Illusory faces are more likely to be perceived as male than female

Susan G. Wardle[a,1,2], Sanika Paranjape[a,1], Jessica Taubert[b], and Chris I. Baker[a]

[a]Section on Learning and Plasticity, Laboratory of Brain & Cognition, National Institute of Mental Health, Bethesda, MD 20814; and [b]Section on Neurocircuitry, Laboratory of Brain & Cognition, National Institute of Mental Health, Bethesda, MD 20814

Edited by Susan Fiske, Psychology Department, Princeton University, Princeton, NJ; received October 1, 2021; accepted November 17, 2021

Despite our fluency in reading human faces, sometimes we mistakenly perceive illusory faces in objects, a phenomenon known as face pareidolia. Although illusory faces share some neural mechanisms with real faces, it is unknown to what degree pareidolia engages higher-level social perception beyond the detection of a face. In a series of large-scale behavioral experiments (n_{total} = 3,815 adults), we found that illusory faces in inanimate objects are readily perceived to have a specific emotional expression, age, and gender. Most strikingly, we observed a strong bias to perceive illusory faces as male rather than female. This male bias could not be explained by preexisting semantic or visual gender associations with the objects, or by visual features in the images. Rather, this robust bias in the perception of gender for illusory faces reveals a cognitive bias arising from a broadly tuned face evaluation system in which minimally viable face percepts are more likely to be perceived as male.

face perception | gender | bias | pareidolia | face evaluation

Human faces convey a rich amount of social information beyond their identity (1–3). We are able to rapidly evaluate the age (4), gender (5, 6), and emotional expression (7) of the faces of individuals, even if they are not known to us, in addition to more abstract traits, such as trustworthiness and aggressiveness (8, 9). Although these judgements are based on visual information, biases have been identified that suggest that both perceptual and cognitive factors are involved in face evaluation (10–13). For example, people tend to judge faces as closer to their own age (10, 13), and damage to the amygdala is associated with perceiving unfamiliar faces as more trustworthy and approachable (12). Biases in face perception have important implications for understanding the neural processing of faces and their role in complex social behaviors (3). However, it is still unknown to what extent these behavioral biases arise from the tuning of the underlying face-processing mechanisms or, alternatively, from the nature of the experimental stimuli and task (10, 11). Here we approach this question from a new angle by examining face evaluation for a different class of faces: illusory faces in inanimate objects.

Face pareidolia is the spontaneous perception of illusory facial features in inanimate objects (Fig. 1), and can be thought of as a natural error of our face detection system (14–18). It has recently been shown that nonhuman primates also experience face pareidolia (14, 15), and that illusory faces engage similar neural mechanisms to real faces in the human brain (18). However, it is unclear to what degree higher-level social perception beyond the detection of a face occurs in pareidolia. Investigation of face evaluation in illusory faces has the potential to reveal new insight into the underlying mechanisms of face perception. A key feature of face pareidolia is that it involves the spontaneous perception of a face in an inanimate object, and consequently it is an example of face perception that is divorced from many characteristics that typically accompany the faces of living organisms, such as the motion of facial muscles (e.g., to form emotional expressions), chronological age, and biological sex. The primary question we address here is whether illusory faces are perceived to have these traits even in the absence of their biological specification. As there is no a priori reason why an illusory face should be perceived to have a specific age, gender or expression, any reliable perception of these attributes would be informative about inherent properties of the underlying system.

Studies using human faces have suggested potential biases in the perceived characteristics of human faces along dimensions such as age (10, 13) and gender (10, 11, 19) under conditions of visual uncertainty. However, determining the potential origin and generality of these biases has proven difficult and highlights the fundamental challenges inherent in understanding how the perception of specific traits is linked to face processing. Human faces are visually complex, and our brains are incredibly well-adapted to processing faces as a cohesive whole (20). Consequently, it is challenging to empirically isolate particular aspects of a human face (e.g., biological sex) from other interdependencies (e.g., identity). Additionally, since human faces have a biologically specified age and gender, it is necessary to introduce uncertainty via deliberate experimental manipulation of the stimuli. Studies of human faces have used various forms of image manipulation, including removing hair (21, 22), showing silhouettes of faces in profile (23), adding visual noise (24), and synthetically generating faces by morphing along stimulus dimensions, such as gender (10, 11, 19). A critical advantage of using pareidolia to probe the tuning of the face-processing system is that no decisions about stimulus manipulation need to

Significance

Face pareidolia is the phenomenon of perceiving illusory faces in inanimate objects. Here we show that illusory faces engage social perception beyond the detection of a face: they have a perceived age, gender, and emotional expression. Additionally, we report a striking bias in gender perception, with many more illusory faces perceived as male than female. As illusory faces do not have a biological sex, this bias is significant in revealing an asymmetry in our face evaluation system given minimal information. Our result demonstrates that the visual features that are sufficient for face detection are not generally sufficient for the perception of female. Instead, the perception of a nonhuman face as female requires additional features beyond that required for face detection.

Author contributions: S.G.W., S.P., J.T., and C.I.B. designed research; S.G.W. and S.P. performed research; S.G.W. and S.P. contributed new reagents/analytic tools; S.G.W., S.P., J.T., and C.I.B. analyzed data; and S.G.W., S.P., J.T., and C.I.B. wrote the paper.

The authors declare no competing interest.

This article is a PNAS Direct Submission.

[1]S.G.W. and S.P. contributed equally to this work.

[2]To whom correspondence may be addressed. Email: susan.wardle@nih.gov.

This article contains supporting information online at http://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2117413119/-/DCSupplemental.

Published January 24, 2022.

PSYCHOLOGICAL AND COGNITIVE SCIENCES

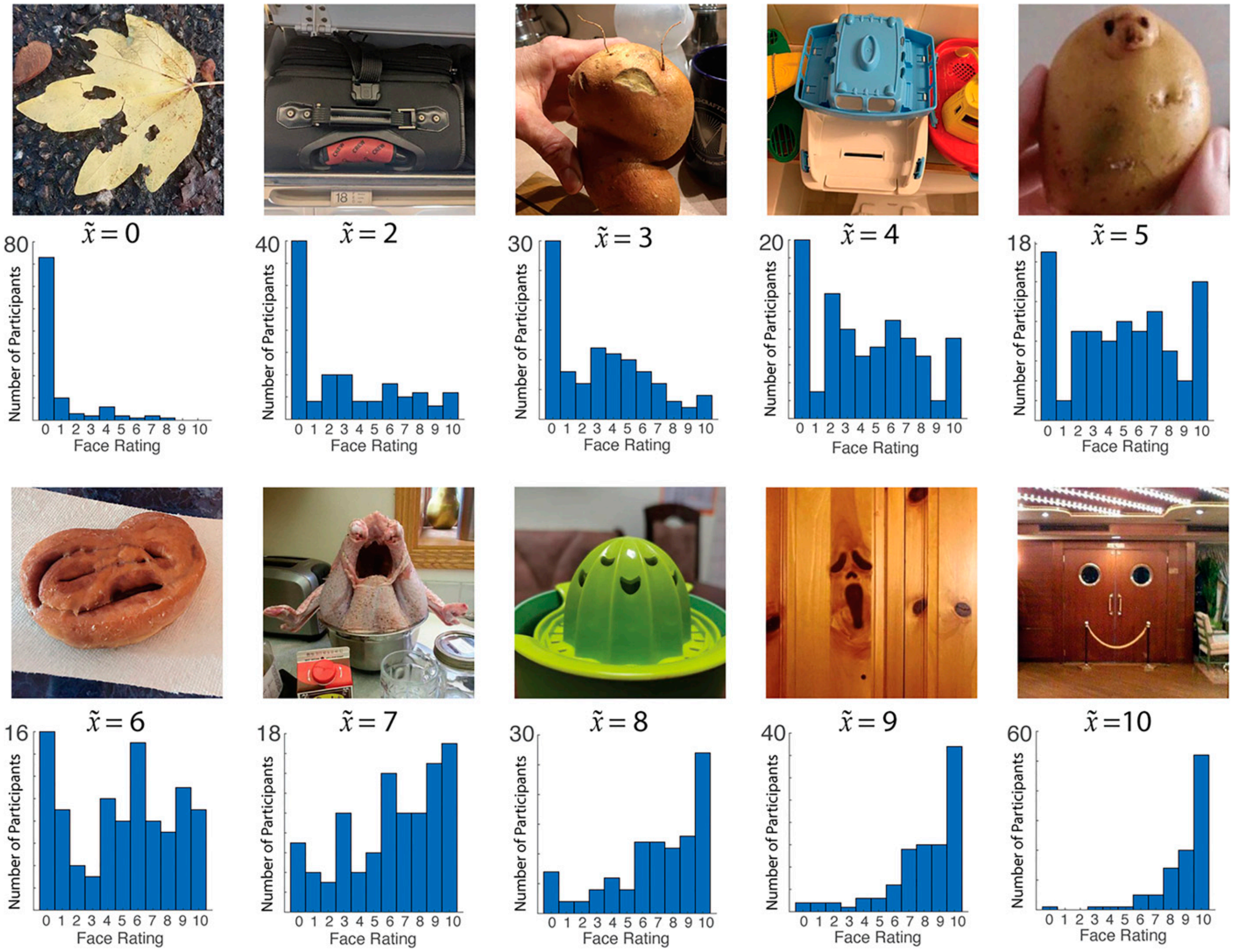

Fig. 1. Distribution of face ratings for 10 representative illusory face images from the set of 256 images used in Exp. 1a. Each image [$n_{(images)} = 256$] received 100 ratings [total $n_{(participants)} = 800$] in Exp. 1a. Below each image is the median ($\tilde{x}$) face-rating score, and a frequency plot of the distribution of face ratings for each image. Note: the scale of the *y* axis is different across frequency plots.

be made, as attributes such as gender and age are unspecified for illusory faces: there is no ground truth. This circumvents the concern that any observable biases are due to choices made in stimulus manipulation (10, 11), and instead any biases observed in the characteristics perceived for these faces are likely to be reflective of the underlying tuning of the face-processing system.

In a series of large-scale behavioral experiments (total n = 3,815) we show that illusory faces in objects are perceived to have a distinct emotional expression, age, and gender.* Furthermore, we discovered a clear bias to perceive illusory faces as male rather than female, at a ratio of ∼4:1. This male bias for pareidolia is highly robust across images and people (Exps. 1a, 3a, and 3b), and cannot be explained by the corresponding object identity (Exps. 1b and 4), object label (Exp. 1b), color (Exps. 2 and 3), or object image content (Exp. 4) of the illusory face images. In contrast, using the same paradigm, we find that human face morphs created from an equal contribution of male and female faces are more likely to be perceived as female than male, although the female bias is smaller in magnitude than the male bias observed for pareidolia (Exp. 5). Together, these results demonstrate that gender evaluation is inextricably linked to face detection, and reveal that these mechanisms are engaged not only by human faces, but also by examples in which the minimal amount of visual information required for face detection occurs. It is important to emphasize that no assignment of gender is necessary for illusory faces as they do not have a biological sex. The existence of a compelling and biased categorization of gender for illusory faces is suggestive of a broadly tuned face evaluation system in which the features that are sufficient for face detection are not generally sufficient for the perception of female.

*We use the term "perceived gender" to refer to the perception of an illusory face as male or female, since illusory faces do not have a biological sex. In contrast, we use the term "perceived sex" to refer to previous research using human faces, in order to distinguish perception of the biological sex from the broader concept of gender. Although we focus on male and female percepts for the purposes of this study, we acknowledge that gender is nonbinary and designed the tasks in all experiments such that a binary gender response (i.e., male or female) was not required.

Results

Illusory Faces are Perceived to Have a Specific Emotional Expression, Age, and Gender. We investigated whether illusory faces are perceived to have a distinct emotional expression, age, and gender using the online crowdsourcing platform Amazon Mechanical Turk (*Materials and Methods*). We collected 256 unique photographs of illusory faces in a diverse set of different natural and man-made objects, such as potatoes, suitcases, and pastries (Fig. 1), sourced from the internet and our personal collection. First, we confirmed that illusory faces were perceived in the images we selected [Exp. 1a; $n_{(total\ participants)} = 800$, $n_{(responses\ per\ image)} = 100$]. For each image, we calculated the median score for how easily participants could see a face as rated on an 11-point scale. The histogram of the median face ratings

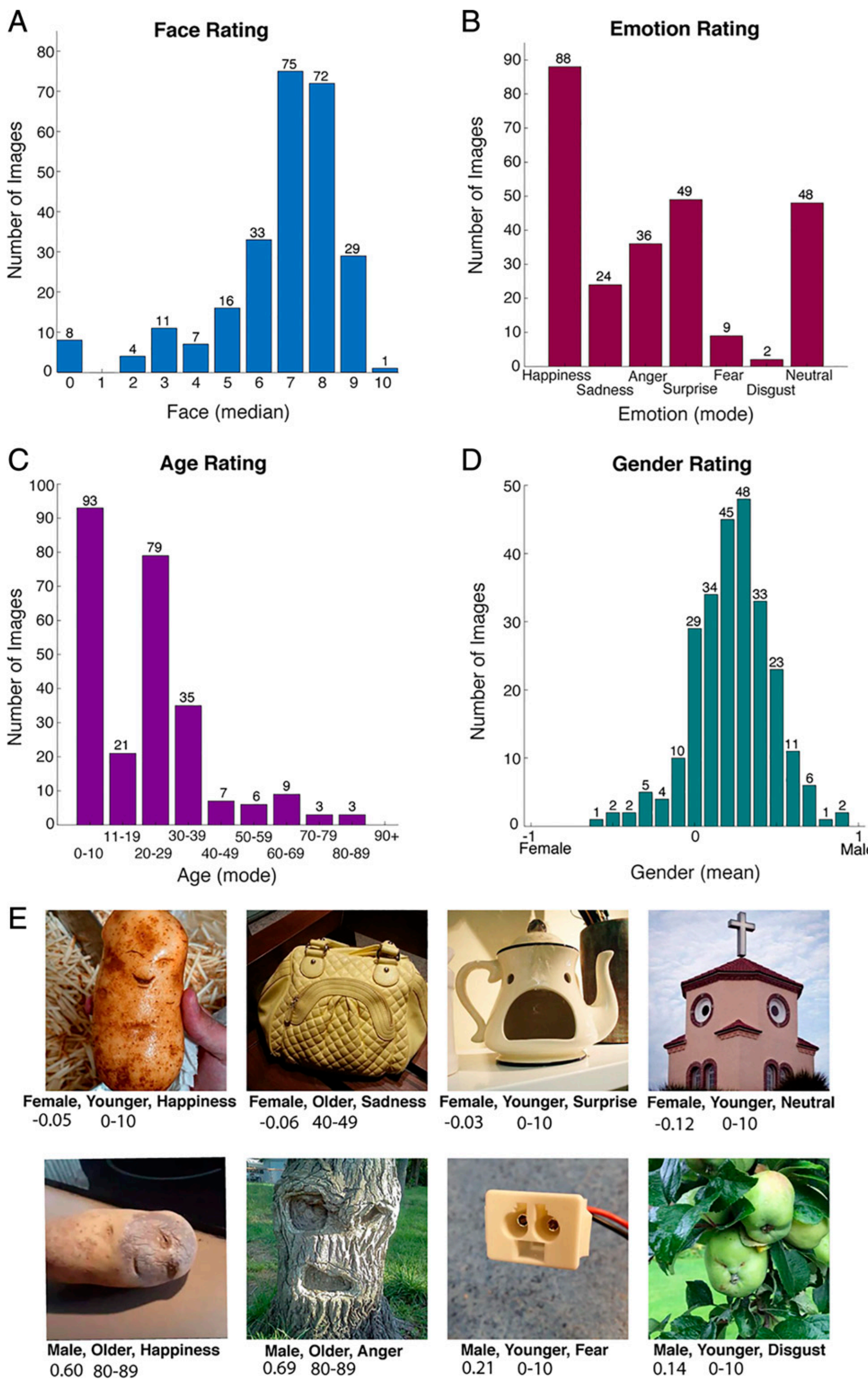

Fig. 2. Illusory faces are perceived to have an emotional expression, age, and gender (Exp. 1a). (*A*) Frequency histogram of the median face rating for each of the 256 illusory face images on an 11-point scale from 0 "cannot see a face" to 10 "easily see a face" [$n_{(total\ participants)} = 800$, $n_{(responses\ per\ image)} = 100$]. (*B*) Frequency histogram of the modal ratings of emotional expression for each illusory face image ($n = 100$ ratings per image) selected from one of seven options. (*C*) Frequency histogram of the modal age rating for each illusory face image selected from 10 options binned in decades from 0 to 90+ y of age [$n_{(responses\ per\ image)} = 100$]. (*D*) Frequency histogram of the mean gender rating for each image on a scale from −1 (female) to +1 (male). (*E*) Eight illusory face images from the total set of 256 images which illustrate images from different gender, age, and emotion categories. For each image, their mean gender rating, modal emotional expression, and modal age rating are listed below.

https://doi.org/10.1073/pnas.2117413119

for each image (Fig. 2*A*) is negatively skewed toward higher ratings, demonstrating that illusory faces were clearly perceived in the majority of the examples we selected [$\tilde{x} = 7$, SD = 1.98, $n_{(images)} = 256$].

Next, we calculated the modal response for the perceived emotional expression and age of each illusory face (Fig. 2 *B* and *C*). A broad range of emotions were perceived across the different illusory faces (Fig. 2*B*): happiness (34%), surprise (19%), anger (14%), sadness (9%), fear (4%), and disgust (1%). Only 19% of illusory faces were rated as having a "neutral" expression. The modal ratings for age were positively skewed toward younger ages (Fig. 2*C*); 75% of the illusory faces had a modal rating under 30 y old. These results show that the majority of illusory faces were perceived as younger rather than older, although there were notably fewer illusory faces perceived as teenage faces (8%, 11 to 19 y of age) compared to the number perceived as either child-like (36%, 0 to 10 y of age) or young adult (31%, 20 to 29 y of age) faces. An own-age bias has been reported for human faces (10, 13), but we found no evidence of a relationship between worker age and the perceived age of illusory faces (*SI Appendix*, Fig. S1).

Finally, we examined whether the illusory faces were perceived to have a gender. Participants rated the gender of each illusory face as "male" (coded as a 1), "female" (−1), or "neutral" (0). The negatively skewed distribution of the mean ratings for each image revealed a clear bias to perceive more of the illusory faces as male rather than female or neutral [mean = 0.28, SD = 0.24, $n_{(images)} = 256$] (Fig. 2*D*). The magnitude of this gender difference was substantial: 90% of illusory face images had a male mean rating (mean > 0), while only 9% of images had a female mean rating (mean < 0). Furthermore, we confirmed that participants also show an overall male bias by replotting the data as a function of participant instead of image (*SI Appendix*, Fig. S2). Here "unbiased" means an equal number of male and female ratings were given by the participant for their set of images, and any deviation from equal (i.e., even by one image) is counted as a bias. For this set of 256 images, 80% of participants had a male bias, and only 3% exhibited a female bias (*SI Appendix*, Fig. S2*A*).

Collectively, these ratings show that illusory faces have a distinct emotional expression, age, and gender. Importantly, there are clear biases in the characteristics perceived for each of these dimensions. Illusory faces are more likely to be perceived as happy, younger, and male, compared to disgusted, older, and female. These results demonstrate the richness of the perception of illusory faces; beyond simple errors of face detection, these faces in inanimate objects are associated with other characteristics relevant for face processing and social perception (Fig. 2*E*). Our most striking observation is a strong bias to perceive illusory faces as male rather than female, even when a neutral option is available. If this bias is robust and reliable across observers, it has important implications for understanding how the perception of sex is processed in the human brain, particularly given that these stimuli do not have a biological sex. For this reason, we next focused on characterizing the extent of this male bias and critically assessing its reliability. Additionally, to understand the nature of the bias, we consider the relative contribution of multiple perceptual (e.g., color, visual associations) and cognitive (e.g., gender of the rater, semantic associations) factors to the perceived gender of illusory faces, which are collated in *SI Appendix*, Table S1.

A Male Bias for Illusory Faces. To evaluate the robustness of the male bias observed in Exp. 1a, we plotted the distribution of all male and female ratings (i.e., excluding neutral responses) made by all participants as a function of image (Fig. 3*A*). Overall, there were significantly more male (81.4%) than female (18.6%) gender ratings [$z = 12.90$, $P = 2.24 \times 10^{-38}$, $n_{(images)} = 256$, one-tailed sign test]. Thus, the male bias observed for the mean rating per image (Fig. 2*D*) also persisted at the level of individual gender ratings made across all images (Fig. 3*A*). Not only were fewer illusory faces rated as female, but those that were, scored relatively lower on the female end of the scale than the equivalent male score for illusory faces rated as male (Fig. 3*B*). A split half analysis revealed very high consistency in the gender ratings attributed to illusory faces across separate subgroups of participants [$r_s(256) = 0.85$, all $P < 0.0001$] (Fig. 3*C*). We found no evidence of a difference in the perceived gender of illusory faces based on the self-identified gender of the rater, as has been reported for human faces (25). Male and female raters gave highly correlated gender ratings for a given illusory face [$r(256) = 0.87$, $P = 5.03 \times 10^{-78}$] (Fig. 3*D*) and had a similar distribution of ratings (Fig. 3*E*). Similarly, male and female workers also gave similar ratings for perceived age and emotional expression (*SI Appendix*, Fig. S3). Overall, these results show that the bias to rate illusory faces as male more often than female is consistent across images, participants, and the gender of the rater.

Finally, we examined the relationship between perceived gender and perceived emotional expression (*SI Appendix*, Fig. S4). Illusory faces perceived as happy were significantly less likely to be perceived as male [$r(256) = -0.33$, $P < 0.001$, $R^2 = 0.11$], and illusory faces perceived to portray anger [$r(256) = 0.34$, $P < 0.001$, $R^2 = 0.12$] or disgust [$r(256) = -0.19$, $P = 0.003$, $R^2 = 0.04$] were significantly more likely to be perceived as male. There was no correlation between ratings of gender and ratings of sadness, surprise, fear, or a neutral expression (all $P > 0.05$). Consistent with these results, associations between gender and emotion have previously been reported for human faces, particularly for angry-male and happy-female (26, 27). Importantly, although we observe a similar association for illusory faces, the effects only explain 4 to 12% in the variability of gender ratings for illusory faces. Furthermore, the direction of causality cannot be inferred from the correlation; for example, it is unclear whether an illusory face is more likely to appear angry because it is male, or more likely to appear male because it is angry.

The Male Bias Is Not Explained by Semantic Object–Gender Associations. To evaluate whether the attribution of gender for illusory faces may have been influenced by semantic associations with the type of object or other visual properties of the images, we conducted a series of follow-up control experiments. First, to assess whether there was any bias in the ratings for gender as a function of the types of objects present in the illusory face images, we generated a list of text labels corresponding to the objects in the 256 illusory face images (*Materials and Methods*). This produced a total of 163 unique object names describing the objects in the set of illusory faces (e.g., "potato") (Fig. 4*A*). The mean gender ratings for object names [$n_{(total\ participants)} = 800$; $n_{(per\ name)} = 100$] are broadly distributed across stimuli and do not show an overall bias for either gender [mean = 0.05, SD = 0.43, $n_{(names)} = 163$] (Fig. 4*B*). Similarly, there was no significant difference in the number of male vs. female responses given to the object names [$z = 1.58$ $P = 0.06$, $n_{(names)} = 163$, one-tailed sign test] (Fig. 4*C*). Furthermore, only 8.5% of the variance in the gender ratings of illusory face images was explained by the gender rating given to the name of the object in the image [$r^2(226) = 0.085$, $P = 7.89 \times 10^{-6}$] (Fig. 4*D*). This contrasts, for example, with the much larger 74.82% of the variance in the gender ratings of illusory face images, which is explained by ratings given by workers of the opposite sex in Exp. 1a (Fig. 3*D*) and the 89.02% in variance explained by gender ratings for color versus grayscale versions of the same image in Exps. 3a and 3b (Fig. 4*D*). These results clearly demonstrate that any preexisting gender associations for the specific objects in the illusory face

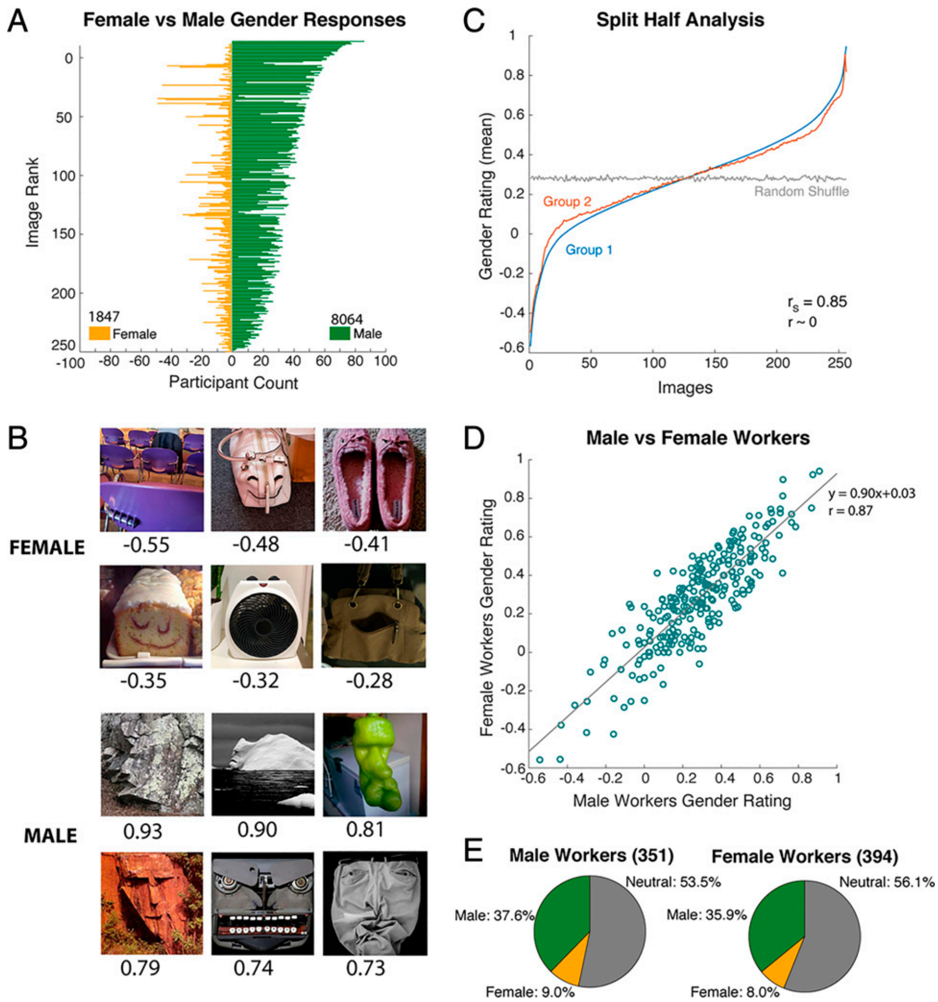

Fig. 3. Bias to perceive illusory faces as male rather than female (Exp. 1a). (*A*) The distribution of female (yellow) and male (green) responses for each image (rows). The order of images (rows) is sorted by the sum of all male and female ratings for each image (image rank). (*B*) The six highest-rated female and male illusory face images. The mean gender rating is below each image [$n_{(responses\ per\ image)} = 100$]. (*C*) Split half consistency analysis of the gender rating scores across participants. The blue line (group 1) shows the scores of a random half of the participants across 1,000 iterations (without replacement), and the red line (group 2) shows the scores of the remaining half of participants, sorted in the same order. The mean rank-order correlation between group 1 and group 2 across all iterations is $r_s(256) = 0.85$, all $P < 0.0001$. For comparison, the gray line (random shuffle) shows the remaining half of the participants when sorted randomly (r ~0). (*D*) Ratings of gender were highly correlated between male and female workers [$r(256) = 0.87$, $P = 5.03 \times 10^{-78}$]. (*E*) Similar distributions of gender ratings in the responses from workers [$n_{(total)} = 800$] who self-identified as male or female. Data for workers who selected "other" ($n = 5$) or who did not select a gender ($n = 50$) are not shown.

images are not sufficient to explain the substantial male bias we observed for illusory faces. Furthermore, we confirmed this pattern of results held for individual participants (*SI Appendix*, Fig. S2*D*).

Visual Object–Gender Associations Do Not Explain the Male Bias. Although we ruled out a preexisting semantic association between object identity and gender underlying the male bias for illusory faces, there may be other visual features of the objects in the images (e.g., color, shape) that have an association with a particular gender. In order to test whether any gender bias exists for visual objects even in the absence of an illusory face, we built and validated a matched set of object images (*Materials and Methods*), which were photographs of the same type of object but which had no perceived face (Fig. 5*A*). The majority of the matched object images were rated as very similar to their corresponding illusory face images on an 11-point scale in an independent stimulus validation experiment (Fig. 5 *B* and *C*) [mean = 7.58, SD = 1.08, $n_{(image\ pairs)} = 227$]. Using the validation data, we selected the 200 highest scoring matches and confirmed in separate experiments (Exps. 2a and 2b) that illusory faces were not perceived in either color [$\tilde{x} = 0$, SD = 0.58, $n_{(images)} = 200$, $n_{(participants)} = 1{,}000$] or grayscale [$\tilde{x} = 0$, SD = 0.48, $n_{(images)} = 200$, $n(_{participants)} = 1{,}000$] versions of the matched images.

Next, we examined the potential role of color by conducting a set of paired experiments with both color and grayscale versions of the selected 200 illusory faces (Fig. 5*A*). The distribution of face scores (Fig. 6 *A* and *B*) confirmed that participants perceived illusory faces in both the color [$\tilde{x} = 7.0$, SD = 1.65, $n_{(images)} = 200$, $n_{(participants)} = 1{,}000$, $n_{(per\ image)} = 100$] and grayscale [$\tilde{x} = 7.0$, SD = 1.65, $n_{(images)} = 200$, $n_{(participants)} = 1{,}000$, $n_{(per\ image)} = 100$] image sets. We took the top 160 illusory face images as calculated based on the face ratings for the color images (*Materials and Methods*) and collected gender ratings. Again, we observed a clear male bias in the mean gender ratings for illusory faces in both the color (Fig. 6*C*) [mean = 0.19, SD = 0.26, $n_{(images)} = 160$, $n_{(per\ image)} = 100$] and grayscale (Fig. 6*D*) [mean = 0.20, SD = 0.27, $n_{(images)} = 160$, $n_{(per\ image)} = 100$] versions. There were also significantly more male than female responses made across all images for both color [$z = 7.04$,

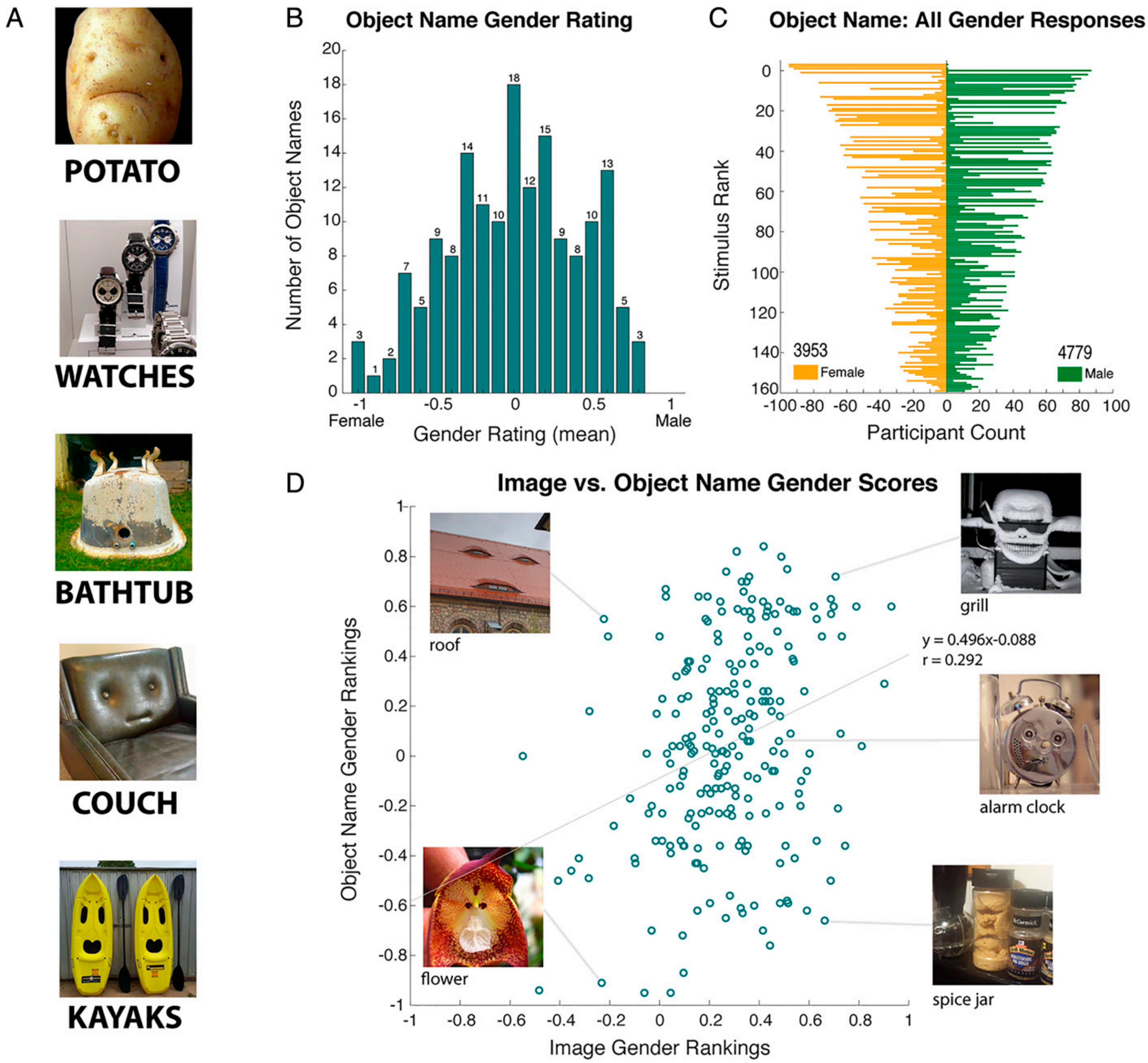

Fig. 4. No bias in the gender attributions for named objects (Exp. 1b). (*A*) Five examples of the 163 object labels used as the text stimuli in Exp. 1b, shown under the corresponding illusory face image from Exp. 1a. No object images were presented in Exp. 1b. (*B*) Frequency histogram of the mean gender ratings for the 163 object labels (n = 100 ratings per label), averaged across participants and coded as female = −1, neutral = 0, male = 1 [mean = 0.05, SD = 0.43, $n_{(names)}$ = 163]. (*C*) The distribution of female (yellow) and male (green) responses across all participants as a function of object label (rows). The order of stimuli in the plot (rows) are sorted by the sum of all male (green) and female (yellow) ratings for each object label (stimulus rank), so that stimuli with the lowest number of neutral ratings are shown at the top. (*D*) Correlation between the gender ratings for the illusory face images (Exp. 1a) and their corresponding object names (Exp. 1b); 8.5% of the variance in the mean gender ratings of the images is explained by the object name mean gender rating [$r^2(226) = 0.085$, $P = 7.89 \times 10^{-6}$].

$P = 9.89 \times 10^{-13}$, $n_{(images)} = 160$, one-tailed sign test] and grayscale [$z = 7.56$, $P = 2.05 \times 10^{-14}$, $n_{(images)} = 160$, one-tailed sign test] illusory faces (Fig. 6 *E* and *F*). These results replicate the male bias in two independent datasets and demonstrate that color is not necessary for the bias.

Finally, we tested whether any gender bias exists for the matched object images that do not contain an illusory face. Using the set of matched object images (Fig. 5*A*) corresponding to the top 160 illusory faces, we asked participants if they associated each nonface object image with male, female, or neutral (*Materials and Methods*). In this case, the distribution of male and female gender ratings for the matching object images was approximately equal [mean = 0.01, SD = 0.38, $n_{(images)} = 160$, $n_{(per\ image)} = 100$] (Fig. 7*A*) and there was no significant difference in the number of male vs. female ratings [$z = 0.24$, $P = 0.41$, $n_{(images)} = 160$, one-tailed sign test] (Fig. 7*B*). We correlated the gender ratings for the illusory face images with those for their corresponding matched object images (Fig. 7*C*), and found that only 10.84% of the variance in the gender ratings of illusory face images was explained by the ratings for the corresponding matched object images [$r^2(160) = 10.84\%$, $P = 2.12 \times 10^{-5}$]. In contrast, there was a very strong

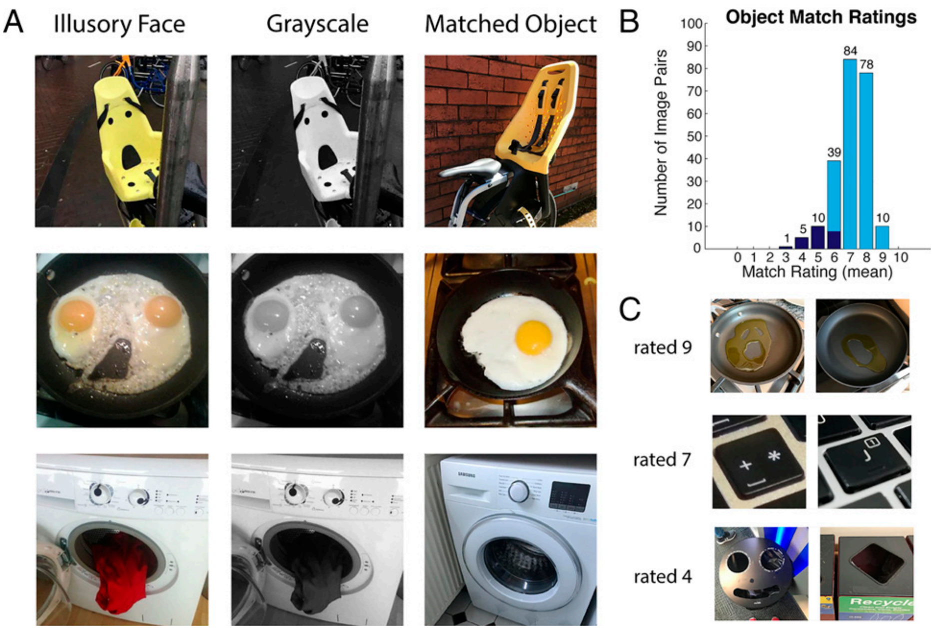

Fig. 5. Example visual stimuli and matched image validation data (Exps. 2 to 4). (*A*) Example stimuli showing illusory faces in color (Exps. 2a and 3a), grayscale (Exps. 2b and 3b), and their corresponding matched object images which do not elicit the perception of an illusory face (Exp. 4). (*B*) Distribution of the mean ratings (n = 260) for the similarity of illusory face images to their matched nonface object images, on an 11-point scale from 0 (different object) to 10 (same object). Only the top 200 matched illusory faces (light blue) were used in Exps. 2a and 2b; images at the bottom of the distribution (dark blue) were excluded (*Materials and Methods*). (*C*) Example image pairs with the mean rating of their similarity. Note that the image pair rated 4 was below the cutoff we adopted for further experiments.

correlation between gender ratings for color versus greyscale versions of the illusory faces (Fig. 7*D*), with 89.02% of the variance in gender ratings for grayscale illusory faces explained by the ratings for their color versions [$r^2(160) = 89.02\%$ $P = 1.08 \times 10^{-77}$]. Together, these data demonstrate that the bias to perceive illusory faces as male is robust even in the absence of color, and cannot be explained by preexisting gender associations with the underlying objects in which the illusory faces are perceived. Instead, the tendency to perceive illusory faces as male appears to arise predominantly from the perception of the face itself.

To rule out the possibility that participants respond "male" for faces more often under all conditions of visual uncertainty, we conducted an analogous experiment with the same paradigm using gender-ambiguous human face morphs to create visual uncertainty (Exp. 5) (*SI Appendix*, Fig. S5). We found there were significantly more female (57.7%) than male (42.3%) gender ratings [$z = -5.58$, $P = 2.35 \times 10^{-8}$, $n_{(images)} = 256$, two-tailed sign test] across all images and participants. This female bias is much smaller in magnitude than the male bias we observed for illusory faces (in Exp. 1a, 81.4% of all nonneutral ratings were male, 18.6% were female), and supports a genuine bias to perceive illusory faces as male more often than female.

Computational Modeling Reveals Visual Features Do Not Explain the Male Bias. Exps. 1 to 4 establish that the male bias for illusory faces is robust and reliable across participants and images. Furthermore, while the gender ratings for object names (Exp. 1b) and matched objects (Exp. 4) suggests that gender attributions for illusory faces cannot be accounted for by visual or semantic associations with the object that the face is perceived in, we were curious whether other visual image features may contribute to the perception of gender. One candidate is the curvature and rectilinearity content of the images, since for example prototypical masculine faces are typically associated with more angular features compared to prototypical feminine faces. For each of the 160 pareidolia images in Exp. 3, we computed separate indices of the curvature and rectilinearity content (28, 29), using the algorithm of Yetter et al. (28). Using multiple linear regression, we found that rectilinearity and curvature indices predicted 9% of the variance in gender ratings for the pareidolia images [$F_{(2,\ 155)} = 7.640$, $P < 0.001$, $R^2 = 0.09$]. When we further added the gender scores for the object names and matched objects as extra predictors in the model, it explained 17.7% of the variance in gender ratings for the corresponding pareidolia images [$F_{(4,\ 153)} = 8.218$, $P < 0.001$, $R^2 = 0.177$, $R^2_{\text{change}} = 0.087$].[†] Both rectilinear indices ($B = 0.391$, $P = 0.008$) and gender ratings for the matched objects ($B = 0.210$, $P = 0.004$) were significant predictors of gender ratings for pareidolia, but curvature ($B = 0.034$, $P = 0.642$) and gender ratings for the object names ($B = -0.003$, $P = 0.962$) were not.

To complement this analysis, we applied four additional classic and state-of-the-art computational visual models (Fig. 8 and *SI Appendix*) to characterize the visual aspects of the stimuli in more detail beyond their curvature and rectilinearity. These include the GBVS and Itti-Koch models of visual saliency (30), the GIST visual feature model (31), and the VGG-19 convolutional neural network (32). Since these models output complex representations of the stimuli that are not reducible to a single index, we used representational similarity analysis to compare the gender ratings and model representations for all pairs of stimuli (Fig. 8). The representational dissimilarity matrices (RDMs) for gender ratings of color (Fig. 8*A*) and grayscale (Fig. 8*B*) pareidolia images are highly similar. However, visual inspection of the model RDMs suggests that none of the model representations (Fig. 8 *C* and *D*) adequately captured the patterns in gender scores across stimuli (Fig. 8*A*). This was confirmed with a multiple regression analysis on the dissimilarity scores for each stimulus pair. A model with the same four

[†]Two of the 160 pareidolia images did not have corresponding object names in Exp. 1b and thus had missing data for this predictor.

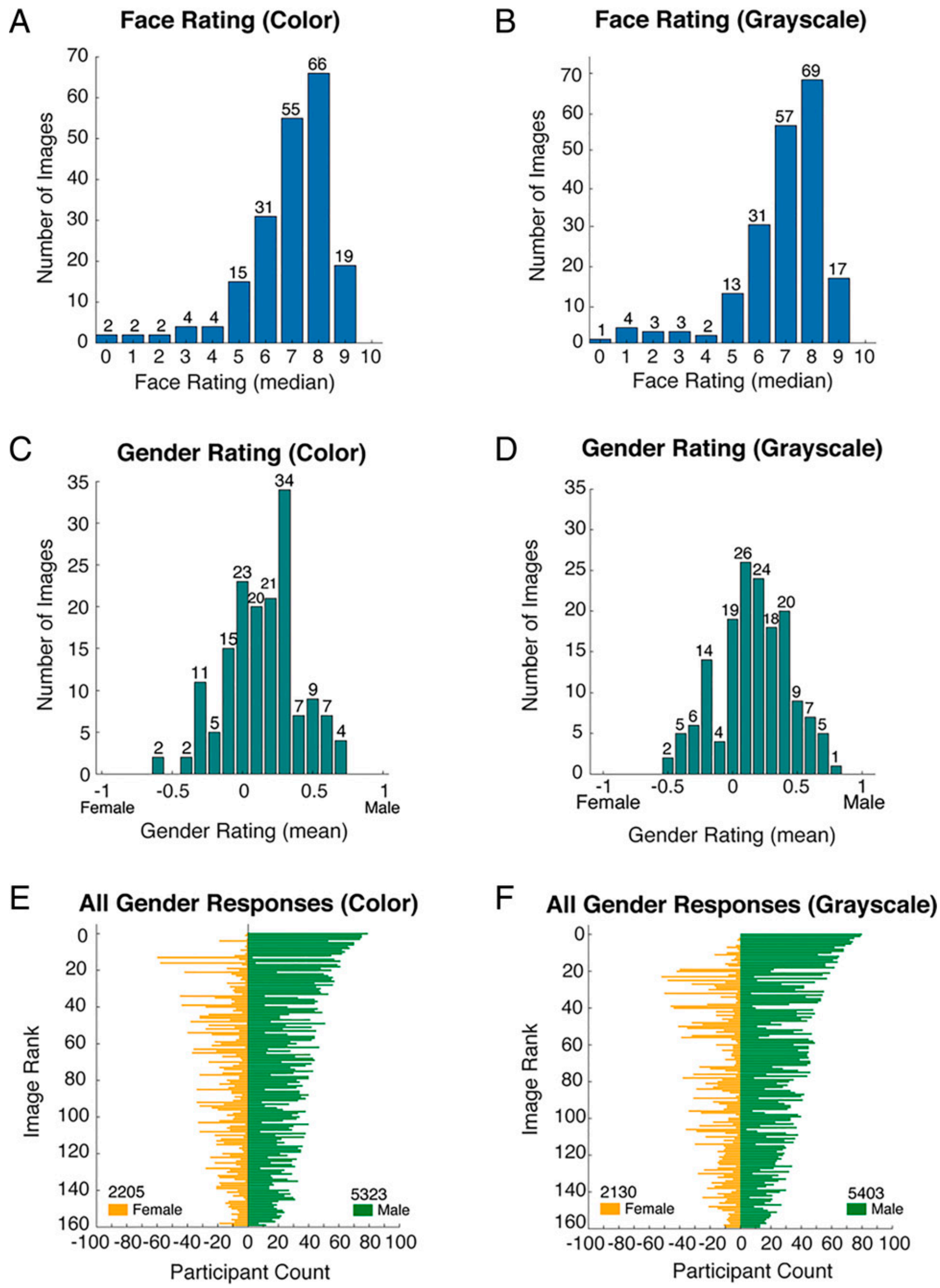

Fig. 6. A male bias for illusory faces in both color and grayscale (Exps. 2 to 3). (*A*) Frequency histogram of the median face ratings [$\tilde{x}$ = 7.0, *SD* = 1.65, $n_{(images)}$ = 160, $n_{(per\ image)}$ = 100) for each illusory face image in color on an 11-point scale from 0 (cannot see a face) to 10 (easily see a face). (*B*) Frequency histogram of the median face ratings [$\tilde{x}$ = 7.0, *SD* = 1.65, $n_{(images)}$ = 160, $n_{(per\ image)}$ = 100] for each illusory face image in grayscale. (*C*) Frequency histogram of the mean gender rating scores for each illusory face image in color, averaged across participants and coded as female = −1, neutral = 0, male = 1 [mean = 0.19, SD = 0.26, $n_{(images)}$ = 160, $n_{(per\ image)}$ = 100]. (*D*) Frequency histogram of the mean gender rating scores for each illusory face image in grayscale [mean = 0.20, SD = 0.27, $n_{(images)}$ = 160, $n_{(per\ image)}$ = 100], as in *C*. (*E*) Distribution of female (yellow) and male (green) responses across all participants as a function of illusory face image (rows) in color. The order of stimuli (rows) is sorted by the sum of all male and female ratings for each image (image rank), so that stimuli with the lowest number of neutral ratings are shown at the top. (*F*) Distribution of female and male responses for the grayscale versions of the images, as in *E*.

predictors as used in the previous regression analysis (gender ratings for object names and matched objects, and the curvature and rectilinear indices) only predicted 1.5% of the variance in the similarity between pairs of stimuli in their gender scores [$F_{(4, 12,398)} = 45.681$, $P < 0.001$, $R^2 = 0.015$]. When we added all 26 predictors from the visual feature models (including 19 layers of VGG-19), only 5.6% of the variance in the dissimilarity of gender scores between stimulus pairs was explained [$F_{(26, 12,376)} = 28.287$, $P < 0.001$, $R^2 = 0.056$, $R^2_{\text{change}} = 0.042$]. Considered together, these results demonstrate that while visual features may explain some of the variance in gender scores, they are not sufficient to explain the male bias for illusory faces.

Discussion

Illusory faces are spontaneous errors of face detection in which the perceived facial features are defined by highly variable visual properties. We conducted a series of large-scale

https://doi.org/10.1073/pnas.2117413119

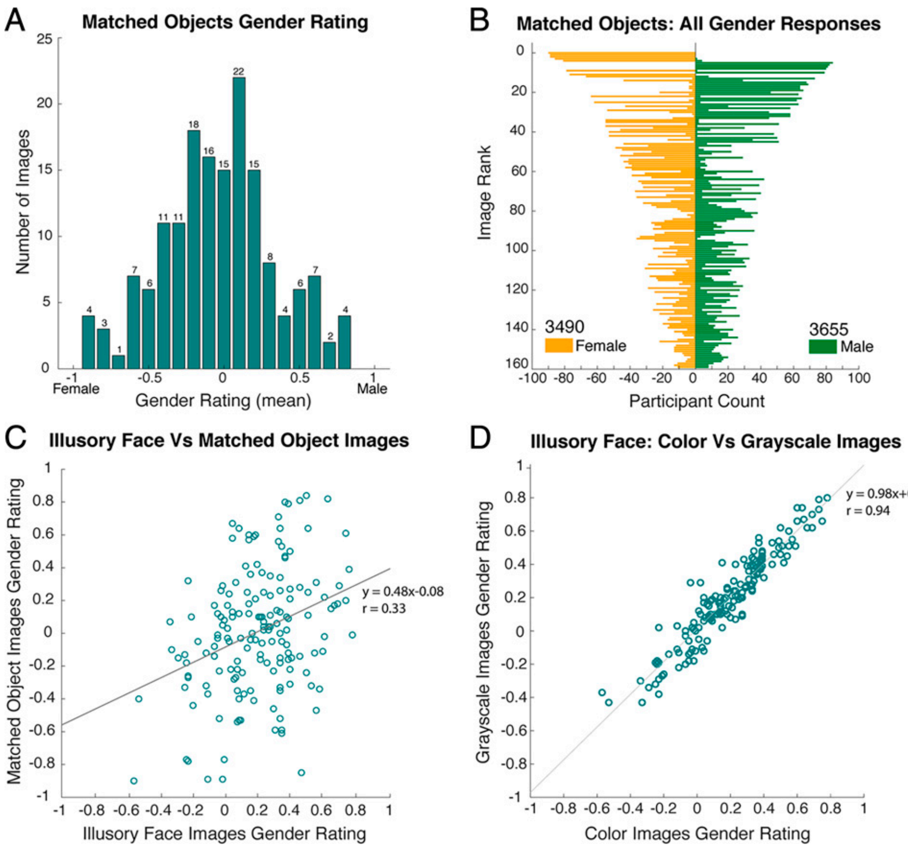

Fig. 7. No gender bias for object images that do not contain an illusory face (Exp. 4). (*A*) Frequency histogram of the mean gender ratings for each matched (nonface) object image, averaged across participants and coded as female = −1, neutral = 0, male = 1 [mean = 0.01, SD = 0.38, $n_{(images)}$ = 160, $n_{(per\ image)}$ = 100]. (*B*) Distribution of female (yellow) and male (green) responses across all participants for each matched object image (rows). The order of stimuli (rows) is sorted by the sum of all male and female ratings for each image (image rank), so that stimuli with the lowest number of neutral ratings are shown at the top. (*C*) Scatterplot showing the correlation between gender ratings for illusory faces (Exp. 3a) and their corresponding matched object image (Exp. 4) [$r^2(160) = 10.84\%$, $P = 2.12 \times 10^{-5}$]. The solid line is the result of the best-fitting linear regression of the form $y = mx + b$. (*D*) Scatterplot showing the strong correlation between gender ratings for the same illusory faces when shown in color (Exp. 3a) and in grayscale (Exp. 3b) [$r^2(160) = 89.02\%$ $P = 1.08 \times 10^{-77}$] as in *C*.

behavioral experiments and found that illusory faces in inanimate objects are readily perceived to have a specific emotional expression, age, and gender. This suggests that illusory faces engage brain mechanisms involved in higher-order aspects of face evaluation beyond the simple detection of a face, and is consistent with recent neuroimaging results that have revealed shared neural mechanisms for human faces and illusory faces (14, 15, 18). Most strikingly, we observed a strong bias to perceive illusory faces in objects as male rather than female, which was replicated across three separate experiments. This bias occurred both at the level of ratings for individual illusory face examples, and for the ratings of individual participants.

In a series of follow-up experiments and using computational modeling, we considered the relative contribution of multiple perceptual and cognitive factors to the male bias (*SI Appendix*, Table S1). In terms of perceptual factors, we were able to rule out alternative explanations based upon: 1) the color of the images, 2) visual associations with the object or image content containing the illusory face, 3) the curvature or rectilinearity of the images, and 4) other complex visual features of the images as captured by computational modeling. Similarly, in terms of cognitive factors, we found that the male bias could not be explained by: 1) the self-identified gender of the participant, 2) semantic associations with the object type or label that contains the illusory face, 3) the task design, 4) a tendency to rate ambiguous stimuli as male in general, or 5) the perceived emotion of the illusory face. Notably, several of these factors did have a modest contribution to explaining some of the variance in gender ratings given to illusory faces, but most of the variance remained unexplained. Together, our data support a robust bias in the perception of gender in illusory faces that is not reducible to other perceptual or cognitive factors.

Based on our results, we speculate that the bias to categorize illusory faces as male more frequently than as female is cognitive rather than perceptual in origin. Even state-of-the-art visual feature models, such as the VGG-19 deep neural network, did not contain a representation of the illusory faces that was strongly predictive of their gender rating, evidence against a strictly perceptual explanation. Instead, we suggest that a cognitive bias to perceive illusory faces as male arises from a broadly tuned face evaluation system in which minimally viable face percepts are more likely to be categorized as male. The origin of the bias—whether from social conditioning or from perceptual factors such as our visual diet of faces during development—remains an open question that is beyond the reach of the current data. Although the exact nature of the cognitive mechanism underlying the bias to perceive illusory faces as male is not possible to determine here, there are several possibilities. One possibility is that it stems from a conceptual or linguistic origin, in that male is the default gender in social communication. By this account, the perception of an illusory face in an object invokes the concept of "person," which in turn invokes the concept of "male," unless additional information suggests otherwise. A related idea is that male is the default

PSYCHOLOGICAL AND COGNITIVE SCIENCES

https://doi.org/10.1073/pnas.2117413119

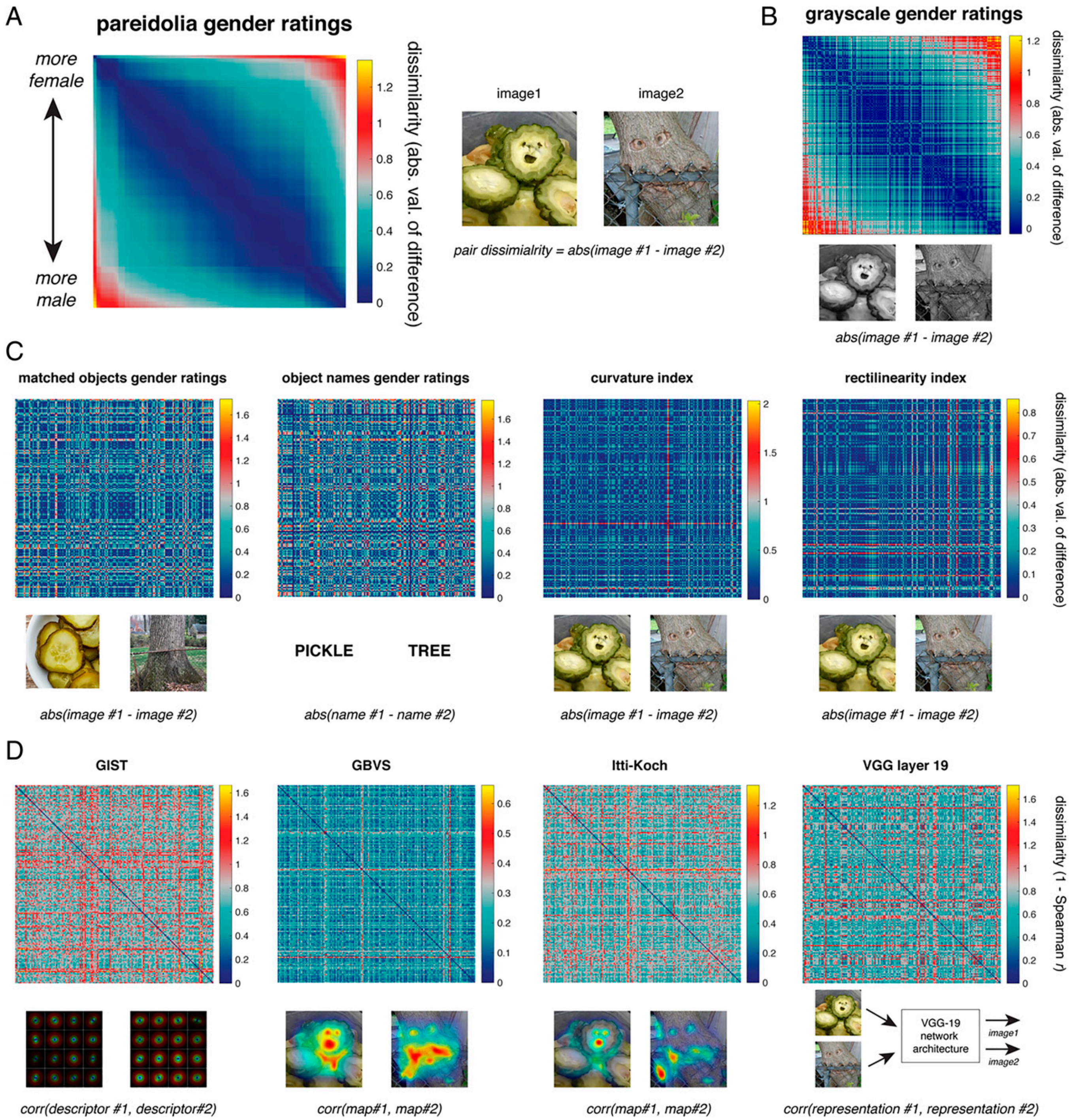

Fig. 8. Computational models of visual features do not explain the male bias for illusory faces. (*A*) We constructed a 160 × 160 RDM based on the gender ratings for the 160 pareidolia stimuli in Exp. 3a. The order of stimuli in all matrices in *A*–*D* is organized according to the mean gender score of the pareidolia images in Exp. 3a, thus a model that explained most of the variance in gender scores for pareidolia would be visually similar to the model in *A*. (*B*) The dissimilarity matrix constructed from independent gender ratings for 160 grayscale pareidolia images in Exp. 3b is similar in structure to the RDM for the Exp. 3a ratings for color images in *A*. (*C*) For comparison, we constructed dissimilarity matrices from the gender scores for the matched objects (Exp. 4) and object names (Exp. 1b) corresponding to the 160 pareidolia images in Exp. 3a, as well as the difference in curvature and rectilinearity scores for each pair of pareidolia images. (*D*) Dissimilarity matrices for the GIST, GBVS, and Itti-Koch visual feature and saliency models were built by calculating 1-Spearman correlation between the model representation of each pair of stimuli. Dissimilarity matrices for the VGG-19 convolutional neural network were constructed by calculating 1- Spearman correlation between the model representation of each pair of stimuli separately for each of the 19 layers of the CNN (only the final fully connected layer is pictured here).

gender for a face, unless other visual details (e.g., eyelashes, long hair, trimmed eyebrows) suggest differently. In this case, given that illusory faces provide only the minimal visual information required for the perception of a face, they are also likely to be perceived as male by default. Regardless of the origin of the male bias for face pareidolia, its existence raises interesting questions about how social norms may interact with visual perception.

Aside from the existence of a bias, it is important to note that illusory faces do not have a biologically specified sex, and consequently there is no a priori reason why they should be perceived to have a gender. Importantly, our experimental paradigm did not require participants to select a gender for the illusory faces, and around half of the responses were neutral. Thus, the bias to see examples of face pareidolia as male rather than female occurred without forcing a binary decision to be made. In contrast to illusory faces, human faces are sexually dimorphic (33), and people are impressively accurate at recognizing the biological sex of an individual by facial cues alone, even when other external cues, such as hair and clothing, are absent (21, 22). Adults can identify a face as male or female in a few hundred milliseconds (5), from presentations as brief as 75 ms (6) and under considerable attentional load (34, 35), consistent with neural specialization appropriate for identifying a characteristic with such biological and social relevance. Despite our skill at determining biological sex from human facial characteristics, there is evidence from both behavioral (10, 11, 19) and neuroimaging (36) studies that male and female human faces may not be processed as equal, discrete categories by the brain. It is possible that the presence of only minimal features (e.g., eyes, mouth) leads to the perception of the face as male for nonhuman faces in which the sexually dimorphic characteristics (22, 33, 37–41) that give clues to sex in human faces are absent.

A reliable bias to perceive illusory faces as male rather than female even though they do not have a biological sex is consistent with the idea that sex perception is not strictly categorical (10, 11, 19). Other known asymmetries, such as faster reaction times for classifying male than female faces by sex (39) and a greater neural response to female than male faces embedded in a stream of images of the opposite sex (36), illustrate differences between the sexes in both perception and neural processing. Similarly, there is evidence that threatening faces emerge faster in consciousness (42), suggesting a prioritization of processing for dimensions that may correlate with masculinity. In a continuous flash-suppression paradigm, faces that evoked a high level of power/dominance broke through to conscious perception the fastest (42). Recent magnetencephalography evidence suggests that sex may be processed before other attributes, such as identity (43), also consistent with a prioritization of processing. Interestingly, infants as young as 3 to 4 mo of age show a bias in preferential looking directed toward female faces rather than male, unless their primary caregiver is male, in which case the bias is reversed toward male faces (44). This initial bias in looking behavior could set the groundwork for differential processing of male and female faces in the developing brain. However, the ability to classify faces by sex continues to develop with age, and adults are better than children at classifying the gender of children's faces (21), suggesting there is continued development of the ability to perceive sex from faces for many years after birth. Characterizing this developmental trajectory in more detail will be critical in understanding the role of development in the perception of sex, and more broadly, the neural development of mechanisms involved in face evaluation.

An advantage of using illusory faces to examine biases in gender categorization is that no stimulus manipulation is required to create gender ambiguity, as is necessary with human faces that have an inherent biological sex. Previous studies using manipulated human faces have reported a tendency to perceive a human face as male more often than female when there is visual ambiguity (10, 11, 19, 22, 23, 39), but it has been difficult to rule out alternative explanations for the bias in these paradigms because of the challenge of creating human faces that are ambiguous along the dimension of sex. Examples of manipulations that produce a bias include photographs with external cues to biological sex such as hair removed (21, 22), silhouettes of faces in profile (23), faces with visual noise added (24), and synthetically generated or morphed male and female human faces (10, 11, 19). For example, one study using morphed human faces found that morphs with a greater contribution (> 50%) from the underlying male than female face are clearly rated as male, but the reverse is not true for morphs with a greater contribution from the female face until the morph is ~70% female (19). In contrast, when we used ambiguous morphs of human faces (i.e., 50% male, 50% female) in our paradigm, we discovered a bias to perceive the morphed faces as female, although the magnitude of the female bias for morphed human faces was less than the male bias for pareidolia. This apparent contradiction with earlier studies is likely explained by differences in the paradigm and morphing methods between studies (10, 11, 19). Specifically, inspection of the visual stimuli reveals that our human face morphs have softer edges, compared to the morphs in the other study that have more angular features, a result of differences in the morphing methodology, which may explain the opposing results. Our result showing a reliable male bias for natural errors of face detection, which do not have an inherent biological sex, is thus important in demonstrating that the bias reflects a fundamental characteristic of face perception, and circumvents the issues inherent in manipulating human faces to create ambiguity.

Our data also inform ongoing debates about the processing of different higher-order characteristics of faces, and the relevance of familiarity in face evaluation (11, 43, 45–48). For example, it has been debated whether the sex and identity of faces in processed in parallel or sequentially (5, 48). The results of a recent study, which measured evolving whole-brain face representations over time using magnetoencephalography, suggest that the sex and age of faces is processed before their identity (43). Here we show with illusory faces that neither a coherent identity nor familiarity with a face are necessary for the perception of age, gender, or emotional expression. While familiarity with a face modulates perception of other characteristics—such as sex, age, and race (5, 11, 19, 43, 45, 47)—our results with illusory faces demonstrate that familiarity is not required to engage face-evaluation mechanisms.

Together, our results reveal that male is overwhelmingly the gender perceived in face pareidolia. Illusory faces are readily perceived in a wide variety of natural and man-made objects, by both humans (18) and rhesus macaques (14, 15), and share neural processing mechanisms with real faces (18). Here, we report that socially relevant higher-order characteristics—such as age, sex, and emotional expression—are also perceived in these errors of face detection, suggesting the engagement of more sophisticated mechanisms involved in face evaluation. Overall, these results demonstrate the richness with which face pareidolia taps into the primate face-processing system, and highlights its ability to reveal new insight into face-processing mechanisms (14–16, 18). Our results provide robust evidence for a broadly tuned face evaluation system in which the features that are sufficient for face detection are generally insufficient for a diverse perception of gender. Instead, the perception of a nonhuman face as female seems to require additional features beyond those required for face detection.

Materials and Methods

Participants. All experiments were approved by the NIH Office of Human Subjects Research Protections. A total of 3,815 adults (mean = 39.6 y, SD = 12.1) participated in the experiments via the online crowdsourcing platform Amazon Mechanical Turk. Participants provided consent by acknowledging their participation on a screen before starting the study on Amazon Mechanical Turk, and they were financially compensated for their time. All participants identified as US residents. Of the participants included in the final analysis, 1,670 self-identified as male, 1,676 as female, 18 as other, and 110 gave no

PSYCHOLOGICAL AND COGNITIVE SCIENCES

https://doi.org/10.1073/pnas.2117413119

response to the gender question. Across all experiments a total of 6,236 datasets were collected. In total, 376 datasets (6%) were excluded based on the responses to predetermined catch trials and attention criteria (*SI Appendix*). We analyzed the remaining 5,860 datasets: $n = 800$ datasets each for Exps. 1a and 1b, $n = 260$ for the stimulus validation experiment; $n = 1,000$ each for Exps. 2a and 2b; $n = 400$ each for Exps. 3a, 3b, and 4; and $n = 800$ for Exp. 5.

Stimuli. All visual stimuli used in these experiments are publicly available on the Open Science Framework (https://osf.io). We selected 256 images of illusory faces in objects from various sources, including Google Images, Reddit, and the authors' personal collection. Images were cropped square and resized (400×400 pixels) but no other manipulations were made. All image manipulations (cropping, resizing, and conversion to grayscale) were completed in Photoshop 2020 (v21.2.0) and MATLAB R2020a. All 256 illusory face images were used in Exp. 1a, and 160 of these images were used in Exps. 2 and 3. For Exp. 1b, we generated text labels corresponding to the names of the objects in the illusory face images. The stimuli in Exp. 4 were a validated matched set of object images which were photographs of the same types of objects that were as visually similar as possible to the illusory face images, but which had no face. The stimuli in Exp. 5 were human face morphs that were equal composites of male and female faces (50% female, 50% male). Full details of stimulus selection, generation and validation procedures are described in *SI Appendix*.

Experimental Design and Data Analysis. The online experiments were coded in HTML, CSS, and JavaScript and conducted on the Amazon Mechanical Turk platform. Stimulus presentation was sequential in all experiments; the image or text label stimuli appeared one at a time in a random order for each participant, with the relevant questions listed either to the right (Exp. 1a) or below (all other experiments) the stimulus on each trial. Full details of the experimental design and data analysis are in *SI Appendix*.

Data Availability. The visual stimuli and data for these experiments are publicly available on the Open Science Framework at https://osf.io/f74xh/.

1. N. N. Oosterhof, A. Todorov, The functional basis of face evaluation. *Proc. Natl. Acad. Sci. U.S.A.* **105**, 11087–11092 (2008).
2. A. Todorov, C. P. Said, A. D. Engell, N. N. Oosterhof, Understanding evaluation of faces on social dimensions. *Trends Cogn. Sci.* **12**, 455–460 (2008).
3. R. M. Stolier, J. B. Freeman, Neural pattern similarity reveals the inherent intersection of social categories. *Nat. Neurosci.* **19**, 795–797 (2016).
4. D. M. Burt, D. I. Perrett, Perception of age in adult Caucasian male faces: Computer graphic manipulation of shape and colour information. *Proc. Biol. Sci.* **259**, 137–143 (1995).
5. V. Bruce, H. Ellis, F. Gibling, A. Young, Parallel processing of the sex and familiarity of faces. *Can. J. Psychol.* **41**, 510–520 (1987).
6. A. J. O'Toole, J. Peterson, K. A. Deffenbacher, An 'other-race effect' for categorizing faces by sex. *Perception* **25**, 669–676 (1996).
7. M. Milders, A. Sahraie, S. Logan, Minimum presentation time for masked facial expression discrimination. *Cogn. Emotion* **22**, 63–82 (2008).
8. M. Bar, M. Neta, H. Linz, Very first impressions. *Emotion* **6**, 269–278 (2006).
9. J. Willis, A. Todorov, First impressions: Making up your mind after a 100-ms exposure to a face. *Psychol. Sci.* **17**, 592–598 (2006).
10. T. L. Watson, Y. Otsuka, C. W. G. Clifford, Who are you expecting? Biases in face perception reveal prior expectations for sex and age. *J. Vis.* **16**, 5 (2016).
11. R. Armann, I. Bülthoff, Male and female faces are only perceived categorically when linked to familiar identities—And when in doubt, he is a male. *Vision Res.* **63**, 69–80 (2012).
12. R. Adolphs, D. Tranel, A. R. Damasio, The human amygdala in social judgment. *Nature* **393**, 470–474 (1998).
13. J. Vestlund, L. Langeborg, P. Sörqvist, M. Eriksson, Experts on age estimation. *Scand. J. Psychol.* **50**, 301–307 (2009).
14. J. Taubert, S. G. Wardle, M. Flessert, D. A. Leopold, L. G. Ungerleider, Face pareidolia in the rhesus monkey. *Curr. Biol.* **27**, 2505–2509.e2 (2017).
15. J. Taubert *et al.*, Amygdala lesions eliminate viewing preferences for faces in rhesus monkeys. *Proc. Natl. Acad. Sci. U.S.A.* **115**, 8043–8048 (2018).
16. C. J. Palmer, C. W. G. Clifford, Face pareidolia recruits mechanisms for detecting human social attention. *Psychol. Sci.* **31**, 1001–1012 (2020).
17. R. T. Keys, J. Taubert, S. G. Wardle, A visual search advantage for illusory faces in objects. *Atten. Percept. Psychophys.* **83**, 1942–1953 (2021).
18. S. G. Wardle, J. Taubert, L. Teichmann, C. I. Baker, Rapid and dynamic processing of face pareidolia in the human brain. *Nat. Commun.* **11**, 4518 (2020).
19. I. Bülthoff, F. Newell, Categorical perception of sex occurs in familiar but not unfamiliar faces. *Vis. Cogn.* **11**, 823–855 (2004).
20. M. J. Farah, K. D. Wilson, M. Drain, J. N. Tanaka, What is "special" about face perception? *Psychol. Rev.* **105**, 482–498 (1998).
21. H. A. Wild *et al.*, Recognition and sex categorization of adults' and children's faces: Examining performance in the absence of sex-stereotyped cues. *J. Exp. Child Psychol.* **77**, 269–291 (2000).
22. V. Bruce *et al.*, Sex discrimination: How do we tell the difference between male and female faces? *Perception* **22**, 131–152 (1993).
23. N. Davidenko, Silhouetted face profiles: A new methodology for face perception research. *J. Vis.* **7**, 6 (2007).
24. A. Cellerino, D. Borghetti, F. Sartucci, Sex differences in face gender recognition in humans. *Brain Res. Bull.* **63**, 443–449 (2004).
25. C. Lewin, A. Herlitz, Sex differences in face recognition—Women's faces make the difference. *Brain Cogn.* **50**, 121–128 (2002).
26. D. V. Becker, D. T. Kenrick, S. L. Neuberg, K. C. Blackwell, D. M. Smith, The confounded nature of angry men and happy women. *J. Pers. Soc. Psychol.* **92**, 179–190 (2007).
27. U. Hess, R. B. Adams, K. Grammer, R. E. Kleck, Face gender and emotion expression: Are angry women more like men? *J. Vis.* **9**, 19.1–19.8 (2009).
28. M. Yetter *et al.*, Curvilinear features are important for animate/inanimate categorization in macaques. *J. Vis.* **21**, 3 (2021).
29. X. Yue, I. S. Pourladian, R. B. H. Tootell, L. G. Ungerleider, Curvature-processing network in macaque visual cortex. *Proc. Natl. Acad. Sci. U.S.A.* **111**, E3467–E3475 (2014).
30. J. Harel, C. Koch, P. Perona, Graph-based visual saliency. *Adv. Neural Inf. Process. Syst.* **19**, 545–552 (2007).
31. A. Oliva, A. Torralba, Modeling the shape of the scene: A holistic representation of the spatial envelope. *Int. J. Comput. Vis.* **42**, 145–175 (2001).
32. K. Simonyan, A. Zisserman, Very deep convolutional networks for large-scale image recognition. arXiv [Preprint] (2014). https://arxiv.org/abs/1409.1556 (Accessed 1 October 2021).
33. W. S. Hunter, S. M. Garn, Disproportionate sexual dimorphism in the human face. *Am. J. Phys. Anthropol.* **36**, 133–138 (1972).
34. L. Reddy, P. Wilken, C. Koch, Face-gender discrimination is possible in the near-absence of attention. *J. Vis.* **4**, 106–117 (2004).
35. J. Haberman, D. Whitney, Rapid extraction of mean emotion and gender from sets of faces. *Curr. Biol.* **17**, R751–R753 (2007).
36. D. Rekow, J.-Y. Baudouin, B. Rossion, A. Leleu, An ecological measure of rapid and automatic face-sex categorization. *Cortex* **127**, 150–161 (2020).
37. A. M. Burton, V. Bruce, N. Dench, What's the difference between men and women? Evidence from facial measurement. *Perception* **22**, 153–176 (1993).
38. E. Brown, D. I. Perrett, What gives a face its gender? *Perception* **22**, 829–840 (1993).
39. A. J. O'Toole *et al.*, The perception of face gender: The role of stimulus structure in recognition and classification. *Mem. Cognit.* **26**, 146–160 (1998).
40. I. S. Penton-Voak, J. Y. Chen, High salivary testosterone is linked to masculine male facial appearance in humans. *Evol. Hum. Behav.* **25**, 229–241 (2004).
41. M. J. L. Smith *et al.*, Facial appearance is a cue to oestrogen levels in women. *Proc. Biol. Sci.* **273**, 135–140 (2006).
42. Y. Abir, A. Y. Sklar, R. Dotsch, A. Todorov, R. R. Hassin, The determinants of consciousness of human faces. *Nat. Hum. Behav.* **2**, 194–199 (2017).
43. K. Dobs, L. Isik, D. Pantazis, N. Kanwisher, How face perception unfolds over time. *Nat. Commun.* **10**, 1258 (2019).
44. P. C. Quinn, J. Yahr, A. Kuhn, A. M. Slater, O. Pascalils, Representation of the gender of human faces by infants: A preference for female. *Perception* **31**, 1109–1121 (2002).
45. G. Besson *et al.*, From face processing to face recognition: Comparing three different processing levels. *Cognition* **158**, 33–43 (2017).
46. V. Goffaux *et al.*, From coarse to fine? Spatial and temporal dynamics of cortical face processing. *Cereb. Cortex* **21**, 467–476 (2011).
47. M. Zhao, W. G. Hayward, Integrative processing of invariant aspects of faces: Effect of gender and race processing on identity analysis. *J. Vis.* **13**, 15 (2013).
48. B. Rossion, Is sex categorization from faces really parallel to face recognition? *Vis. Cogn.* **9**, 1003–1020 (2002).