Familiar faces rendered strange: Why inconsistent realism drives characters into the uncanny valley

Debaleena Chattopadhyay
Department of Computer Science,
College of Engineering, University of Illinois at Chicago,
Chicago, IL, USA

Karl F. MacDorman
Department of Human-Centered Computing,
Indiana University School of Informatics and Computing,
Indianapolis, IN, USA

Computer-modeled characters resembling real people sometimes elicit cold, eerie feelings. This effect, called the uncanny valley, has been attributed to uncertainty about whether the character is human or living or real. Uncertainty, however, neither explains why anthropomorphic characters lie in the uncanny valley nor their characteristic eeriness. We propose that realism inconsistency causes anthropomorphic characters to appear unfamiliar, despite their physical similarity to real people, owing to perceptual narrowing. We further propose that their unfamiliar, fake appearance elicits cold, eerie feelings, motivating threat avoidance. In our experiment, 365 participants categorized and rated objects, animals, and humans whose realism was manipulated along consistency-reduced and control transitions. These data were used to quantify a Bayesian model of categorical perception. In hypothesis testing, we found reducing realism consistency did not make objects appear less familiar, but only animals and humans, thereby eliciting cold, eerie feelings. Next, structural equation models elucidated the relation among realism inconsistency (measured objectively in a two-dimensional Morlet wavelet domain inspired by the primary visual cortex), realism, familiarity, eeriness, and warmth. The fact that reducing realism consistency only elicited cold, eerie feelings toward anthropomorphic characters, and only when it lessened familiarity, indicates the role of perceptual narrowing in the uncanny valley.

Introduction

Resolving category uncertainty and perceptual mismatch theories

Human replicas in computer-animated films and in robotics prompt us with increasing frequency to examine how they make us feel. These replicas may elicit cold, eerie feelings because of the perceived salience of their nonhuman features. Mori (1970/2012) graphed this effect as a valley of uncanniness in an otherwise positive relation between human likeness and affinity (Mathur & Reichling, 2016).

Kätsyri, Förger, Mäkäräinen, and Takala (2015) identified two main groups of theories in the uncanny valley literature: category uncertainty and perceptual mismatch (also reviewed in MacDorman & Chattopadhyay, 2016). While the former concerns the entity as a whole, the latter concerns relations among its features (Moore, 2012; Pollick, 2010). Category uncertainty theories propose uncanniness is caused by *doubt about what an entity is*, such as whether it is human or nonhuman, living or inanimate, real or simulated (Jentsch, 1906/1997; MacDorman & Ishiguro, 2006). Category uncertainty theories, as broadly construed, include explanations based on categorical perception (Burleigh, Schoenherr, & Lacroix, 2013; Cheetham, Pavlovic, Jordan, Suter, & Jäncke, 2013; Looser & Wheatley, 2010), category ambiguity (Burleigh & Schoenherr, 2015), conflicting representations (Ferrey, Burleigh, & Fenske, 2015), cognitive dissonance (Hanson, 2005; MacDorman, Green, Ho, & Koch, 2009a; MacDorman, Vasudevan, & Ho, 2009b), balance theory (Tondu & Bardou, 2011), sorites paradoxes (Ramey, 2005), and categorization difficulty (Cheetham, Wu, Pauli, & Jäncke, 2015; Yamada, Kawabe, & Ihaya, 2013).

Perceptual mismatch theories propose uncanniness is caused by a mismatch in the human likeness of an entity's features, such as in their visual or multimodal realism (e.g., human skin paired with computer-modeled eyes, MacDorman et al., 2009a; robot head with human voice, Meah & Moore, 2014; Mitchell et

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

doi: 10.1167/16.11.7 Received March 4, 2016; published September 9, 2016 ISSN 1534-7362

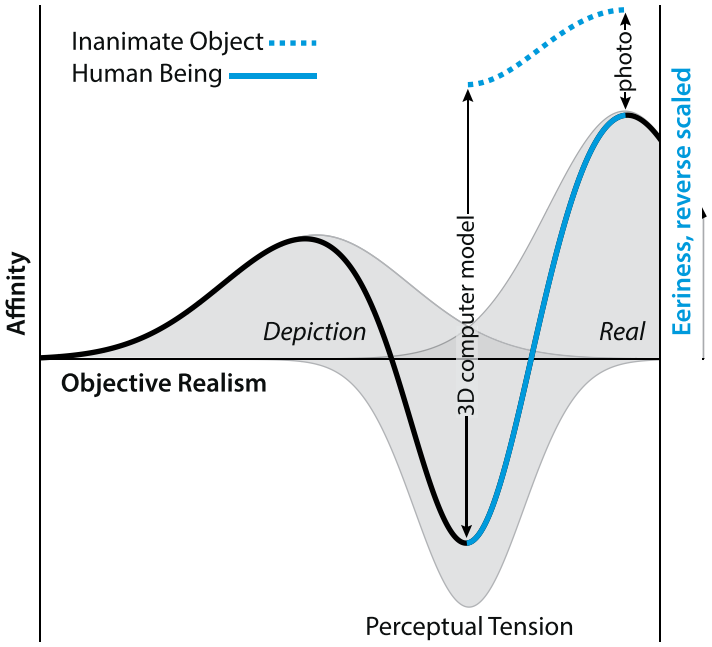

Figure 1. Moore (2012) proposed that affinity for a stimulus is its objective familiarity (probability of occurrence) minus its perceptual tension, weighted by the viewer's sensitivity to perceptual tension. The affinity curve resembles Mori's (1970/2012) uncanny valley graph. This study situates Moore's model for the observational dimension *objective realism* (fraction of real): Real entities are more familiar than their artistic depictions and thus are predicted to engender greater affinity; however, stimuli lying between *depiction* and *real* may have some features that appear more real than others, thus causing perceptual tension. This study found reverse-scaled eeriness ratings of human faces transitioning from *computer modeled* to *real* arose from a valley of eeriness with the predicted curve (solid blue); by contrast, inanimate objects cleared the valley completely (dashed blue).

al., 2011; or human skin with enlarged eyes, Seyama & Nagayama, 2007). We have previously proposed realism inconsistency theory, a kind of perceptual mismatch theory, as an explanation of the uncanny valley. Realism inconsistency theory predicts viewers will experience cold, eerie feelings only when perceiving anthropomorphic entities that possess features at different levels of realism (see Figure 1; MacDorman & Chattopadhyay, 2016). Our study found that increasing an entity's overall category uncertainty did *not* cause cold, eerie feelings. These feelings were instead caused by reducing consistency in the realism of its features. Kätsyri and colleagues (2015) likewise found scant support for category uncertainty theories but strong support for perceptual mismatch theories. We assume that any computer model that can be distinguished from reality is inherently inconsistent in realism because the achievement of realism varies in difficulty by feature, thus resulting in some features appearing more realistic than others.

Moore (2012) proposed a Bayesian model that resolves the apparent incompatibility of category uncertainty and perceptual mismatch theories. Cold, eerie feelings are not caused by uncertainty about an entity's category but by unequal levels of uncertainty about the category of its features. Moore's proposal is based on categorical perception: A physical difference in a feature at a category boundary appears much larger than an equal-sized difference within a category (viz., the perceptual magnet effect; Feldman, Griffiths, & Morgan, 2009). In simultaneously perceiving features of an entity that are closer to and farther from a category boundary, categorical perception causes differences in perceptual distortion among features, which Moore defines as *perceptual tension*. In other words, perceptual tension results from features that differ in their level on a perceptual dimension, such as humanness, animacy, or realism. Although Moore attributes the valley of negative affinity in Mori's graph to perceptual tension, he explains the broader positive trend using an objective measure of familiarity, the probability of occurrence of the stimulus. Thus, a viewer's affinity for the stimulus is graphed by subtracting from its probability of occurrence the perceptual tension it elicits, weighted by the viewer's sensitivity to perceptual tension (Figure 1; Equation 1 in the Method section of Quantitative evaluation of the Bayesian model).

Operationalizing shinwakan

Moore's (2012) model addresses another conundrum—the multidimensionality of the dependent var-

iable in the uncanny valley graph. The dependent variable must measure not only affinity for the humanlike but also the uncanniness of the nearly human—its characteristic eeriness—because this is the experience a theory of the uncanny valley must explain (Mangan, 2015).

In his original essay, Mori (1970/2012) referred to the dependent variable as *shinwakan*, which is a Japanese neologism. *Shinwakan* has been variously translated as *familiarity, affinity, comfort level, likability,* and *rapport* (Ho & MacDorman, 2010). However, as MacDorman and Ishiguro (2006) observed, it is incoherent to conceive of the dependent variable simply as familiarity, because the zero crossing in the graph is total novelty, and an entity cannot cross below total novelty into negative familiarity. Given the conceptual problems with operationalizing *shinwakan* as familiarity, it is unsurprising that Cheetham and colleagues (Cheetham, Suter & Jäncke, 2014; Cheetham et al., 2015) were unable to find an uncanny valley effect: In their studies familiarity increased with human likeness with no decrease for ambiguously humanlike entities.

Ho and MacDorman (2010) found the eeriness of the uncanny valley to be conceptually and empirically distinct from the primary dimension of person perception, typically denoted by *warmth* or *likability* (Cuddy, Fiske, & Glick, 2007). People dislike villainous characters, regardless of whether they are human actors, computer models, or hand-drawn cartoons, and regardless of whether they appear eerie or unfamiliar or neither. Likability could also be confounded with human likeness, the independent variable in Mori's graph. Ho and MacDorman (2010) found human traits are typically more likable than nonhuman traits; indeed, the high correlation between human likeness and likability ($r = .73$, $p < .001$) raises concerns for studies that only use likability as the dependent variable in the uncanny valley graph (e.g., Ferrey et al., 2015; Yamada et al., 2013). Thus, although Mori (1970/2012) and his translators use the layman's term *affinity* to describe the dependent variable, this paper will sidestep the resulting complications by treating affinity as a technical term defined by Moore's (2012) equation (Equation 1). We evaluate predicted affinity primarily in terms of perceived eeriness because eeriness is essential to the uncanny valley experience (Mangan, 2015).

Evaluating Moore's model quantitatively by comparing predicted affinity and perceived eeriness (reverse scaled) can provide supporting or contrary evidence to refine perceptual mismatch theories. To this end, we conducted a pilot study with a subset of data from prior work (MacDorman & Chattopadhyay, 2016). However, Moore's model failed to predict perceived eeriness in part because it did not include as a factor the entity's level of anthropomorphism. Only for human and other vertebrate faces had reducing realism consistency increased the uncanny valley effect; for plants and artifacts no such effect was found (MacDorman & Chattopadhyay, 2016). These results align with the finding that the face perception network is highly sensitive to realism in human faces but not to realism in other kinds of objects (James et al., 2015). What Moore's equation does not explain is why an anthropomorphic appearance is necessary for perceptual tension to cause the uncanny valley effect (Figure 1).

Anthropomorphism, perceptual narrowing, and threat avoidance

We propose that, within Moore's Bayesian framework, the effect of anthropomorphic appearance on eeriness could partly be explained by perceptual narrowing. Although human infants start out discriminating human faces and those of other primates equally well, as they mature they become better at discriminating human faces and worse at discriminating those of other primates (Lewkowicz & Ghazanfar, 2006; Pascalis, de Haan, & Nelson, 2002; Scott, Pascalis, & Nelson, 2007). Perceptual narrowing also occurs in perceiving faces of other races and faces of nonprimate animals (Kelly et al., 2007; Simpson, Varga, Frick, & Fragaszy, 2010). Elsewhere we have proposed various mechanisms that are compatible with perceptual narrowing to explain how perceiving small deviations from human appearance could produce large prediction errors in brain regions honed for recognizing human faces, hands, and bodies, thus engendering the cold, eerie feelings associated with the uncanny valley (MacDorman & Chattopadhyay, 2016; MacDorman et al., 2009a; MacDorman & Ishiguro, 2006; also see Saygin, Chaminade, Ishiguro, Driver, & Frith, 2012).

Thus, Moore's equation needs to be revised because perceptual narrowing can cause a stimulus to appear unfamiliar even though its appearance is physically similar to familiar anthropomorphic stimuli. To overcome this issue, in our analysis we replace probability of occurrence in Moore's equation (an objective measure) with perceived familiarity (a subjective measure). We argue that small deviations from norms look more unfamiliar in anthropomorphic entities than deviations of the same magnitude in nonanthropomorphic entities. Thus, people may be inclined to rate stimuli that are relatively similar to human as unfamiliar because of perceptual narrowing. That effect needs to be captured by the dependent variable *affinity*.

A second issue with Moore's equation is that affinity increases with probability of occurrence, which implies we feel the greatest affinity for what is most familiar. However, the familiar is not necessarily

pleasant but could be boring, while the extremely novel could be terrifying. The relation between novelty and pleasure has typically been graphed as an inverted *U*. Pleasure rises as novelty increases to peak at a moderate degree of novelty and then rapidly falls; thus, high levels of novelty elicit unpleasant sensations and aversive responses (Berlyne, 1970, 1971; Lang, Bradley, Sparks, & Lee, 2007; Zuckerman, 1976, 2014). A prerequisite for perceiving novelty and, in turn, for responding aversively to excessive or unanticipated novelty is the perceptual narrowing that comes with the formation of new categories (Hebb, 1946; exposure learning, Gottlieb, 2002; van Kampen, 2015; for a review, see Bronson, 1968a). For example, human infants develop a fear of strangers only after they can distinguish a stranger from their mother (Bronson, 1968b).

In this paper, we extend realism inconsistency theory. After introducing our general method (see the General method section), we quantify Moore's Bayesian model of the uncanny valley, replacing probability of occurrence with perceived familiarity. We evaluate the model by comparing its affinity predictions with participants' reverse-scaled eeriness ratings of the stimulus (see the Quantitative evaluation of the Bayesian model section). The revised model's predictions closely followed participants' responses for humans (anthropomorphism: high). Based on this result, we set out to test whether, owing to perceptual narrowing, realism inconsistency inhibits the perceived familiarity of anthropomorphic entities (H1), and in turn causes the uncanny valley effect (H2; see the Unfamiliar anthropomorphic entities elicit eerie feelings section). The results support these hypotheses. Finally, we define *realism inconsistency* objectively at multiple scales and orientations in a two-dimensional (2-D) Morlet wavelet domain inspired by the primary visual cortex and use this concept to elucidate the role of familiarity in the perception of objects, animals, and humans by constructing corresponding structural equation models (see the Why unfamiliar anthropomorphic entities elicit eerie feelings section).

General method

Design

This study adopted a within-group design with three rounds of experiments. In each round unique participants observed task stimuli derived from four to five real entities and their three-dimensional (3-D) computer models. The experiment consisted of a two-alternative forced choice categorization task and ratings of the task stimuli.

Participants

Participants were recruited by email from a randomized exhaustive list of undergraduates attending a Midwestern public university system. Participants received no compensation. The study was approved by the Indiana University Office of Research Administration (Study No. 1210009909).

A total of 365 participants were recruited. In the first round, Zlatko, Ingrid, dog, and parrot were presented to 118 participants ($Mdn_{age} = 21$, $IQR_{age} = 1$; 67% female); in the second round, Clint, Emelie, Juliana, Simona, and Ferrari were presented to 104 participants ($Mdn_{age} = 21$, $IQR_{age} = 4.5$; 58% female); and in the third round, camera, washer, and water lily were presented to 143 participants ($Mdn_{age} = 22$, $IQR_{age} = 4$; 56% female).

Participants reflected the demographics of the university system's undergraduate population: 74% White, 7% African American, 6% Hispanic, 3% Asian, 3% multiracial, 6% foreign, and 1% unknown. Participants had no or mild visual impairment with corrective lenses (100%, $Mdn =$ no impairment), were mostly right-handed (82%), native English speakers (91%), raised (95%) and residing (98%) in the United States. The three rounds of experiments were conducted sequentially between October 2013 and March 2014.

Stimuli

The task stimuli from MacDorman and Chattopadhyay (2016) were used. The dataset contained images (600×600 pixels) of 12 entities representing three levels of anthropomorphism: low (four nonanimal objects: car, camera, washing machine, and water lily), intermediate (two nonhuman animals: dog and parrot), and high (six humans: Clint, Zlatko, Emelie, Ingrid, Simona, and Juliana).

For each of the 12 entities, a 3-D replica was modeled using computer software (Figure 2), and 17 representations were derived using the original photograph and the computer model as follows. First, two observational dimensions were defined for each of the 12 entities: Feature Set 1 and 2 (Table 1). Then, for both feature sets, the fraction of real was systematically varied by sixths along three transitions from 0% real (100% computer modeled) to 100% real (0% computer modeled; Figure 3). For instance, the fraction of real of Feature Set 1 could be 0, 1/6, 1/3, 1/2, 2/3, 5/6, or 1. The three transitions entail only 17 of 49 possible combinations to reduce participant fatigue and attrition.

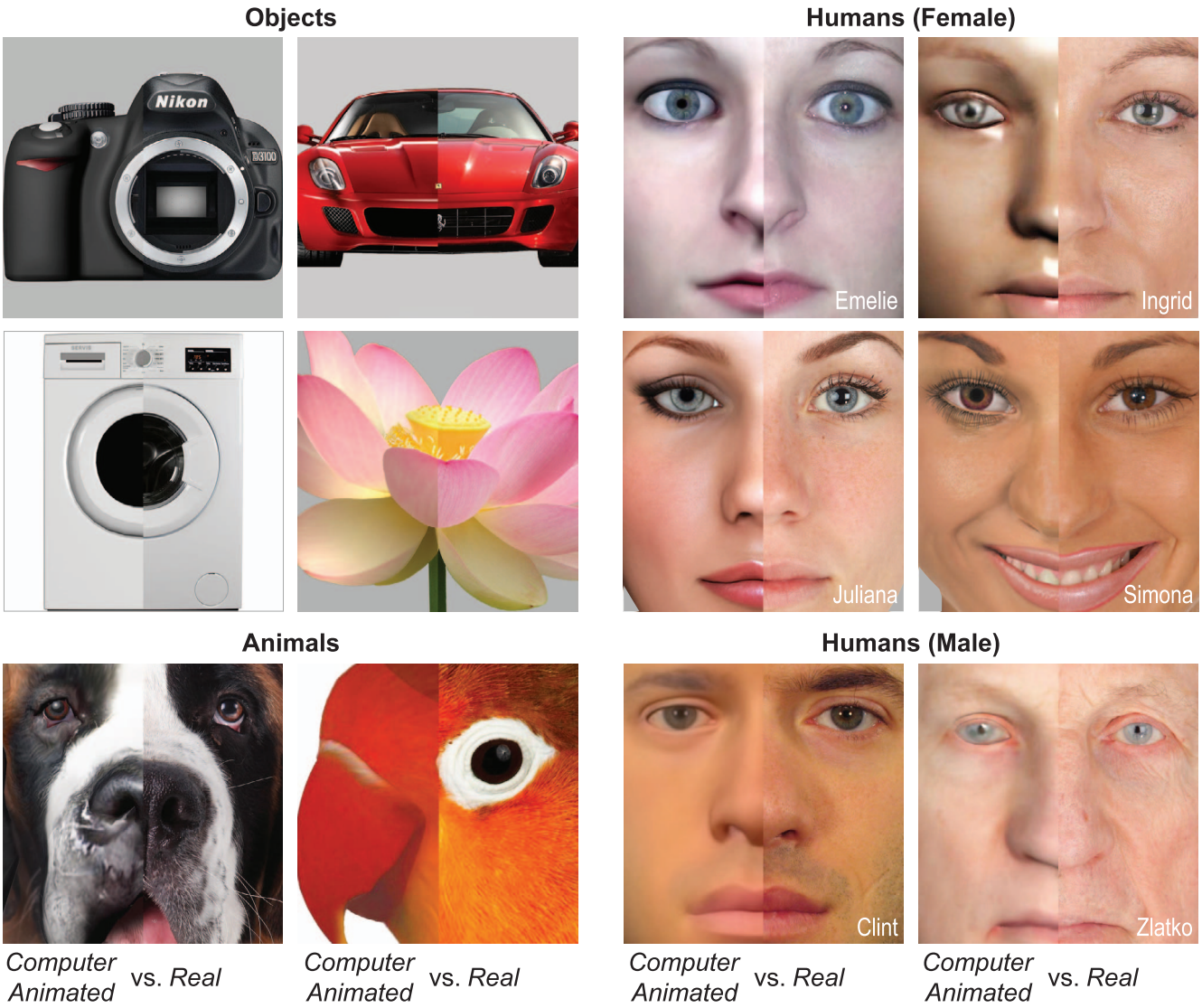

Figure 2. Objects, animals, and humans constitute the low, intermediate, and high anthropomorphism groups, respectively. The right half of a photograph of each entity is shown beside the left half of its 3-D computer model.

Procedure

Using an access-controlled website, participants completed the following sequence of activities: informed consent, the dichotomous thinking inventory (Oshio, 2009), categorization tasks, task stimuli ratings (see Dependent variables section), and a demographics survey. A pilot study found each round requires approximately 15–30 min to complete.

Entity	Feature Set 1 (FS1)	Feature Set 2 (FS2)
Camera	lens mount, reflex mirror, parts visible behind the lens	front, lens release, mode dial, shutter button
Car	headlights, front grilles	bumper, hood, outside mirrors, windshield
Flower	stamen, stigma, style	petals, stem
Washing machine	door, door handle, window	front panel, controls, filter cover, dispenser
Dog	eyes, flew, tongue	cheek, foreface, nose, stop
Parrot	eyes, eyelids, beak, cere, left nostril	cheek, forehead, lore
Humans	eyes, eyelashes, mouth	skin, nose, eyebrows

Table 1. Feature sets for objects, animals, and humans.

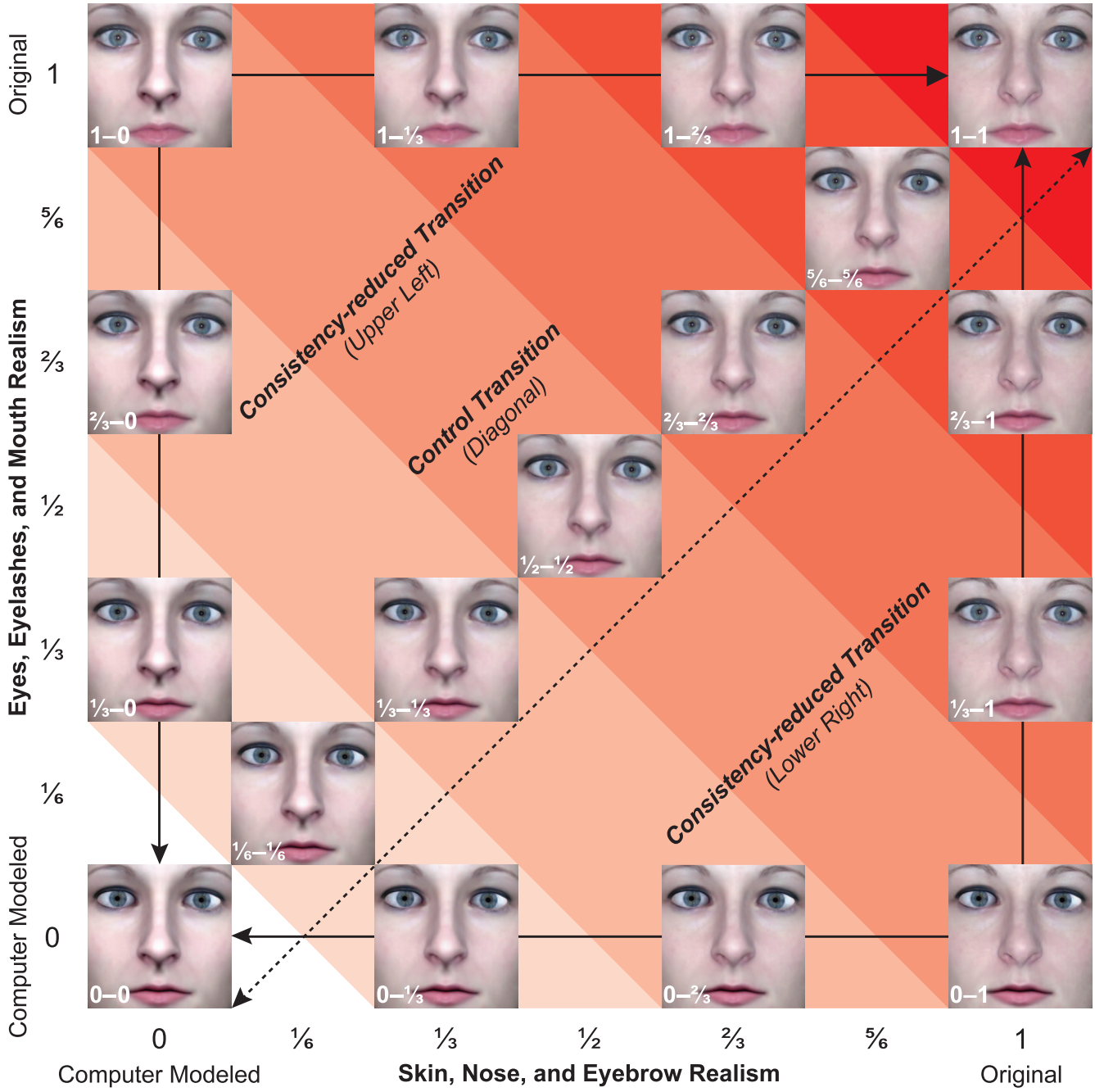

Figure 3. The diagonal depicts a consistent change in the objective realism (fraction of real) of all features of an entity, from the 3-D model to the original. The lower-right path depicts an inconsistent change in which Feature Set 2 (e.g., skin, nose, and eyebrows) changes first and then Feature Set 1 (e.g., eyes, eyelashes, and mouth). The upper-left depicts an inconsistent change in which Feature Set 1 changes first and then Feature Set 2. The colored bands indicate the consistency-reduced representations and control being compared.

In the categorization task, participants categorized the task stimuli as either computer animated or real. (The term *computer animated* was used instead of the more accurate descriptor *computer modeled* because participants were familiar with it.) The order of stimulus presentation was randomized for each participant, and the first two trials of each block were practice trials. Each trial began with the presentation of a task stimulus. On the left of the stimulus was an anchor (e.g., *real*) with the instruction "Press *e*," and on the right of the stimulus was another anchor (e.g., *computer animated*) with the instruction "Press *i*." Participants were instructed to categorize the stimuli as quickly and as accurately as possible. They could pause and restart the task by pressing the spacebar. If no response was provided within 3000 ms, the task paused automatically. The task

restarted with the next stimulus, and the current stimulus was randomly inserted into the remaining trials. Each stimulus was presented twice to swap the left–right order of the anchors. The task continued until responses to all stimuli were provided. Furthermore, after each trial, three masks were presented in sequence for 100 ms each to suppress the afterimage of the stimulus (a 19 × 19 checkerboard, a 38 × 38 checkerboard, and a 50% gray panel). Thus, the minimum interval between trials was 300 ms. For each trial, the categorization response was recorded as was the response time (elapsed time in milliseconds between the presentation of the image and the registration of an *e* or *i* keypress).

For task stimuli ratings, participants rated realism, eeriness, warmth, and familiarity for 68 stimuli (4 entities × 17 representations) in the first round, 85 stimuli (5 entities × 17 representations) in the second, and 51 (3 entities × 17 representations) in the third (see the Dependent variables section).

Independent variables

The independent variables for this study were anthropomorphism (low, intermediate, and high) and fraction of real. Fraction of real was varied along one control transition (diagonal) and two consistency-reduced transitions (lower right and upper left). For the diagonal transition, the stimuli had consistent fractions of real for Feature Set 1 and 2 (FS1–FS2): 0–0, 1/6–1/6, 1/3–1/3, 1/2–1/2, 2/3–2/3, 5/6–5/6, and 1—1. For the lower-right consistency-reduced transition, FS1–FS2 were 0–0, 0–1/3, 0–2/3, 0–1, 1/3–1, 2/3–1, and 1–1, and for the upper-right consistency-reduced transition, 0–0, 1/3–0, 2/3–0, 1–0, 1–1/3, 1–2/3, and 1–1.

Dependent variables

The dependent variables were the percentage of times each task stimulus was categorized as *real,* its mean response times, and its rating on indices of realism, eeriness, warmth, and familiarity. The indices were used experimentally in hypothesis testing. Each index used three 7-point semantic differential items, ranging from −3 to +3 (i.e., from left to right: very, moderately, slightly, neutral, slightly, moderately, very). The anchors were, for realism: computer animated–real, replica–original, and digitally copied–authentic; for eeriness: ordinary–creepy, plain–weird, and predictable–eerie; for warmth: cold-hearted–warm-hearted, hostile–friendly, and grumpy–cheerful; and for familiarity: rarely seen–common, unfamiliar–recognizable, and unique–familiar. Task stimuli ratings followed the categorization tasks to reduce bias from exposure to the realism, eeriness, warmth, and familiarity indices.

Quantitative evaluation of the Bayesian model

Method

Moore (2012) based his Bayesian model of the uncanny valley on Feldman and colleagues' (2009) model of categorical perception. In his model, Moore defines *perceptual tension* as the perception of stimuli whose features differ in the certainty of their category membership across observational dimensions. Moore's model defined *affinity* (the dependent variable in Mori's graph) as the probability of occurrence of a stimulus minus the perceptual tension arising from its conflicting features:

$$F[S] = p(S) - kV[S] \quad (1)$$

where $F[S]$ represents affinity for stimulus S, $p(S)$ the probability of its occurrence, k the viewer's sensitivity to perceptual tension, and $V[S]$ the perceptual tension. Perceptual tension is measured as the variance in perceptual distortions for each dimension on which a stimulus can be perceived:

$$V[S_i] = E\left[(D[S_i])^2\right] - (E[D[S_i]])^2 \quad (2)$$

where $D[S]$ is along dimension i (Feldman et al., 2009; Moore, 2012). $D[S]$ is calculated using the variance of a feature's category and the probability of perceiving the feature as belonging to that category (see Appendix A).

This model's predictions were compared with the eeriness and warmth ratings from experiments in MacDorman and Chattopadhyay (2016). These data were selected because objects, animals, and humans (both male and female) were represented by at least two individuals of each group and because the two observation dimensions, Feature Set 1 and 2, were controlled with seven equidistant levels of objective realism. Features were designed to be perceived as either real or computer modeled. Stimuli had either consistency-reduced or control levels of realism across the two feature sets. This enabled the model's predictions to be tested with both matched cues (less perceptual tension) and mismatched cues (more perceptual tension). We next explain how model parameters were computed to predict the empirical data. Model parameters were calculated separately for each of the three transitions (two consistency-reduced and one control transition, see the Stimuli section).

To parametrize the Bayesian model, we adapted the calculations of Feldman and colleagues (2009, pp. 13–15). In our experimental setup, realism was the dimension along which features varied and thus were perceived as either real (category 1) or computer animated (category 2). The two observation dimensions were Feature Set 1 and 2 (Table 1).

The parameters to be set were

μ_{real}: mean of the category *real*
$\mu_{computer\_modeled}$: mean of the category *computer animated*
σ_c^2: category variance
σ_s^2: uncertainty in the observed signal
k: viewer's sensitivity to perceptual tension, and
$p(S)$: probability of occurrence of the stimulus.

Similar to the computation of goodness ratings of categories in Iverson and Kuhl (1995), we computed the mean of the category *real* using the categorization responses and realism ratings. For objects, animals, and humans, μ_{real} was the fraction of real of the stimulus's Feature Set 1 and 2 that received the highest realism ratings, when we rank-ordered realism ratings for all stimuli categorized as *real*.

The means of the category *computer animated* for objects, animals, and humans were calculated based on μ_{real} and the category identification curves (probability of belonging to the category *computer animated* or *real,* see figure 5 in Feldman et al., 2009). Curves were plotted using the percentage of times stimuli were categorized as real (Appendix C, Figure C1). To identify the category boundary, we fitted a logistic function and calculated the gain g and the bias b of the logistic function. Equations for these calculations are described in Appendix B. Feldman and colleagues' computations used the empirical data of Iverson and Kuhl (1995), whose stimuli, speech sounds, were expressed orthogonally—in terms of their first two formants (F_1–F_2 space, see figures 1 and 2 in Iverson & Kuhl, 1995). Similarly, in calculating our logistic parameters, we used eigenvectors to represent orthogonally the two observational dimensions of a stimulus. For each entity, first an eigenface was generated using all 17 variants (Turk & Pentland, 1991). Then each image was represented by their reconstruction errors from projection onto the first and second eigenvectors, sorted by decreasing eigenvalues. Finally, assuming high category variance, we set the free parameter, σ_S^2/σ_c^2 to 0.1. (Predictions, however, did not change significantly for $0.1 \leq \sigma_S^2/\sigma_c^2 \leq 1$.) To measure individual differences in the viewer's sensitivity to perceptual tension k, we used the dichotomous thinking inventory (Oshio, 2009) as a rough analogue, averaging all 15 items for each participant. We measured $p(S)$ as the perceived familiarity ratings of the stimulus.

Predictions of the Bayesian model

Predicted affinity and perceived eeriness (reverse scaled) were compared for matched stimuli for low, intermediate, and high levels of anthropomorphism (i.e., objects, animals, and humans) and for one control (diagonal) and two consistency-reduced (lower right and upper left) transitions. The Bayesian model closely fit the data (Figures 4 through 6). Table 2 lists the parameters used in the model (derived from the equations in Appendix A and B). Predicted perceptual tension and sensitivity to perceptual conflict were rescaled from 0 to 1. Rescaled sensitivity was approximated using the dichotomous thinking inventory (round 1: $M = 0.48$, $SD = 0.03$, round 2: $M = 0.57$, $SD = 0.04$, and round 3: $M = 0.54$, $SD = 0.03$).

		Maximum real		Maximum computer modeled	
Entity	Transition	Feature Set 1	Feature Set 2	Feature Set 1	Feature Set 2
Humans	Diagonal	100	100	96.41	93.60
	Lower right	100	100	94.62	90.82
	Upper left	100	100	96.85	94.39
Animals	Diagonal	83	83	95.87	98.17
	Lower right	100	100	74.33	76.90
	Upper left	100	100	75.81	79.07
Objects	Diagonal	100	100	101.62	99.08
	Lower right	100	100	99.54	99.47
	Upper left	100	100	100.81	98.85

Table 2. Parameters used to evaluate the Bayesian model of the uncanny valley.

We computed the mean distance between predicted affinity and perceived eeriness (reverse scaled) for all three transitions in perceiving objects: diagonal, $L^1 = -0.04$, lower right, $L^1 = -0.16$, and upper left, $L^1 = -0.03$; animals: diagonal, $L^1 = -0.41$, lower right, $L^1 = -0.63$ and upper left, $L^1 = -0.47$; and humans: diagonal, $L^1 = -0.27$, lower right, $L^1 = -0.28$, and upper left, $L^1 = -0.34$. To assess curve similarity, Fréchet distance (Alt & Godau, 1995) between predicted affinity and perceived eeriness (reverse scaled) was computed after subtracting their mean L^1 distance. The curves were similar for all three transitions in perceiving objects: diagonal, $d = 0.31$, lower right, $d = 0.43$, and upper left, $d = 0.29$; animals: diagonal, $d = 0.19$, lower right, $d = 0.24$, and upper left, $d = 0.30$; and humans: diagonal, $d = 0.09$, lower right, $d = 0.07$, and upper left, $d = 0.16$. As these results indicate, the Bayesian model of the uncanny valley predicted perceived eeriness from the fraction of real of the stimulus for Feature Set 1 and 2, its perceived realism and familiarity, its percentage categorized as real, and the viewer's sensitivity to perceptual tension, measured by the dichotomous thinking inventory. (See Appendix D for a comparison of predicted affinity and perceived warmth.)

Excluding the 100% real representation, realism across the 16 remaining representations was rated higher for object ($M = -0.26$, $SD = 0.89$) than animal ($M = -0.70$, $SD = 1.75$) or human models ($M = -0.87$, $SD = 1.56$). However, 100% real human (2.04) and animal faces (2.14) were perceived as more real than 100% real objects (1.18). The familiarity across 100% or partly

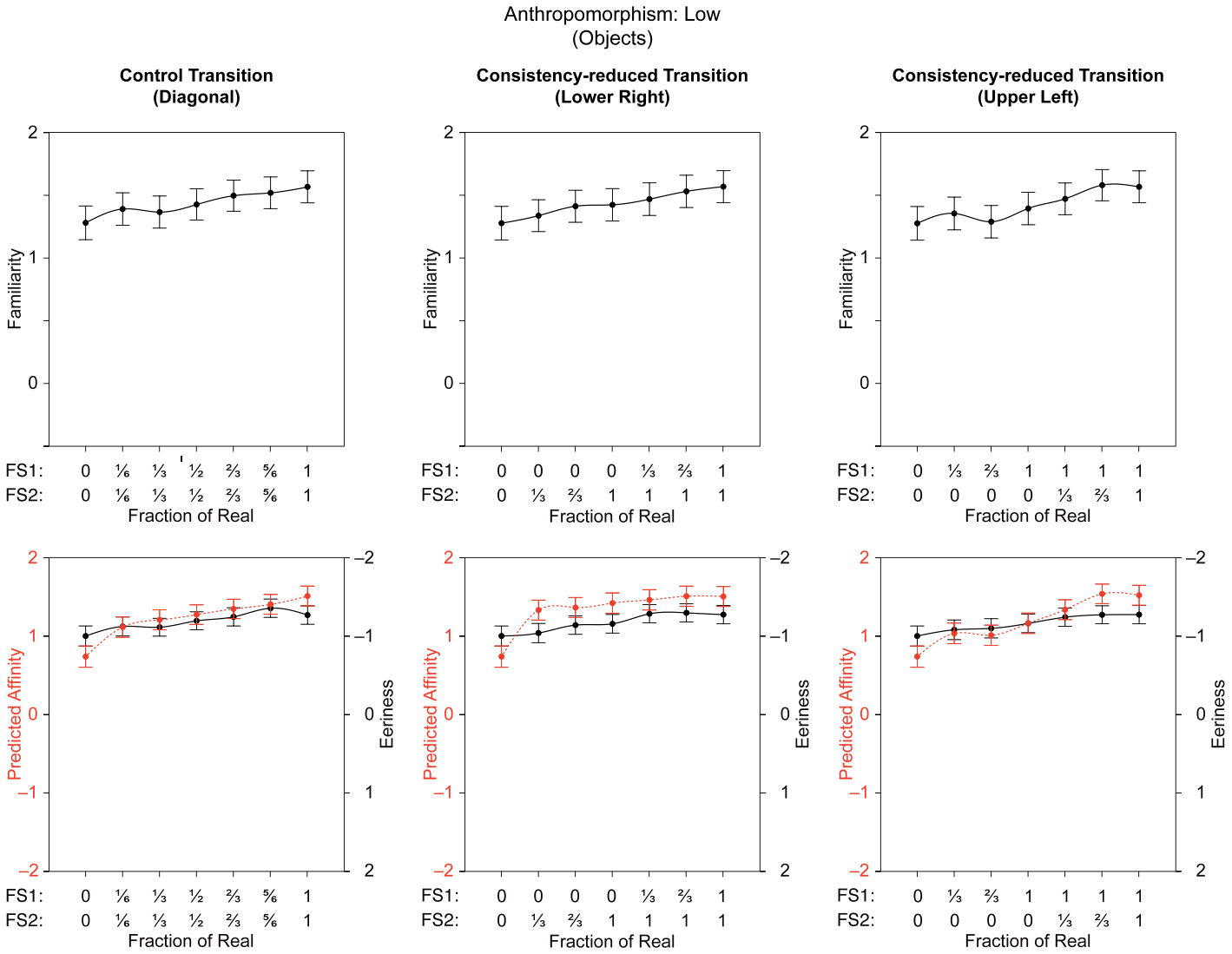

Figure 4. For objects, perceived familiarity and eeriness ratings of the stimulus are plotted against its fraction of real for the control (diagonal) and the consistency-reduced transitions (lower right, upper left). The dashed line represents affinity as predicted by the revised Bayesian model. (The lines in Figures 4 through 6 depict cubic spline interpolation.)

computer-modeled representations was rated higher in object ($M = 1.42$, $SD = 0.09$) than animal ($M = 0.77$, $SD = 0.60$) and human models ($M = 0.53$, $SD = 0.41$). The difference in familiarity between fully real objects (1.56) and fully real humans (1.19) and animals (1.75) was much less.

Unfamiliar anthropomorphic entities elicit eerie feelings

Theory and hypotheses

Realism inconsistency theory predicts that features at inconsistent levels of realism in an anthropomorphic entity cause perceptual processes in viewers to make conflicting inferences regarding whether the entity is real, thus resulting in the uncanny valley effect (MacDorman & Chattopadhyay, 2016). Prior research has found realism inconsistency increases eeriness (MacDorman et al., 2009a; Mitchell et al., 2011) or unpleasantness (Seyama & Nagayama, 2007). This study extends realism inconsistency theory. We hypothesize that, owing to perceptual narrowing, higher levels of anthropomorphism amplify the inhibitory effect of realism inconsistency on perceived familiarity. Put briefly, realism inconsistency makes anthropomorphic entities look unfamiliar.

However, this does not explain why anthropomorphic entities that appear unfamiliar should elicit cold, eerie feelings and avoidance behavior. Eeriness and creepiness are associated with basic emotions, elicited

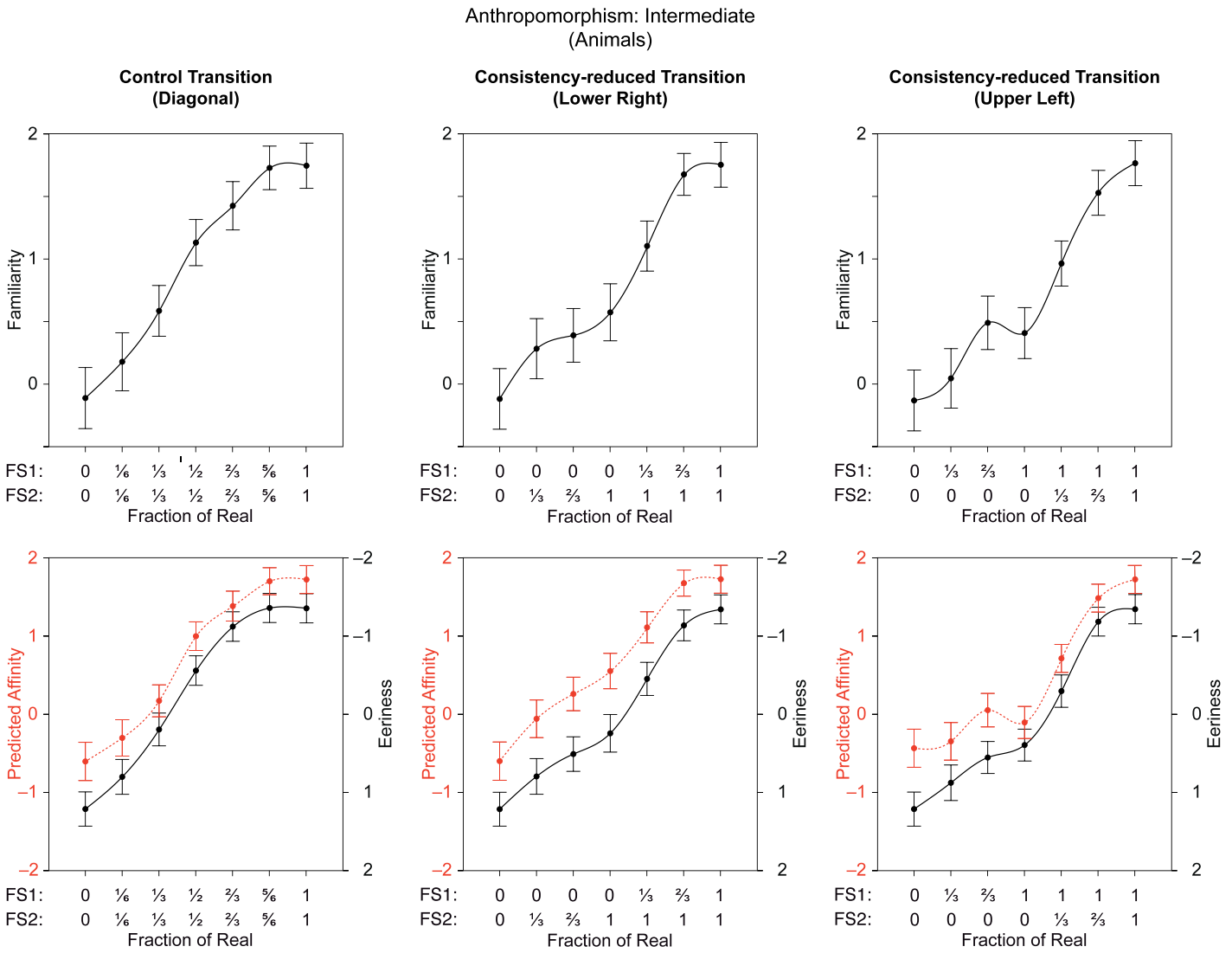

Figure 5. For animals, perceived familiarity and eeriness ratings of the stimulus are plotted against its fraction of real for the control (diagonal) and the consistency-reduced transitions (lower right, upper left). The dashed line represents affinity as predicted by the revised Bayesian model.

by threats, that motivate avoidance behavior—emotions like fear, anxiety, and disgust (Ho, MacDorman, & Pramono, 2008).

The uncanny valley effect could be a symptom of an adaptive mechanism for threat avoidance

Proposed threat avoidance mechanisms vary widely in their origin and function, from preserving the biological self and its genes to preserving the socially constructed self and its cultural worldview (MacDorman & Ishiguro, 2006). For example, we could be more disturbed by abnormalities in an anthropomorphic entity because they activate a mechanism for pathogen avoidance; the entity's physical and behavioral similarity indicate genetic relatedness and, thus, a higher risk of infection from exposure (Curtis et al., 2004; MacDorman et al., 2009a; Moosa & Ud-Dean, 2010; Park, Faulkner, & Schaller, 2003). Alternatively, abnormalities could indicate low fitness in a potential mate, thus activating a mechanism for unfit mate avoidance (MacDorman & Ishiguro, 2006; MacDorman et al., 2009a). A propensity of threat avoidance mechanisms to register false positives for human replicas could be explained by the high fitness cost of *not* registering existential and reproductive threats (Burleigh et al., 2013; Nesse, 2005) and the absence of human replicas in our evolutionary environment. From a sociological standpoint, a human replica could threaten an individual's religious and cultural worldview and sense of personal and human identity, thus eliciting psychological defense mechanisms (MacDorman & Entezari, 2015; MacDorman et al., 2009b; Ramey, 2005).

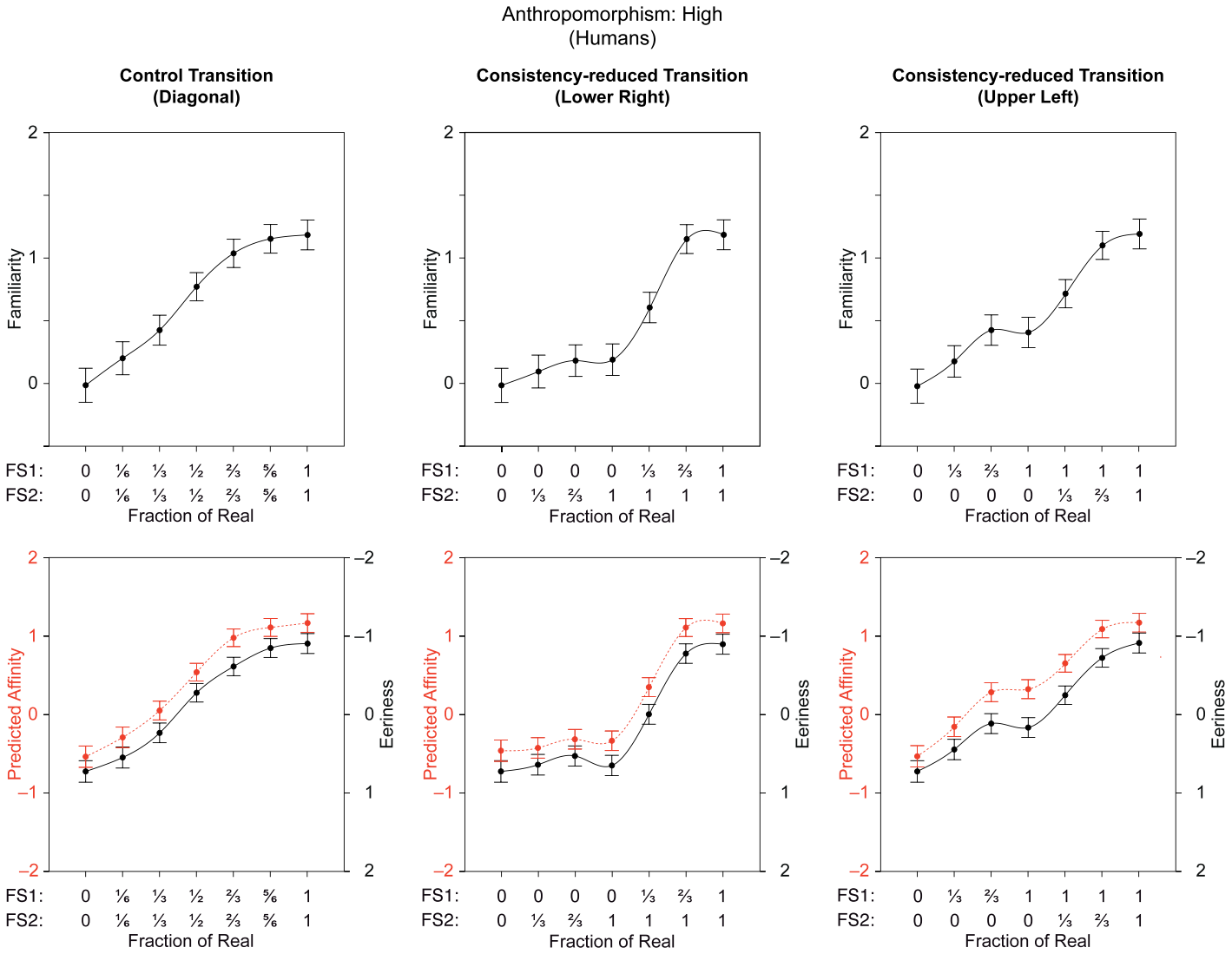

Figure 6. For humans, perceived familiarity and eeriness ratings of the stimulus are plotted against its fraction of real for the control (diagonal) and the consistency-reduced transitions (lower right, upper left). The dashed line represents affinity as predicted by the revised Bayesian model.

We tested our extension of realism inconsistency theory with two hypotheses:

H1. The more anthropomorphic the entity, the more reducing consistency in feature realism decreases perceived familiarity.

H2. Only when reducing realism consistency reduces perceived familiarity does it also increase the uncanny valley effect.

Method

The independent variables remained anthropomorphism and fraction of real, while the dependent variable familiarity was used from the task stimuli ratings to test H1, and eeriness and familiarity were used to test H2 (see the General method section). The uncanny valley effect is operationalized as an increase in eeriness because eeriness is the experience that a theory of the uncanny valley must explain (see the Operationalizing *Shinwakan* section). The index used to measure eeriness has both discriminant and content validity (Ho & MacDorman, 2010). (An analysis of perceived warmth is reported in Appendix E.)

Data analysis preliminaries

H1 was tested using a mixed-design analysis of variance (ANOVA). H2 was tested using paired-samples *t* tests and within-group repeated-measures ANOVAs. Because a preliminary analysis using a linear mixed-effects model found little effect of entity as a random factor, the hypotheses were tested using

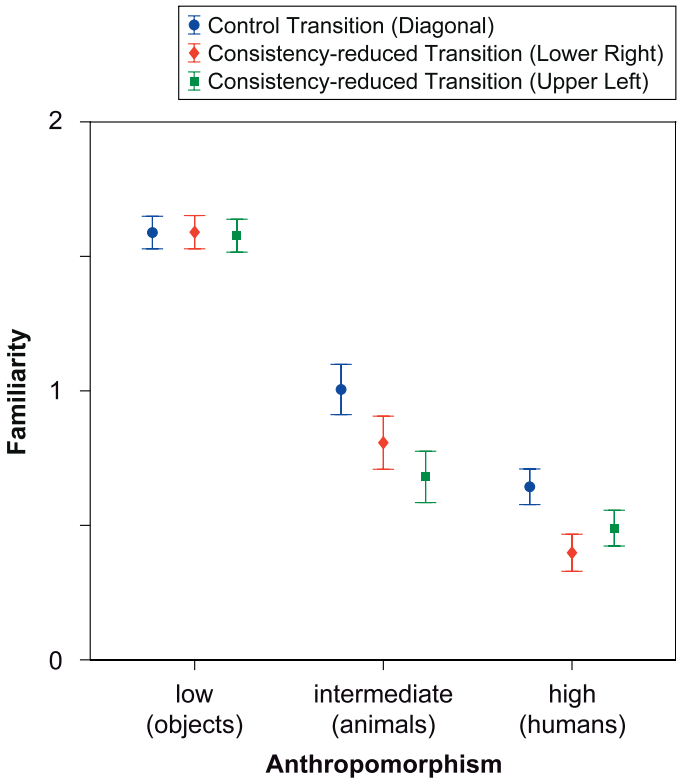

Figure 7. Familiarity ratings are plotted against the level of anthropomorphism (low: objects, intermediate: animals, and high: humans) for the control (diagonal) and the consistency-reduced transitions (lower right, upper left).

only the fixed factor: fraction of real (MacDorman & Chattopadhyay, 2016, supplementary material, section 1).

Response times were $\log_{10}$-transformed to remove positive skew. Q-Q plots were linear, confirming normality (Thode, 2002). Test statistics were interpreted with a significance threshold of $\alpha = .05$. Effect sizes for ANOVAs are reported as partial eta-squared (η_p^2) with the thresholds small = .01, medium = .06, and large = .14 and for t tests as Cohen's d with the thresholds small = .20, medium = .50, and large = .80 (Cohen, 1992). Pairwise comparisons are reported with Bonferroni–Holm correction, and violations of sphericity with Greenhouse–Geisser correction.

Testing H1

H1 states that the more anthropomorphic the entity, the more reducing consistency in feature realism decreases familiarity. To test H1, we conducted a mixed-design ANOVA with anthropomorphism as a between-groups factor and realism consistency as a within-group factor. We analyzed low, intermediate, and high levels of anthropomorphism, namely, objects (camera, washer, and water lily, $n = 2{,}145$, Round 3), animals (dog and parrot, $n = 1{,}180$, Round 2), and humans (Clint, Emelie, Juliana, and Simona, $n = 2{,}080$, Round 1; Figure 7), respectively, and control (diagonal) and consistency-reduced (lower right, upper left) transitions. The ANOVA confirmed the level of anthropomorphism significantly affected the difference in eeriness ratings between the control and consistency-reduced transitions. An interaction effect of anthropomorphism × realism consistency was found for both the lower-right, $F(2, 5{,}391) = 20.06$, $MSE = 0.87$, $p < .001$, $\eta_p^2 = .007$, and upper-left transition, $F(2, 5{,}388) = 21.16$, $MSE = 0.89$, $p < .001$, $\eta_p^2 = .008$. These results indicate that reduced realism consistency affected familiarity ratings of objects, animals, and humans differently.

For the lower-right transition, a post hoc Tukey's honest significant difference (HSD) test found that reduced realism consistency decreased familiarity significantly more for humans than for objects and significantly more for animals than for objects, $p < .001$. For the upper-left transition, a post hoc Tukey's HSD found that reduced realism consistency decreased familiarity significantly more for humans than for objects, $p < .001$, and significantly more for animals than for objects, $p < .001$, but significantly more for animals than for humans, $p = .001$.

H1 was supported for four of six comparisons (3 levels of anthropomorphism × 2 transitions). H1 was not supported in comparing humans and animals in the upper-left and lower-right transition.

Testing H2

A nested-factor ANOVA (familiarity ~ transition/fraction of real with five levels of fraction of real, except the two endpoints) found that transition significantly affected familiarity in perceiving humans, $F(2, 9{,}753) = 26.02$, $MSE = 2.44$, $p < .001$, $\eta_p^2 = .005$, and animals, $F(2, 3{,}517) = 12.43$, $MSE = 2.52$, $p < .001$, $\eta_p^2 = .007$, but not objects, $F(2, 7{,}975) = 0.12$, $p = .885$.

Object models

For the four object models ($n = 533$) in the low anthropomorphism group, fraction of real significantly affected familiarity ratings for all three transitions with a small effect size: diagonal, $F(5.53, 2{,}911) = 7.55$, $MSE = 0.75$, $p < .001$, $\eta_p^2 = .01$; lower right, $F(5.40, 2{,}866) = 8.49$, MSE = 0.73, $p < .001$, $\eta_p^2 = .02$; and upper left, $F(5.64, 2{,}995) = 12.03$, $MSE = 0.73$, $p < .001$, $\eta_p^2 = .02$ (Figure 4). Fraction of real also significantly affected eeriness ratings for each transition with a small effect size: diagonal, $F(5.18, 2{,}724) = 9.78$, $MSE = 0.86$, $p < .001$, $\eta_p^2 = .02$; lower right, $F(5.00, 2{,}654) = 9.84$, $MSE =$

Consistency-reduced vs. control transition	Familiarity				Eeriness			
	t	*df*	*p*	*d*	*t*	*df*	*p*	*d*
0–1/3 vs. 1/6–1/6	−1.19	234	.117		0.09	234	.465	
0–2/3 vs. 1/3–1/3	2.17	235	.016		−3.65	235	<.001	0.24
0–1 vs. 1/2–1/2	5.45	235	<.001	0.35	−6.91	235	<.001	0.45
1/3–1 vs. 2/3–2/3	3.22	235	.001	0.21	−5.83	235	<.001	0.38
2/3–1 vs. 5/6–5/6	0.70	233	.243		−2.36	233	.010	0.15
1/3–0 vs. 1/6–1/6	1.49	234	.069		−0.87	234	.192	
2/3–0 vs. 1/3–1/3	0.87	235	.192		−3.00	235	.002	0.20
1–0 vs. 1/2–1/2	5.95	233	<.001	0.39	−7.55	233	<.001	0.49
1–1/3 vs. 2/3–2/3	4.29	233	<.001	0.28	−6.69	233	<.001	0.44
1–2/3 vs. 5/6–5/6	2.06	232	.021		−1.60	232	.056	

Table 3. Familiarity and eeriness for all 10 comparisons in perceiving animals. *Notes*: Boldface indicates reduced realism consistency significantly increasing eeriness. Fraction pairs indicate the fraction of real for Feature Set 1 and 2. Effect sizes are provided for significant outcomes in Tables 3 and 4.

0.94, $p < .001$, $\eta_p^2 = .02$; and upper left, $F(5.11, 2{,}711) = 7.69$, $MSE = 0.92$, $p < .001$, $\eta_p^2 = .01$ (Figure 4).

Pairwise comparisons showed that for objects, reduced realism consistency did not significantly affect eeriness or familiarity for any pairs, either between the lower-right or upper-left transition and the control.

Animal models

For the two animal models ($n = 236$) in the intermediate anthropomorphism group, fraction of real significantly affected familiarity ratings for all three transitions with a large effect size: diagonal, $F(3.85, 886) = 80.79$, $MSE = 2.49$, $p < .001$, $\eta_p^2 = .26$; lower right, $F(3.92, 912) = 82.12$, $MSE = 2.26$, $p < .001$, $\eta_p^2 = .26$; and upper left, $F(3.83, 885) = 73.98$, $MSE = 2.57$, $p < .001$, $\eta_p^2 = .24$ (Figure 5). Fraction of real also significantly affected eeriness ratings for each transition with a large effect size: diagonal, $F(4.12, 947) = 149$, $MSE = 2.53$, $p < .001$, $\eta_p^2 = .39$; lower right, $F(4.32, 1{,}006) = 129$, $MSE = 2.40$, $p < .001$, $\eta_p^2 = .36$; and upper left, $F(4.08, 944) = 131$, $MSE = 2.6$, $p < .001$, $\eta_p^2 = .36$ (Figure 5).

Pairwise comparisons showed that for animals reduced realism consistency significantly increased eeriness for four of five pairs in the lower-right transition: 1/3–1/3 and 0–2/3, 1/2–1/2 and 0–1, 2/3–2/3 and 1/3–1, and 5/6–5/6 and 2/3–1; and three of five pairs in the upper-left transition: 1/3–1/3 and 2/3–0, 1/2–1/2 and 1–0, and 2/3–2/3 and 1–1/3 (Table 3). Pairwise comparisons further showed that for animals reduced realism consistency significantly decreased familiarity for two pairs in the lower-right transition: 1/2–1/2 and 0–1, and 2/3–2/3 and 1/3–1; and two pairs in the upper-left transition: 1/2–1/2 and 1–0, and 2/3–2/3 and 1–1/3. Thus, H2 was supported for all pairs except one between the lower-right transition (2/3–0 fraction of real) and the control and two between the upper-left transition (0–2/3 and 2/3–1) and the control.

Human models

For the six human models in the high anthropomorphism group ($n = 652$), three within-group repeated-measures ANOVAs confirmed that fraction of real significantly affected familiarity ratings for all three transitions with a large effect size: diagonal, $F(3.91, 2{,}529) = 106$, $MSE = 2.17$, $p < .001$, $\eta_p^2 = .14$; lower right, $F(4.14, 2{,}664) = 114$, $MSE = 2.09$, $p < .001$, $\eta_p^2 = .15$; and upper left, $F(4.11, 2{,}634) = 102$, $MSE = 1.93$, $p < .001$, $\eta_p^2 = .14$ (Figure 6). Fraction of real also significantly affected eeriness ratings for each transition with a large effect size: diagonal, $F(3.71, 2{,}396) = 186$, $MSE = 2.53$, $p < .001$, $\eta_p^2 = .22$; lower right, $F(3.96, 2{,}553) = 190$, $MSE = 2.52$, $p < .001$, $\eta_p^2 = .23$; and upper left, $F(4.32, 2{,}792) = 150$, $MSE = 2.16$, $p < .001$, $\eta_p^2 = .19$ (Figure 6).

Pairwise comparisons showed that for humans reduced realism consistency significantly increased eeriness for three pairs in the lower-right transition, 1/3–1/3 and 0–2/3, 1/2–1/2 and 0–1, and 2/3–2/3 and 1/3–1, and three pairs in the upper-left transition, 1/2–1/2 and 1–0, 2/3–2/3 and 1–1/3, and 5/6–5/6 and 1–2/3 (Table 4). Pairwise comparisons further showed that for humans reduced realism consistency significantly decreased familiarity for three pairs in the lower-right transition, 1/3–1/3 and 0–2/3, 1/2–1/2 and 0–1, and 2/3–2/3 and 1/3–1, and two pairs in the upper-left transition, 1/2–1/2 and 1–0, and 2/3–2/3 and 1–1/3. Thus, H2 was supported in all except one in the upper-left transition (5/6–5/6).

These results support H2. Reducing consistency in feature realism decreased familiarity significantly in perceiving humans and animals with a small effect size but not objects. Reducing consistency in realism also

Consistency-reduced vs. control transition	Familiarity				Eeriness			
	t	*df*	*p*	*d*	*t*	*df*	*p*	*d*
0−1/3 vs. 1/6−1/6	2.09	650	.019		−1.48	650	.069	
0−2/3 vs. 1/3−1/3	3.66	649	<.001	0.14	−4.46	649	<.001	0.17
0−1 vs. 1/2−1/2	8.92	650	<.001	0.35	−13.22	650	<.001	0.52
1/3−1 vs. 2/3−2/3	7.36	649	<.001	0.29	−9.60	649	<.001	0.38
2/3−1 vs. 5/6–5/6	0	648	.500		−1.37	648	.086	
1/3−0 vs. 1/6–1/6	0.38	650	.353		1.79	650	.037	
2/3−0 vs. 1/3−1/3	−0.08	650	.470		1.84	650	.033	
1−0 vs. 1/2−1/2	6.01	650	<.001	0.24	−6.98	650	<.001	0.27
1−1/3 vs. 2/3−2/3	6.40	649	<.001	0.25	−6.59	649	<.001	0.26
1−2/3 vs. 5/6−5/6	1.40	649	.081		−2.88	649	.002	0.11

Table 4. Familiarity and eeriness for all 10 comparisons in perceiving humans. *Note*: Boldface indicates reduced realism consistency significantly increasing eeriness.

caused a small uncanny valley effect, shown as increased eeriness, in perceiving humans and animals but not objects. In humans and animals, the results show that for nine out of 13 pairwise comparisons, only when reduced realism consistency decreased perceived familiarity did it also increase the uncanny valley effect. (See Appendix E for a comparison of perceived familiarity and warmth.)

Why unfamiliar anthropomorphic entities elicit eerie feelings

To confirm the relations among realism inconsistency, realism, familiarity, eeriness, and warmth, a structural equation model was calculated for each level of anthropomorphism: low (objects), intermediate (animals), and high (humans). A first step in this process was to define realism inconsistency objectively. We define it in a 2-D Morlet wavelet domain functionally analogous to how the primary visual cortex represents physical properties of the stimulus.

A multiscale and orientation measure for realism inconsistency

Daugman (1980) proposed that a parametrized family of 2-D Morlet basis functions (Gabor kernels) could model the anisotropic receptive field profiles of single neurons in several areas of the primary visual cortex, including their selective tuning for characteristic scale, localization, orientation, and quadrature phase relations. Daugman (1985) found using chi-squared tests that this wavelet family fit the profiles of 97% of single neurons in the cat visual cortex.

To compare a target representation to its 100% real representation, both representations were decomposed in a family of 2-D Morlet basis functions, analogous to neurons, resulting in wavelet coefficients, analogous to their firing rates. By definition, the realism inconsistency of the 100% real representation is zero. For the 16 remaining representations, we define *realism inconsistency* as the standard deviation of the differences between corresponding wavelet coefficients of the target representation and the 100% real representation.

Prior to computing its distance from the 100% real representation, each image was converted to gray-scale and decomposed in 2-D Morlet basis functions at four scales (1, 2, 3, and 4) and six orientations (0° to 150° at 30° intervals) using *cwtft2* in Matlab (MATLAB 8.0 [R2016a] and Statistics Toolbox 8.1, The MathWorks, Inc., Natick, MA) (Arivazhagan, Ganesan, & Priyal, 2006; Lee, 1996). Figure 8 shows an example. For each of the 24 decompositions of the target and 100% real representation, the difference in the values of corresponding wavelet coefficients was computed to obtain their standard deviation. Then, the 24 standard deviations were averaged to derive the realism inconsistency of each representation.

Parameter estimates

Figure 9 presents our model, reporting the standardized path coefficients. The model depicts a causal relation from the stimulus variable *realism inconsistency* to the perceptual variables *realism* and *familiarity* (and from realism to familiarity) to the affective variables *eeriness* and *warmth*.

Model fit and parameter estimates were computed with maximum likelihood estimation using the *lavaan* package in R (Rosseel, 2012). The model had good fit (root mean square error of approximation [RMSEA] <

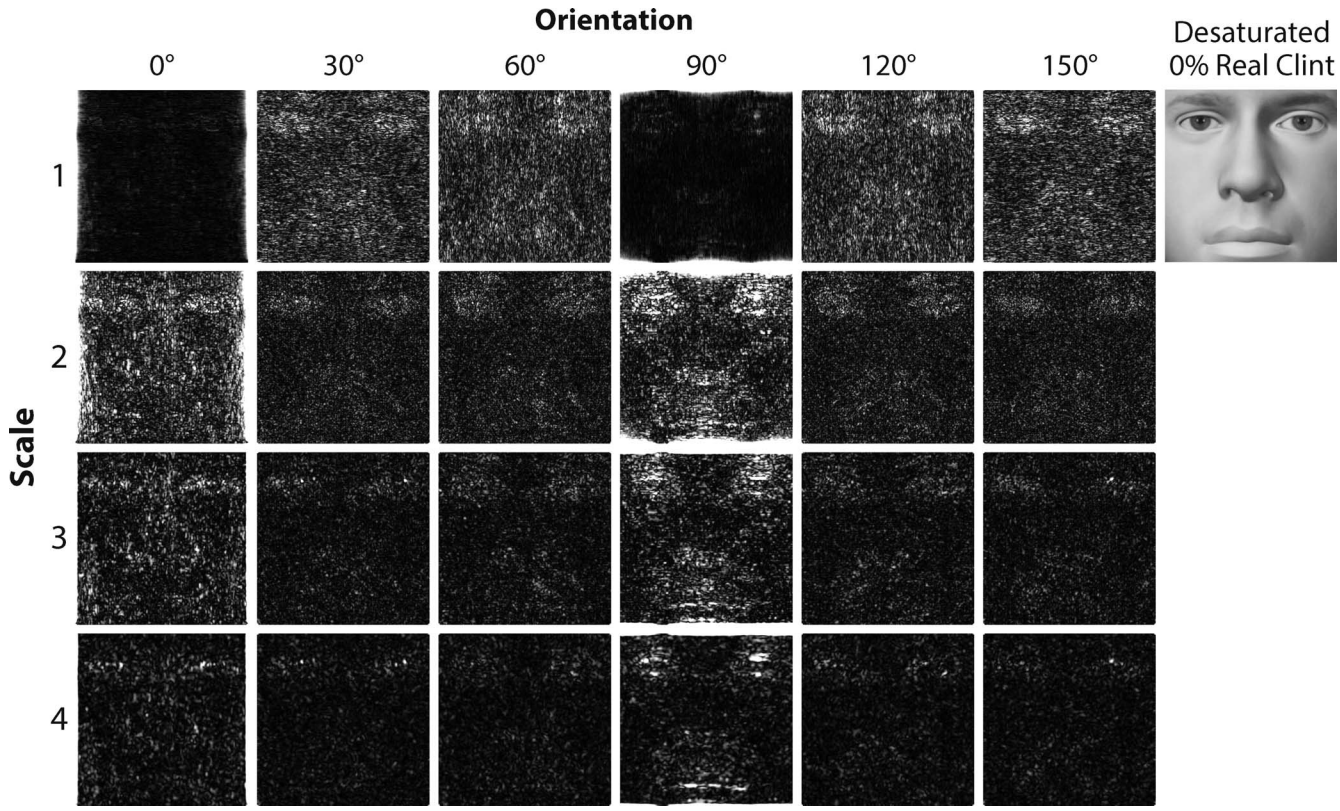

Figure 8. The desaturated 0% real representation of Clint was decomposed in 2-D Morlet basis functions at four scales (1, 2, 3, and 4) and six orientations (0° to 150° at 30° intervals). Coefficients are visualized from black to white for scales 1–4 in the range 0–0.01, 0–0.5, 0–1.0, and 0–2.0, respectively.

.06, nonnormed fit index [NNFI] ≥ .95, comparative fit index [CFI] ≥ .90; Hu & Bentler, 1999) for humans (n = 11,070, df = 2, χ^2 = 59.53, p < .001, RMSEA = .051, NNFI = .978, CFI = .996, expected cross validation index [ECVI] = .0008) but poor fit for animals (n = 4,002, df = 2, χ^2 = 357.91, p < .001, RMSEA = .211, NNFI = .750, CFI = .950, ECVI = .095) and objects (n = 9,055, df = 2, χ^2 = 675.87, p < .001, RMSEA = .193, NNFI = .406, CFI = .881, ECVI = .077).

We proposed that perceptual narrowing enhances the discriminability of physically similar anthropomorphic entities and, thus, small deviations from norms can cause an entity to appear unfamiliar. We further proposed that, owing to threat avoidance, anthropomorphic entities that look fake or unfamiliar elicit cold, eerie feelings and avoidance behavior. Thus, we hypothesized that anthropomorphism increases the inhibitory effect of realism inconsistency on familiarity and of familiarity on eeriness. The structural equation models support these hypotheses.

As the structural equation models show, perceived realism increases familiarity, especially the familiarity of humans and animals. This may explain why studies that did not manipulate perceptual tension and operationalized affinity solely as familiarity were unable to find an uncanny valley effect (Cheetham et al., 2014; Cheetham et al., 2015).

Discussion

Perceptual tension and perceived familiarity predict the uncanny valley

Moore (2012) proposed a Bayesian model of the affinity curve in Mori's (1970/2012) uncanny valley graph. The model resolves an apparent incompatibility between category uncertainty and perceptual conflict theories of why human replicas appear uncanny. The uncanniness is not caused by uncertainty about an entity's category but by perceptual tension, that is, unequal levels of uncertainty about the category of the entity's features. In Moore's (2012) model, a viewer's affinity for a stimulus is graphed by subtracting from its probability of occurrence the perceptual tension it elicits, weighted by the viewer's sensitivity to perceptual tension. But using probability of occurrence to operationalize perceived familiarity fails to account for perceptual narrowing during early

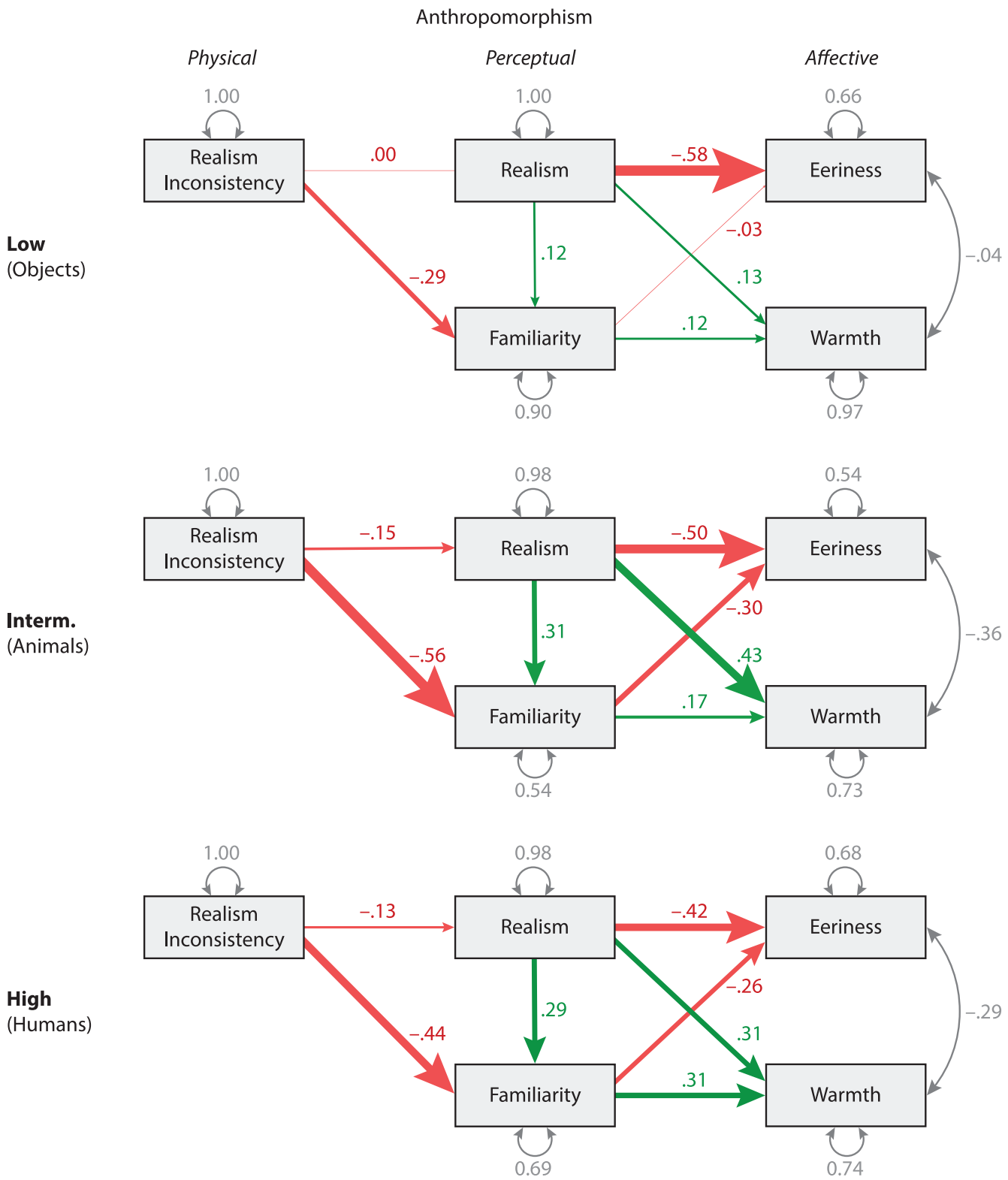

Figure 9. To explore the relations among physical, perceptual, and affective variables at low, intermediate, and high levels of anthropomorphism, a structural equation model was calculated for objects, animals, and humans, respectively. Only the model for humans had a good fit (RMSEA = .051, NNFI = .978, CFI = .996). All standardized gammas were significant ($p < .001$) except realism inconsistency → realism for objects.

childhood development. Our pilot study found perceived familiarity in fact deviates from probability of occurrence.

Thus, in this study, we revised and quantified Moore's (2012) model of affinity, replacing probability of occurrence with perceived familiarity:

$$F[S] = U[S] - kV[S] \quad (3)$$

where $F[S]$ represents affinity for stimulus S, $U[S]$ familiarity with it, k the viewer's sensitivity to perceptual tension, and $V[S]$ the perceptual tension, operationalized as reduced realism consistency.

Affinity, as predicted by the revised model, tracked perceived eeriness (reverse scaled) more closely for humans (mean Fréchet distance 0.11) than for animals (0.24) or objects (0.34, Figures 4 through 6). These results indicate higher levels of anthropomorphism cause perceptual tension to inhibit perceived familiarity, thus amplifying cold, eerie feelings.

The effect of perceptual narrowing on familiarity

The familiarity of representations decreased as anthropomorphism increased (Figure 7). This result can be explained by perceptual narrowing (Pascalis et al., 2002). Perceptual narrowing enhances perceptual discrimination, rendering abnormalities more apparent. Perceptual narrowing was strongest in perceiving humans and animals, which could enhance the sensitivity of threat avoidance mechanisms to their replicas. For humans the difference in realism between the 100% real representation and the rest (2.04 vs. −0.87) was much greater than for objects (1.18 vs. −0.26); the difference in familiarity followed a similar pattern (100% real humans 1.19 vs. rest 0.53, 100% real objects 1.56 vs. rest 1.42).

Perceptual narrowing may make it more difficult to model realistic looking eyes, skin, fur, or feathers than parts of plants and artifacts. Indeed, in comparing 3-D computer models, realism ratings for objects were almost double those for humans and animals, indicating viewers find object models far more realistic than human and animal models. A similar trend appeared for familiarity. In sum, computer-modeled features did not reduce the realism and familiarity of object models as much as they reduced the realism and familiarity of human and animal models.

We hypothesized that, due to perceptual narrowing, reduced realism consistency would decrease familiarity more in anthropomorphic (humans and animals) than nonanthropomorphic entities (objects). Our results bore this out (Figure 7 and the Testing H1 section). For both humans and animals, the familiarity of representations along the consistency-reduced transitions was significantly lower than for paired representations along the control transition. This effect was not found in perceiving objects. Thus, H1 was supported. In addition, reduced realism consistency increased the uncanny valley effect only when it also reduced perceived familiarity (see the Testing H2 section). Again, this effect was not found in perceiving objects. Thus, H2 was also supported.

A role for threat avoidance

We assume 3-D computer models are inherently realism inconsistent. Realism varies by feature, because some features are more difficult than others to model, texture, light, and render realistically. When a model falls short of 100% realism, it is highly improbable that every feature would fall short to the same degree. This means only the 100% real representation is realism consistent. The 100% computer-modeled representation and the 15 representations derived from it are all realism inconsistent. The consistency reduction manipulation used in the experiment only offers a relative comparison of consistency between paired representations in the consistency-reduced and control transitions.

To address this limitation, we sought to define realism inconsistency objectively, drawing inspiration from how the primary visual cortex represents physical properties of a stimulus at multiple scales and orientations. Both the target and 100% real representation were decomposed in 2-D Morlet basis functions (Gabor kernels), which are analogous to the receptive field profiles of single neurons (Daugman, 1980, 1985). Each basis function's wavelet coefficient, analogous to a neuron's firing rate, for the target representation was subtracted from the corresponding wavelet coefficient for the 100% real representation. The standard deviation of these differences provided a measure of realism inconsistency. This enabled us to explore the relation among physical, perceptual, and affective variables in structural equation models at three levels of anthropomorphism (Figure 9).

The structural equation models indicate that perceptual narrowing alone cannot explain the uncanny valley effect. As an indicator of perceptual narrowing, realism inconsistency inhibited familiarity more in perceiving humans ($\gamma = -.44$) than objects but, nevertheless, still inhibited familiarity in perceiving objects ($\gamma = -.29$). However, familiarity had a negligible effect on the eeriness ($\gamma = -.03$) and only a small effect on the warmth ($\gamma = .12$) of objects. By contrast, familiarity inhibited the eeriness ($\gamma = -.26$) of humans and increased their warmth ($\gamma = .31$) with a medium effect size. Judging from the models, even an unfamiliar object would be unlikely to elicit cold, eerie feelings. These results suggest that, although perceptual narrowing is likely to make abnormalities more salient in anthropomorphic entities, an

additional mechanism is necessary to explain why those abnormalities are perceived—not only as unfamiliar—but as strange and uncanny.

Elsewhere we have proposed a mechanism for threat avoidance could be elicited by human replicas (MacDorman & Entezari, 2015; MacDorman & Ishiguro, 2006; MacDorman et al., 2009a; MacDorman et al., 2009b). However, an important area to explore is how cold, eerie feelings differ from other emotions that motivate threat avoidance, such as fear and disgust. Mangan (2015) explained these feelings as fringe experience, whereby a threat is unconsciously sensed, but its cause is never consciously identified. This analysis echoes Freud's (1919/2003) conception of the uncanny as something familiar, but long repressed, which elicits morbid anxiety. Ohman (2000) identified this kind of anxiety as undirected alarm caused by preattentive mechanisms falsely locating a threat that remains unconfirmed and unresolved (e.g., by a fight or flight response, cf. Burleigh et al., 2013; Misselhorn, 2009; Nesse, 2005; Zuckerman & Spielberger, 2015). The relation between the uncanny valley, its peculiar phenomenology, and the mechanism of threat avoidance are left to future work.

Keywords: anthropomorphism, computer animation, face perception, familiarity, Gabor kernels, perceptual narrowing, threat avoidance

Acknowledgments

We would like to thank Tyler J. Burleigh, Asif A. Ghazanfar, Stevan Harnad, Bruce B. Mangan, Maya B. Mathur, Roger K. Moore, Frank E. Pollick, Christopher H. Ramey, Bertrand Tondu, and the anonymous reviewers for helpful advice on revising the manuscript. We are grateful to Ryan Sukale for developing and maintaining the experimental applications and website used in this study. We greatly appreciate those who have created the 3-D models and artwork used in this study: Clinton T. Koch for "Emelie," Guile Lindroth (CEO, NextOS, Curitiba, Brazil, www.nextos.com) for "Blond Ann Avatar," Vivek Varma Kosuri (Rapra Designing Solutions, Hyderabad, India, www.rapra.in) for "Nikon," "washing machine," and "water lily," and Thomas Rudat (Illustrator and 3-D Artist, Hannover, Germany, thomas-rudat.cgsociety.org) for "Simona." The following royalty free models were obtained from TurboSquid (www.turbosquid.com): "Ferrari 599 GTB 2006" (ID 483862), "Saint Bernard" (ID 421846), and "Lovebird" (ID 659591). We would also like to thank Clinton T. Koch for the photographs "Clint" and "Emelie." The following royalty photographs were obtained: "Saint Bernard" (ID 10629611) from 123RF (www.123rf.com) and "Ingrid" (ID 209863), "Juliana" (ID 461057), "Simona" (ID 60775) and "Zlatko" (ID 473816) from 3D.sk. This research is supported by the US National Institutes of Health (P20 GM066402).

Commercial relationships: none.
Corresponding author: Karl F. MacDorman.
Email: kmacdorm@indiana.edu.
Address: Indiana University School of Informatics and Computing, Indianapolis, IN, USA.

References

Alt, H., & Godau, M. (1995). Computing the Fréchet distance between two polygonal curves. *International Journal of Computational Geometry & Applications*, *5*(1), 75–91, doi:10.1142/S0218195995000064.

Arivazhagan, S., Ganesan, L., & Priyal, S. P. (2006). Texture classification using Gabor wavelets based rotation invariant features. *Pattern Recognition Letters*, *27*(16), 1976–1982, doi:10.1016/j.patrec.2006.05.008.

Berlyne, D. E. (1970). Novelty, complexity and hedonic value. *Perception and Psychophysics*, *8*, 279–286, doi:10.3758/BF03212593.

Berlyne, D. E. (1971). *Aesthetics and psychobiology*. New York: Appleton-Century-Crofts.

Bronson, G. W. (1968a). The development of fear in man and other animals. *Child Development*, *39*(2), 409–431, doi:10.2307/1126955.

Bronson, G. W. (1968b). The fear of novelty. *Psychological Bulletin*, *69*(5), 350–358, doi:10.1037/h0025706.

Burleigh, T. J., & Schoenherr, J. R. (2015). A reappraisal of the uncanny valley: Categorical perception or frequency-based sensitization? *Frontiers in Psychology*, *5*(1488), 1–19, doi:10.3389/fpsyg.2014.01488.

Burleigh, T. J., Schoenherr, J. R., & Lacroix, G. L. (2013). Does the uncanny valley exist? An empirical test of the relationship between eeriness and the human likeness of digitally created faces. *Computers in Human Behavior*, *29*(3), 759–771, doi:10.1016/j.chb.2012.11.021.

Cheetham, M., Pavlovic, I., Jordan, N., Suter, P., & Jäncke, L. (2013). Category processing and the human likeness dimension of the uncanny valley hypothesis: Eye-tracking data. *Frontiers in Psychology*, *4*(108), 1–12, doi:10.3389/fpsyg.2013.00108.

Cheetham, M., Suter, P., & Jäncke, L. (2014).

Perceptual discrimination difficulty and familiarity in the uncanny valley: More like a 'happy valley.' *Frontiers in Psychology*, *5*(1219), 1–15, doi:10.3389/fpsyg.2014.01219.

Cheetham, M., Wu, L., Pauli, P., & Jäncke, L. (2015). Arousal, valence, and the uncanny valley: Psychophysiological and self-report findings. *Frontiers in Psychology*, *6*(981), 1–15, doi:10.3389/fpsyg.2015.00981.

Cohen, J. (1992). A power primer. *Psychological Bulletin*, *112*(1), 155–159, doi:10.1037/0033-2909.112.1.155.

Cuddy, A. J. C., Fiske, S. T., & Glick, P. (2007). The BIAS map: Behaviors from intergroup affect and stereotypes. *Journal of Personality and Social Psychology*, *92*(4), 631–648, doi:10.1037/0022-3514.92.4.631.

Curtis, V., Aunger, R., & Rabie, T. (2004). Evidence that disgust evolved to protect from risk of disease. *Proceedings of the Royal Society of London: Biological Sciences*, *271*(Suppl. 4), S131–S133, doi:10.1098/rsbl.2003.0144.

Daugman, J. G. (1980). Two-dimensional spectral analysis of cortical receptive field profiles. *Vision Research*, *20*(10), 847–856, doi:10.1016/0042-6989(80)90065-6

Daugman, J. G. (1985). Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by two-dimensional visual cortical filters. *Journal of the Optical Society of America A*, *2*(7), 1160–1169, doi:10.1364/JOSAA.2.001160.

Feldman, N. H., Griffiths, T. L., & Morgan, J. L. (2009). The influence of categories on perception: Explaining the perceptual magnet effect as optimal statistical inference. *Psychological Review*, *116*(4), 752–782, doi:10.1.1.211.3309.

Ferrey, A. E., Burleigh, T. J., & Fenske, M. J. (2015). Stimulus-category competition, inhibition, and affective devaluation: A novel account of the uncanny valley. *Frontiers in Psychology*, *6*(249), 1–15, doi:10.3389/fpsyg.2015.00249.

Freud, S. (2003). *The uncanny* (D. McLintock, Trans.). New York: Penguin. (Original work published in 1919)

Gottlieb, G. (2002). Developmental-behavioral initiation of evolutionary change. *Psychological Review*, *109*(2), 211–218, doi:10.1037/0033-295X.109.2.211.

Hanson, D. (2005). Expanding the aesthetic possibilities for humanoid robots. *Proceedings of the Views of the Uncanny Valley Workshop*, IEEE-RAS International Conference on Humanoid Robots, Tsukuba, Japan.

Harnad, S. (1987). Category induction and representation. In S. Harnad (Ed.), *Categorical perception: The groundwork of cognition* (pp. 535–565). New York: Cambridge University Press.

Hebb, D. O. (1946). On the nature of fear. *Psychological Review*, *53*(5), 259–276, doi:10.1037/h0061690.

Ho, C.-C., & MacDorman, K. F. (2010). Revisiting the uncanny valley theory: Developing and validating an alternative to the Godspeed indices. *Computers in Human Behavior*, *26*(6), 1508–1518, doi:10.1016/j.chb.2010.05.015.

Ho, C.-C., MacDorman, K. F., & Pramono, Z. D. (2008). Human emotion and the uncanny valley: A GLM, MDS, and Isomap analysis of robot video ratings. In *Proceedings of the Third ACM/IEEE International Conference on Human–Robot Interaction* (pp. 169–176). New York: ACM, doi:10.1145/1349822.1349845.

Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, *6*(1), 1–55, doi:10.1080/10705519909540118.

Iverson, P., & Kuhl, P. K. (1995). Mapping the perceptual magnet effect for speech using signal detection theory and multidimensional scaling. *Journal of the Acoustical Society of America*, *97*(1), 553–562, doi:10.1121/1.412280.

James, T. W., Potter, R. F., Lee, S. K., Kim, S., Stevenson, R. A., & Lang, A. (2015). How realistic should avatars be? An initial fMRI investigation of activation of the face perception network by real and animated faces. *Journal of Media Psychology*, *27*(3), 109–117, doi:10.1027/1864-1105/a000156.

Jentsch, E. (1997). *Zur Psychologie des Unheimlichen* [On the psychology of the uncanny] (R. Sellars, Trans.). *Angelaki*, *2*(1), 7–16. (Original work published in 1906), doi:10.1080/09697259708571910.

Kätsyri, J., Förger, K., Mäkäräinen, M., & Takala, T. (2015). A review of empirical evidence on different uncanny valley hypotheses: Support for perceptual mismatch as one road to the valley of eeriness. *Frontiers in Psychology*, *6*(390), 1–16, doi:10.3389/fpsyg.2015.00390.

Kelly, D. J., Quinn, P. C., Slater, A. M., Lee, K., Ge, L., & Pascalis, O. (2007). The other-race effect develops during infancy: Evidence of perceptual narrowing. *Psychological Science*, *18*(12), 1084–1089, doi:10.1111/j.1467-9280.2007.02029.x.

Lang, A., Bradley, S. D., Sparks, J. V., Jr., & Lee, S. (2007). The motivation activation measure (MAM): How well does MAM predict individual differences

in physiological indicators of appetitive and aversive activation? *Communication Methods and Measures*, *1*(2), 113–136, doi:10.1080/19312450701399370.

Lee, T. S. (1996). Image representation using 2D Gabor wavelets. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *18*(10), 959–971, doi:10.1109/34.541406.

Lewkowicz, D. J., & Ghazanfar, A. A. (2006). The decline of cross-species intersensory perception in human infants. *Proceedings of the National Academy of Sciences*, *103*(17), 6771–6774, doi:10.1073/pnas.0602027103.

Looser, C. E., & Wheatley, T. (2010). The tipping point of animacy: How, when, and where we perceive life in a face. *Psychological Science*, *21*(12), 1854–1862, doi:10.1177/0956797610388044.

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

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

MacDorman, K. F., Green, R. D., Ho, C.-C., & Koch, C. (2009a). Too real for comfort: Uncanny responses to computer generated faces. *Computers in Human Behavior*, *25*(3), 695–710, doi:10.1016/j.chb.2008.12.026.

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

MacDorman, K. F., Vasudevan, S. K., & Ho, C.-C. (2009b). Does Japan really have robot mania? Comparing attitudes by implicit and explicit measures. *AI & Society*, *23*(4), 485–510, doi:10.1007/s00146-008-0181-2.

Mangan, B. B. (2015). The uncanny valley as fringe experience. *Interaction Studies*, *16*(2), 193–199, doi:10.1075/is.16.2.05man.

Mathur, M. B., & Reichling, D. B. (2016). Navigating a social world with robot partners: A quantitative cartography of the uncanny valley. *Cognition*, *146*, 22–32, doi:10.1016/j.cognition.2015.09.008.

Meah, L. F. S., & Moore, R. K. (2014). The uncanny valley: A focus on misaligned cues. In M. Beetz, B. Johnston, & M.-A. Williams (Eds.), *Social Robotics* (LNAI, vol. 8755, pp. 256–265). Cham, Switzerland: Springer, doi:10.1007/978-3-319-11973-1_26

Misselhorn, C. (2009). Empathy with inanimate objects and the uncanny valley. *Minds and Machines*, *19*, 345–459, doi:10.1007/s11023-009-9158-2.

Mitchell, W. J., Szerszen, Sr., K. A., Lu, A. S., Schermerhorn, P. W., Scheutz, M., & MacDorman, K. F. (2011). A mismatch in the human realism of face and voice produces an uncanny valley. *i-Perception*, *2*(1), 10–12, doi:10.1068/i0415.

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

Moosa, M. M., & Ud-Dean, S. M. M. (2010). Danger avoidance: An evolutionary explanation of the uncanny valley. *Biological Theory*, *5*(1), 12–14, doi:10.1162/BIOT_a_00016.

Mori, M. (2012). The uncanny valley (K. F. MacDorman & N. Kageki, Trans.). *IEEE Robotics and Automation*, *19*(2), 98–100. (Original work published in 1970) doi:10.1109/MRA.2012.2192811

Nesse, R. M. (2005). Natural selection and the regulation of defenses. A signal detection analysis of the smoke detector principle. *Evolution and Human Behavior*, *26*(1), 88–105, doi:10.1016/j.evolhumbehav.2004.08.002.

Ohman, A. (2000). Fear and anxiety: Evolutionary, cognitive, and clinical perspectives. In *Handbook of emotions* (2nd ed., pp. 573–593). New York: Guilford Press.

Oshio, A. (2009). Development and validation of the dichotomous thinking inventory. *Social Behavior and Personality*, *37*(6), 729–741, doi:10.2224/sbp.2009.37.6.729.

Park, J. H., Faulkner, J., & Schaller, M. (2003). Evolved disease-avoidance processes and contemporary anti-social behavior: Prejudicial attitudes and avoidance of people with disabilities. *Journal of Nonverbal Behavior*, *27*(2), 65–87, doi:10.1023/A:1023910408854.

Pascalis, O., de Haan, M., & Nelson, C. A. (2002). Is face processing species-specific during the first year of life? *Science*, *296*(5571), 1321–1323, doi:10.1126/science.1070223.

Pollick, F. E. (2010). In search of the uncanny valley. *Lecture Notes of the Institute for Computer Sciences*: *Social Informatics and Telecommunications Engineering*, *40*(4), 69–78, doi:10.1007/978-3-642-12630-7_8.

Ramey, C. H. (2005). The uncanny valley of similarities concerning abortion, baldness, heaps of sand, and humanlike robots. *Views of the Uncanny Valley Workshop, IEEE-RAS International Conference on Humanoid Robots*, Tsukuba, Japan.

Rosseel, Y. (2012). lavaan: An R package for structural equation modeling. *Journal of Statistical Software*, *48*(2), 1–36, doi:10.18637/jss.v048.i02.

Saygin, A. P., Chaminade, T., Ishiguro, H., Driver, J., & Frith, C. (2012). The thing that should not be: Predictive coding and the uncanny valley in perceiving human and humanoid robot actions. *Social Cognitive and Affective Neuroscience*, *7*(4), 413–422, doi:10.1093/scan/nsr025.

Scott, L. S., Pascalis, O., & Nelson, C. A. (2007). A domain-general theory of the development of perceptual discrimination. *Current Directions in Psychological Science*, *16*(4), 197–201, doi:10.1111/j.1467-8721.2007.00503.x.

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

Simpson, E. A., Varga, K., Frick, J. E., & Fragaszy, D. (2010). Infants experience perceptual narrowing for nonprimate faces. *Infancy*, *16*(3), 318–328, doi:10.1111/j.1532-7078.2010.00052.x.

Thode, H. C., Jr. (2002). Testing for normality. In *Statistics: Textbooks and monographs* (vol. 164). New York: CRC.

Tondu, B., & Bardou, N. (2011). A new interpretation of Mori's uncanny valley for future humanoid robots. *International Journal of Robotics and Automation*, *26*(3), 337–348, doi:10.2316/Journal.206.2011.3.206-3348.

Turk, M., & Pentland, A. P. (1991). Face recognition using eigenfaces. In *Proceedings of the IEEE Computer Society on Computer Vision and Pattern Recognition* (pp. 586–591). Piscataway, NJ: IEEE, doi:10.1109/CVPR.1991.139758

van Kampen, H. S. (2015). Violated expectancies: Cause and function of exploration, fear, and aggression. *Behavioural Processes*, *117*, 12–28, doi:10.1016/j.beproc.2014.05.005.

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

Zuckerman, M. (1976). Sensation seeking and anxiety, traits and states, as determinants of behavior in novel situations. In I. G. Sarason & C. D. Spielberger (Eds.), *Stress and anxiety* (pp. 141–170). Oxford, England: Hemisphere.

Zuckerman, M. (2014). *Sensation seeking: Beyond the optimal level of arousal*. Chicago: Psychology Press.

Zuckerman, M., & Spielberger, C. D. (Eds.). (2015). *Emotions and anxiety: New concepts, methods, and applications*. New York: Psychology Press.

Appendix A: Perceptual magnet effect

Following Feldman, Griffiths, and Morgan's (2009, appendix A) calculations for the expected category of the target stimulus, the perceptual magnet effect along a single dimension is calculated as

$$D[S] = E[T|S] - S \quad \text{(A1)}$$

where $D[S]$ is the displacement function (Equation 2) and $E[T|S]$ is the expected value of the perceptual target T when a physical stimulus S is observed. Numerically, S consists of the values describing the feature set in the observational dimensions. For multiple categories, this expected value is

$$E[T|S] = \sum_c p(c|S)\frac{\sigma_c^2 S + \sigma_S^2 \mu_c}{\sigma_c^2 + \sigma_S^2} \quad \text{(A2)}$$

which can be rewritten as

$$E[T|S] = \frac{\sigma_c^2}{\sigma_c^2 + \sigma_S^2} S + \frac{\sigma_S^2}{\sigma_c^2 + \sigma_S^2} \sum_c p(c|S)\mu_c \quad \text{(A3)}$$

which is derived from the posterior probability of membership of a given category c, with μ_c as the category mean, σ_c^2 as the category variance, and σ_S^2 as the uncertainty of the observed signal. The posterior probability, according to Bayes' theorem is

$$p(c|S) = \frac{p(S|c)p(c)}{\sum_c p(S|c)p(c)} \quad \text{(A4)}$$

where the stimulus belongs to the category c with a normal distribution N:

$$S|c \sim N(\mu_c, \sigma_c^2 + \sigma_S^2) \quad \text{(A5)}$$

$D[S]$ measures the perceptual distortion of a stimulus towards or away from a category boundary along a given dimension. A zero value indicates no perceptual distortion.

Appendix B: Category parameters from identification curves

In the two-alternative forced choice categorization task, participants categorized stimuli as either computer animated or real. The percentages categorized as

real were used to plot the logistic identification curves. To identify the category boundary, the gain g and the bias b of the logistic function were calculated (Feldman, Griffiths, & Morgan, 2009). The mean of the category *computer animated* was then computed using g, b, and the mean of the category *real* (see Equation B4).

The logistic identification function is an empirical measure of $p(c_1|S)$, which according to Bayes' rule is

$$p(c_1|S) = \frac{p(S|c_1)p(c_1)}{p(S|c_1)p(c_1) + p(S|c_2)p(c_2)} \quad \text{(B1)}$$

Dividing each part of the fraction by the numerator and applying two inverse functions to the last term results in

$$p(c_1|S) = \frac{1}{1 + \exp\left(\log \frac{p(S|c_2)p(c_2)}{p(S|c_1)p(c_1)}\right)} \quad \text{(B2)}$$

Now assuming that both categories c_1 and c_2 have equal prior probability, and using the distribution for $p(S|c)$ from Equation A5, we can simplify Equation B2 to

$$p(c_1|S) = \frac{1}{1 + e^{-gS+b}} \quad \text{(B3)}$$

where $g = (\mu_1 - \mu_2) / (\sigma_c^2 + \sigma_S^2)$ and $b = (\mu_1^2 - \mu_2^2) / (2(\sigma_c^2 + \sigma_S^2))$. Hence, we can compute the mean of the second category μ_2 given the values of g, b, μ_1 as

$$\mu_2 = \frac{2b}{g} - \mu_1 \quad \text{(B4)}$$

We used Equation B4 to calculate the mean of the category *computer animated,* and Equation A3 to calculate perceptual tension with σ_S^2/σ_c^2 set to 0.1. Percentages categorized as *computer animated* and *real* were used for $p(c_1|S)$ and $p(c_2|S)$, respectively.

Appendix C: Response times and categorization responses

A nested-factor ANOVA (categorization responses ~ transition/fraction of real with five levels of fraction of real, except the two endpoints) found that transition significantly affected categorization responses: Humans, $F(2, 9{,}765) = 24.9$, $MSE = 0.18$, $p < .001$, $\eta_p^2 = .005$; animals, $F(2, 3{,}525) = 70.48$, $MSE = 0.15$, $p < .001$, $\eta_p^2 = .038$; and objects, $F(2, 7{,}980) = 44.51$, $MSE = 0.24$, $p < .001$, $\eta_p^2 = .011$. However, a similar nested-factor ANOVA found that transition significantly affected response times ($\log_{10}$-transformed) in perceiving humans, $F(2, 19{,}500) = 13.11$, $MSE = 0.04$, $p < .001$, $\eta_p^2 = .001$, and animals, $F(2, 7{,}035) = 4.00$, $MSE = 0.03$, $p = .018$, $\eta_p^2 = .001$, but not objects, $F(2, 15{,}930) = 1.27$, $p = .282$. Figure C1 shows how categorization responses and response times varied with the fraction of real in humans, animals, and objects. The figures depict a cubic spline interpolating the mean responses to facilitate a visual comparison with the standard S-shaped logistic function of categorical perception (Harnad, 1987).

Appendix D: Comparison of predicted affinity and perceived warmth

Figure D1 compares predicted affinity and perceived warmth for matched stimuli at low, intermediate, and high levels of anthropomorphism for one control and two consistency-reduced transitions. The mean distance between predicted affinity and perceived warmth for all three transitions was, in perceiving objects, diagonal, $L^1 = -1.46$, lower right, $L^1 = -1.59$, and upper left, $L^1 = -1.41$; animals: diagonal, $L^1 = -0.77$, lower right, $L^1 = -0.66$, and upper left, $L^1 = -0.30$; and humans: diagonal, $L^1 = -0.61$, lower right, $L^1 = -0.17$, and upper left, $L^1 = -0.53$. Curves were similar for all three transitions in perceiving objects: diagonal, $d = 0.64$, lower right, $d = 0.77$, and upper left, $d = 0.60$; animals: diagonal, $d = 2.12$, lower right, $d = 2.01$, and upper left, $d = 2.05$; and humans: diagonal, $d = 1.63$, lower right, $d = 1.72$, and upper left, $d = 1.55$.

Appendix E: Comparison of perceived familiarity and warmth

In perceiving objects, fraction of real also significantly affected warmth ratings for each transition: diagonal, $F(4.83, 2{,}540) = 15.73$, $MSE = 0.51$, $p < .001$, $\eta_p^2 = .03$; lower right, $F(4.93, 2{,}620) = 16.09$, $MSE = 0.49$, $p < .001$, $\eta_p^2 = .03$; and upper left, $F(5.02, 2{,}666) = 16.98$, $MSE = 0.46$, $p < 0.001$, $\eta_p^2 = .03$. Pairwise comparisons showed that for objects reduced realism consistency did not significantly affect warmth or familiarity for any pairs in either the lower-right or upper-left transition.

In perceiving animals, fraction of real also significantly affected warmth ratings for each transition: diagonal, $F(4.16, 957) = 69.20$, $MSE = 1.50$, $p < .001$, $\eta_p^2 = .23$; lower right, $F(4.57, 1{,}064) = 67.79$, $MSE = 1.32$, $p < .001$, $\eta_p^2 = .23$; and upper left, $F(3.44, 794) = 61.59$, $MSE = 2.07$, $p < .001$, $\eta_p^2 = .21$. Pairwise comparisons showed that reduced realism consistency significantly decreased warmth for one pair in the lower-right transition: 1/2–1/2 and 0–1; and three pairs in the upper-left transition: 1/3–1/3 and 2/3–0, 1/2–1/2 and 1–0, and 2/3–2/3 and 1–1/3 (Table E1).

In perceiving humans, fraction of real also significantly affected warmth ratings for each transition: diagonal, $F(3.42, 2{,}211) = 167$, $MSE = 1.77$, $p < .001$, $\eta_p^2 = .21$; lower right, $F(3.88, 2{,}499) = 152$, $MSE = 1.56$, $p < .001$, $\eta_p^2 = .19$; and upper left, $F(3.99, 2{,}574) = 128$, $MSE = 1.44$, $p < .001$, $\eta_p^2 = .17$. Pairwise comparisons showed that reduced realism consistency significantly decreased warmth for four pairs in the lower-right transition, 1/3–1/3 and 0–2/3, 1/2–1/2 and 0–1, 2/3–2/3 and 1/3–1, and 5/6–5/6 and 2/3–1, and three pairs in the upper-left transition, 1/2–1/2 and 1–0, 2/3–2/3 and 1–1/3, and 5/6–5/6 and 1–2/3 (Table E2).

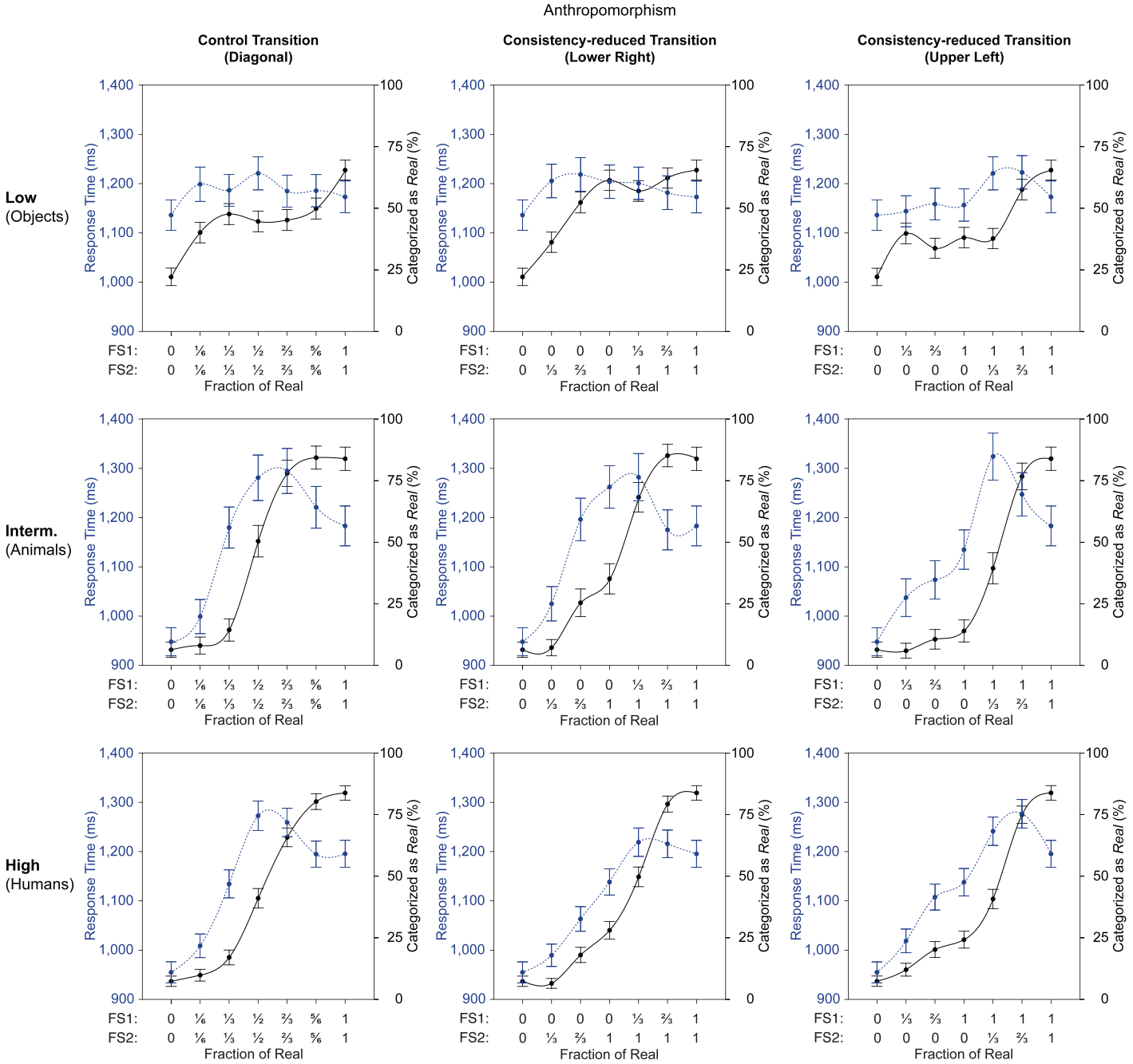

Figure C1. For objects, animals, and humans, the mean response time in milliseconds (dashed curve) and percentage of times the stimulus was categorized as *real* (vs. *computer animated,* solid curve) are plotted against its fraction of real for the control (diagonal) and reduced-consistency transitions (lower right, upper left). Error bars indicate the 95% confidence interval of the true mean.

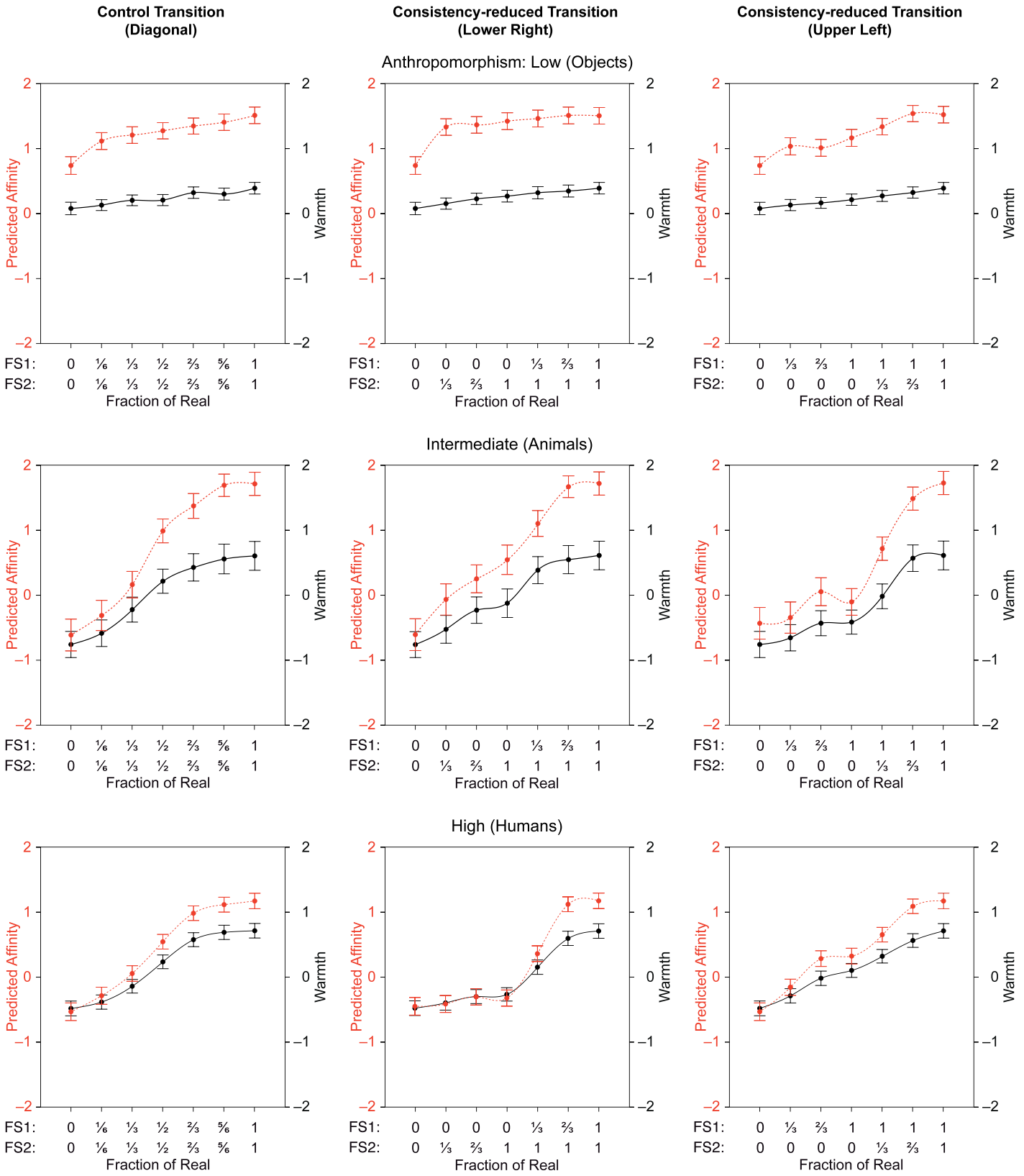

Figure D1. For objects, animals, and humans, warmth ratings of the stimulus are plotted against its fraction of real for the control (diagonal) and the consistency-reduced transitions (lower right, upper left). The dashed line represents affinity as predicted by the revised Bayesian model. Error bars indicate the 95% confidence interval of the true mean.

Consistency-reduced vs. control transition	Familiarity				Warmth			
	t	*df*	*p*	*d*	*t*	*df*	*p*	*d*
0−1/3 vs. 1/6−1/6	−1.19	234	.117		−0.82	234	.206	
0−2/3 vs. 1/3−1/3	2.17	235	.016		0.06	235	.475	
0−1 vs. 1/2−1/2	5.45	235	<.001	0.35	4.15	235	<.001	0.27
1/3−1 vs. 2/3−2/3	3.22	235	.001	0.21	0.51	235	.307	
2/3−1 vs. 5/6–5/6	0.70	233	.243		0.43	233	.334	
1/3−0 vs. 1/6–1/6	1.49	234	.069		0.86	234	.196	
2/3−0 vs. 1/3−1/3	0.87	235	.192		2.58	235	.006	0.17
1−0 vs. 1/2−1/2	5.95	233	<.001	0.39	6.41	233	<.001	0.42
1−1/3 vs. 2/3−2/3	4.29	233	<.001	0.28	4.46	233	<.001	0.29
1−2/3 vs. 5/6−5/6	2.06	232	.021		−0.14	232	.447	

Table E1. Familiarity and warmth for all 10 comparisons in perceiving animals. *Note*: Fraction pairs indicate the fraction of real for Feature Set 1 and 2. Effect sizes are provided for significant outcomes in Tables E1 and E2.

Consistency-reduced vs. control transition	Familiarity				Warmth			
	t	*df*	*p*	*d*	*t*	*df*	*p*	*d*
0−1/3 vs. 1/6−1/6	2.09	650	.019		0.43	650	.335	
0−2/3 vs. 1/3−1/3	3.66	649	<.001	0.14	3.42	649	<.001	0.13
0−1 vs. 1/2−1/2	8.92	650	<.001	0.35	9.46	650	<.001	0.37
1/3−1 vs. 2/3−2/3	7.36	649	<.001	0.29	8.57	648	<.001	0.34
2/3−1 vs. 5/6–5/6	0	648	.500		2.37	648	.009	0.09
1/3−0 vs. 1/6–1/6	0.38	650	.353		−2.12	650	.014	
2/3−0 vs. 1/3−1/3	−0.08	650	.470		−2.42	650	.008	
1−0 vs. 1/2−1/2	6.01	650	<.001	0.24	2.82	650	.003	0.11
1−1/3 vs. 2/3−2/3	6.40	649	<.001	0.25	5.98	649	<.001	0.23
1−2/3 vs. 5/6−5/6	1.40	649	.081		3.19	649	.001	0.13

Table E2. Familiarity and warmth for all 10 comparisons in perceiving humans.