology: Genera.* **117**, 34-50 (1988).

44. Macmillan, N.A. & Creelman, C.D., Jr. *Detection theory: A user's guide.*, Cambridge University Press, Cambridge, (1991).
45. Snodgrass, J.G., Levy-Berger, G., & Haydon, M. *Human experimental psychology.*, Oxford University Press, New York (1985).
46. Donaldson, W. Accuracy of d' and A' as estimates of sensitivity. *Bulletin of the Psychonomic Society.* **31**, 271-274 (1993).
47. Donaldson, W. Measuring recognition memory. *Journal of Experimental Psychology: General.* **121**, 275-278 (1992).
48. Stanislaw, H. & Todorov, N. Calculation of signal detection theory measures. *Behavior Research Methods, Instruments, & Computers.* **31**, 137-149 (1999).
49. Wright, D.B., Horry, R., & Skagerberg, E.M. Functions for traditional and multilevel approaches to signal detection theory. *Behavior Research Methods.* **41**, 257-267 (2009).
50. Sorkin, R.D. Spreadsheet signal detection. *Behavior Research Methods, Instruments, & Computers.* **31** (1), 6-54 (1999).
51. Verde, M.F., Macmillan, N.A., & Rotello, C.M. Measures of sensitivity based on a single hit rate and false-alarm rate: The accuracy, precision, and robustness of d', Az, and A'. *Perception & Psychophysics.* **68**, 643-654 (2006).
52. Swets, J.A. Form of empirical ROCs in discrimination and diagnostic tasks: Implications for theory and measurement of performance. *Psychological Bulletin.* **99**, 181-198 (1986).
53. Pollack, I. & Norman, D.A. A nonparametric analysis of recognition experiments. *Psychonomic Science.* **1**, 125-126 (1964).
54. Macmillan, N.A. Signal detection theory as data analysis method and psychological decision model. In: *A handbook for data analysis in the behavioral sciences: Methodological issues.*, Keren, G. & Lewis, C., Eds., Erlbaum, Hillsdale, NJ, 21-57 (1993).
55. Todorov, A. & Engell, A. The role of the amygdala in implicit evaluation of emotionally neutral faces. *Soc. Cogn. Affect. Neurosci.* **3**, 303-312 (2008).
56. Phelps, E.A. & LeDoux, J.E. Contributions of the amygdala to emotion processing: from animal models to human behavior. *Neuron.* **48**, 175-187 (2005).
57. Herwig, U., Kaffenberger, T., Baumgartner, T., & Jancke, L. Neural correlates of a "pessimistic" attitude when anticipating events of unknown emotional valence. *Neuroimage.* **34**, 848-858 (2007).
58. Levy, I., Snell, J., Nelson, A.J., Rustichini, A., & Glimcher, P.W. Neural representation of subjective value under risk and ambiguity. *J. Neurophysiol.* **103**, 1036-1047 (2010).
59. Neta, M. & Whalen, P.J. The primacy of negative interpretations when resolving the valence of ambiguous facial expressions. *Psychol. Sci.* **21**, 901-907 (2010).
60. Seger, C.A. & Miller, E.K. Category learning in the brain. *Annu. Rev. Neurosci.* **33**, 203-219 (2010).
61. Volz, K.G., Schubotz, R.I., & von Cramon, D.Y. Predicting events of varying probability: uncertainty investigated by fMRI. *Neuroimage.* **19** (1), 271-280 (2003).
62. Grinband, J., Hirsch, J., & Ferrera, V.P. A neural representation of categorization uncertainty in the human brain. *Neuron* **49**, 757-763 (2006).
63. Heekeren, H.R., Marrett, S., & Ungerleider, L.G. The neural systems that mediate human perceptual decision making. *Nat. Rev. Neurosci.* **9**, 467-479 (2008).
64. Gray, M.A. & Critchley, H.D. Interoceptive basis to craving. *Neuron.* **54**, 183-186 (2007).
65. Wager, T.D., Phan, K.L., Liberzon, I., & Taylor, S.F. Valence, gender, and lateralization of functional brain anatomy in emotion: a meta-analysis of findings from neuroimaging. *Neuroimage.* **19**, 513-531 (2003).
66. Poldrack, R.A. & Rodriguez, P. How do memory systems interact? Evidence from human classification learning. *Neurobiol. Learn. Mem.* **82**, 324-332 (2004).
67. Seger, C.A. & Cincotta, C.M. The roles of the caudate nucleus in human classification learning. *J. Neurosci.* **25**, 2941-2951 (2005).
68. Seger, C.A., Peterson, E.J., Cincotta, C.M., Lopez-Paniagua, D., & Anderson, C.W. Dissociating the contributions of independent corticostriatal systems to visual categorization learning through the use of reinforcement learning modeling and Granger causality modeling. *Neuroimage.* **50**, 644-656 (2010).
69. Filoteo, J.V., Maddox, W.T., Salmon, D.P., & Song, D.D. Information-integration category learning in patients with striatal dysfunction. *Neuropsychology.* **19**, 212-222 (2005).
70. Bonnet, C., Fauquet Ars, J., & Estaún Ferrer, S. Reaction times as a measure of uncertainty. *Psicothema.* **20** (1), 43-8 (2008).
71. Hyman, R. Stimulus information as a determinant of reaction time. *J. Exp. Psychol.* **45**, 188-196 (1953).
72. Wickelgren, W.A. Speed-accuracy tradeoff and information processing dynamics. *Acta. Psychologica.* **41**, 67-85 (1977).
73. Zacksenhouse, M. & Bogacz, R., Holmes P. Robust versus optimal strategies for two-alternative forced choice tasks. *Journal of Mathematical Psychology.* **54** (2), 230-246 (2010).
74. Campanella, S., Chrysochoos, A., & Bruyer, R. Categorical perception of facial gender information: Behavioural evidence and the face-space metaphor. *Visual Cognition.* **8**, 237-262 (2001).
75. Farrell, B. 'Same - different' judgments: A review of current controversies in perceptual comparisons. *Psychological Bulletin.* **98**, 419-456 (1985).
76. Palmer, J., Huk, A.C., *et al.* The effect of stimulus strength on the speed and accuracy of a perceptual decision. *J. Vision.* **5**, 376-404 (2005).
77. Mori, M. The uncanny valley., MacDorman, K.F. & Norri Kageki, Trans. *IEEE Robotics and Automation.* **19** (2), 98-100, Original work published in 1970, (2012).
78. MacDorman, K., Green, R., Ho, C.-C., & Koch, C. Too real for comfort? Uncanny responses to computer generated faces. *Comput. Hum. Behav.* **25**, 695-710 (2009).
79. Grier, J.B. Nonparametric indexes for sensitivity and bias - computing formulas. *Psychological Bulletin.* **75** (6), 424-429 (1971).
80. Repp, B.H. Categorical perception: Issues, methods and findings. In: *Speech and language: Advances in basic research and practice.*, Lass, N., Ed., **vol. 10**, Academic Press, Orlando, FL, 244-335 (1984).
81. de Gelder, B., Teunisse, J.P., & Benson, P.J. Categorical perception of facial expressions: Categories and their internal structure' *Cognition and Emotion.* **11** (1), 1-23 (1997).
82. Liberman, A.M., Harris, K., Hoffmann, H.S., & Griffith, B. The discrimination of speech sounds within and across phoneme boundaries. *J. Exp. Psychol.* **54**, 358-368 (1957).