Article

PERCEPTION

Perception
2025, Vol. 54(4) 252–265

DOI: 10.1177/03010066251316456
journals.sagepub.com/home/pec
Sage

Inversion produces opposite size illusions for faces and food

Henry Millbank
Birkbeck, University of London, UK

Eamonn Walsh
Institute of Psychiatry, Psychology & Neuroscience, King's College London, UK

Matthew R. Longo
Birkbeck, University of London, UK

Abstract

Faces are important communicative signals in humans and face perception is believed to involve specialised mechanisms in the visual system. Several other categories of stimuli are also thought to involve specialised processes, including bodies, letters, places, and food. A recently described face size illusion shows that upright faces appear physically smaller than identical inverted faces. This illusion appears to be highly face-specific, not occurring for other stimulus categories, such as bodies, letters, and hands. In this study, we investigated whether an analogous size inversion illusion occurs for items of food, a category which has recently been found to also involve specialised processes in the visual system. The results provided a clear replication of the face size illusion, with upright faces seen as smaller than inverted faces. In contrast, items of food and everyday objects showed an effect in the opposite direction, appearing larger when upright than when inverted. These results provide further evidence for the highly face-selective nature of the face size illusion. They also provide evidence for a different size illusion which affects visual perception of food.

Keywords

face perception, spatial cognition, object recognition, perception

Date Received: 21 June 2024; accepted: 8 January 2025

Introduction

Face perception is fundamental to our social interactions with other people, and relies on highly-specialised visual processes. Face perception is believed to rely on a specific form of "configural"

Corresponding author:
Matthew R. Longo, School of Psychological Sciences, Birkbeck, University of London, Malet Street,
London WC1E 7HX, UK.
Email: m.longo@bbk.ac.uk

processing, qualitatively different than the "featural" processing used for recognising other objects (Piepers & Robbins, 2012). A common approach to investigating configural processing in faces is using inversion effects. Whereas upright faces are thought to be processed configurally, inverted faces are processed more featurally, whereby the eyes, nose, mouth, and other facial features are analysed separately (Carey & Diamond, 1977; Farah et al., 1995; Maurer et al., 2002). Inversion has a range of effects on face perception, disrupting recognition of identity (Carey & Diamond, 1977; Yin, 1969), emotion recognition (Bombari et al., 2013; McKelvie, 1995; Prkachin, 2003), detection of distorted features (Thompson, 1980), and perception of size (Araragi et al., 2012; Walsh et al., 2018). In terms of this last effect, Araragi and colleagues (2012) described a size illusion in which upright faces are perceived as smaller than inverted faces. This effect has been replicated in a number of subsequent studies (Walsh et al., 2018; Walsh, Moreira et al., 2024; Walsh, Whitby et al., 2024; Zhang et al., 2021), which have also shown a high degree of specificity for the illusion to faces.

Neuroimaging research has linked the behavioural face inversion effect to the face-selective fusiform face area or FFA (Yovel & Kanwisher, 2005). The FFA is one of a relatively small set of regions in the ventral visual pathway, typically found bilaterally, which appear to be highly selective for specific categories of visual stimuli (Kanwisher, 2010). Neighbouring areas include another face selective region known as the occipital face area (Pitcher et al., 2011; Tsantani et al., 2021), two body selective regions known as the extrastriate body areas (Downing et al., 2001) and the fusiform body area (Peelen & Downing, 2005), a region selective for places and natural scenes also known as the parahippocampal place area (Epstein & Kanwisher, 1998), a region selective for written letters and words (Cohen et al., 2000), and a hand-selective region (Bracci et al., 2010, 2012).

One additional category of visual stimulus that has received substantial attention is food (Rumiati & Foroni, 2016), which consists of substances which are consumed by organisms and provide the nutrition and energy necessary for life. Images of food activate a broad, distributed brain network, including the orbitofrontal cortex (Killgore & Yurgelun-Todd, 2007; Mengotti et al., 2019; Simmons et al., 2005), medial prefrontal cortex (Goldstone et al., 2009; Killgore & Yurgelun-Todd, 2007; Plassmann et al., 2010), the anterior cingulate cortex (Frank et al., 2010), and the insula (Avery et al., 2021; Simmons et al., 2013). For meta-analyses, see Van Der Laan and colleagues (2011) and Zheng and colleagues (2022).

Images of food also produce robust activations within the ventral visual pathway, including regions on the aforementioned fusiform gyrus (Adamson & Troiani, 2018; Killgore & Yurgelun-Todd, 2007; Van Meer et al., 2015). These responses have been shown to relate both to the nutritional content of the depicted food items (Frank et al., 2010; Siep et al., 2009; Uher et al., 2006), how hungry the participant is (Frank et al., 2010; Führer et al., 2008; Siep et al., 2009; Uher et al., 2006), and whether the participants are of healthy weight or overweight (Frankort et al., 2012; Nummenmaa et al., 2012). Nevertheless, until recently, food was not thought to produce the sort of category-selective response that has been described for categories such as faces, bodies, and words (Kanwisher, 2010).

Three recent studies of large fMRI datasets have challenged this view, indicating there may be additional selective responses for images of food (Jain et al., 2023; Khosla et al., 2022; Pennock et al., 2023). Each of these studies took advantage of a large publicly available dataset in which fMRI data was collected while participants were shown a wide variety of naturalistic images of many categories of objects, people, and animals (Allen et al., 2022). For example, Khosla and colleagues (2022) used a data-driven factorization method to identify maximally-independent groupings of stimuli in terms of the patterns of neural activations they produced within the ventral visual pathway, identifying five such components. Four of these components corresponded to known category selective brain regions, namely faces, bodies, scenes, and letters. Interestingly, the final component corresponded to food. Another study by Jain and colleagues (2023) similarly identified food selectivity and found the spatial localisation of food responses to be more variable across participants than responses to other categories, which could explain why previous studies failed to identify food-selective brain regions.

Another recent study has provided a different interpretation of these putatively food-specific responses (Ritchie et al., 2024). Ritchie and colleagues compared responses in the ventral visual pathway to food and to tools, finding a high level of similarity in terms of the location, magnitude, and spatial pattern of activations to these two types of stimulus. They suggest that apparent food-specificity may actually relate to the manual affordances of food, rather than reflecting a true category-selective response.

As mentioned above, Araragi and colleagues (2012) showed that faces appear physically smaller when upright than when inverted. This face size illusion has been replicated in several recent studies (Walsh et al., 2018; Walsh, Moreira et al., 2024; Walsh, Whitby et al., 2024; Zhang et al., 2021), which have also shown that it has a high level of specificity to faces. Walsh and colleagues (2018) found clear underestimation of the size of upright faces, but no similar effect for hands or everyday objects, and an effect in the opposite direction for headless bodies (i.e., inverted bodies were perceived as smaller than upright ones). In a subsequent study, Walsh, Moreira et al. (2024) again found a clear underestimation of upright faces, but an effect in the opposite direction for letters and words, similar to that for bodies. To date, the size underestimation of upright stimuli originally described by Araragi and colleagues (2012) has been found only for faces, but not for other categories of stimuli.

Given recent studies showing category-selective neural responses to food in the ventral visual pathway (Jain et al., 2023; Khosla et al., 2022; Pennock et al., 2023), as well as evidence that responses to food and to faces may partly overlap within the fusiform gyrus (Adamson & Troiani, 2018), this study investigated whether there is a size illusion for food, analogous to that found for faces. On each trial, participants saw the same face, or food item, or everyday object, presented as a pair, one of which was upright and the other inverted (Figure 1). The participant's task was to judge which of the two objects appeared physically larger. By varying the size of the inverted stimulus across trials, we estimated whether there was a size bias for inverted relative to upright stimuli. We expected that inverted faces should appear larger than upright ones, consistent with previous studies (Araragi et al., 2012; Walsh et al., 2018; Walsh, Moreira et al., 2024; Zhang et al., 2021), and that no such effect should occur for everyday objects as observed in our recent study (Walsh et al., 2018). The key question was whether a face-like inversion effect would occur for food, or whether this size illusion is face-specific. Finally, given evidence that hunger modulates neural responses to seeing images of food (Frank et al., 2010; Führer et al., 2008; Siep et al., 2009; Uher et al., 2006), we collected information about how hungry participants were.

Method

Participants

Fifty-four individuals in the United Kingdom, recruited from the social network of the researchers ($N = 4$) and from the Prolific service (www.prolific.co; $N = 50$), participated after giving informed consent. Participants were naïve to the purpose of the study. Ethical approval was provided by the School of Psychological Science Research Ethics Committee, Birkbeck, University of London. Data from an additional seven participants was excluded from analyses due to low model fit (see below); these participants were replaced to meet our pre-determined sample size (see below). The participants ranged from 20 to 69 years of age (*M*: 38.1, *SD*: 12.6); 27 were female, 26 were male, and one self-described their gender as 'other'.

Figure 1. Example stimuli, including faces (top row), food (middle row), and objects (bottom row).

Prior to sample collection, an *a priori* power analysis was conducted using G*Power version 3.1.9.7 (Faul et al., 2007) to determine the minimum sample size required to test our hypotheses. In our previous studies using this paradigm (Walsh et al., 2018; Walsh, Moreira et al., 2024; Walsh, Whitby et al., 2024), the effect size for the size illusion for faces has been large. A weighted-average of the Cohen's d for t-tests comparing PSEs to 0 gave an overall average of $d = 0.907$. As we suspected that any potential effect for items of food might be smaller than this, we aimed to have 95% power to detect an effect of Cohen's $d = 0.50$, at a significance criterion of $\alpha = .05$, which produced an estimate of 54 participants.

Stimuli

The stimulus set consisted of 24 images: 8 different faces, 8 food items, and 8 inanimate objects. Figure 1 shows examples from each stimulus category. Face stimuli were from the Karolinska Directed Emotional Faces database (Lundqvist et al., 1998) and included the neutral facial expression of 4 men and 4 women. The neck and background were cropped so only the face was visible.

Both food and object stimuli were taken from the *Food Pics Extended* database (Blechert et al., 2019). For both food and object stimuli, we selected items which were familiar and typically observed in a canonical upright orientation. The food items selected were: a slice of cake, bars of chocolate, an ice cream cone, a brownie, strawberries, a slice of pizza, a piece of nigiri sushi, and a hamburger. The objects selected were: a sofa, a chair, a basket, an umbrella, a kettle, a coffee pot, a bedside table, and a watering pail. All stimuli were in colour given the association between food and colour selectivity in the ventral visual pathway (Pennock et al., 2023). All images had contextual environmental features removed and were cropped and centralised to maximise occupied square space, and on a white background (Figure 1).

To check that our intuitions about the canonical upright orientation of stimuli matched those of other people, we asked eight individuals to make judgments of which orientation appeared to be 'upright'. Each of the 24 images was shown with upright and inverted versions side by side. The participant was asked to click on the stimulus that appeared to be upright. Overall, participants clicked on what we had labelled the 'upright' stimulus on 100% of face trials, and 98.4% of object and food trials, showing overwhelming agreement with the assignments we made based on our own intuitions.

Procedure

The experiment was conducted online using the Gorilla platform (https://gorilla.sc/) (Anwyl-Irvine et al., 2020). Participants were physically located in the UK, and could complete the study using desktop computers, laptops, and tablets, but not mobile phones.

Participants judged which of two images presented either side of a central cross appeared to be physically larger. Each trial began with a centrally displayed cross for 500 ms, followed by the two images (see Figure 2). The upright stimulus was always shown at 500 pixels in height (the physical size and viewing distance varied depending on the specific computer used by the participant). The

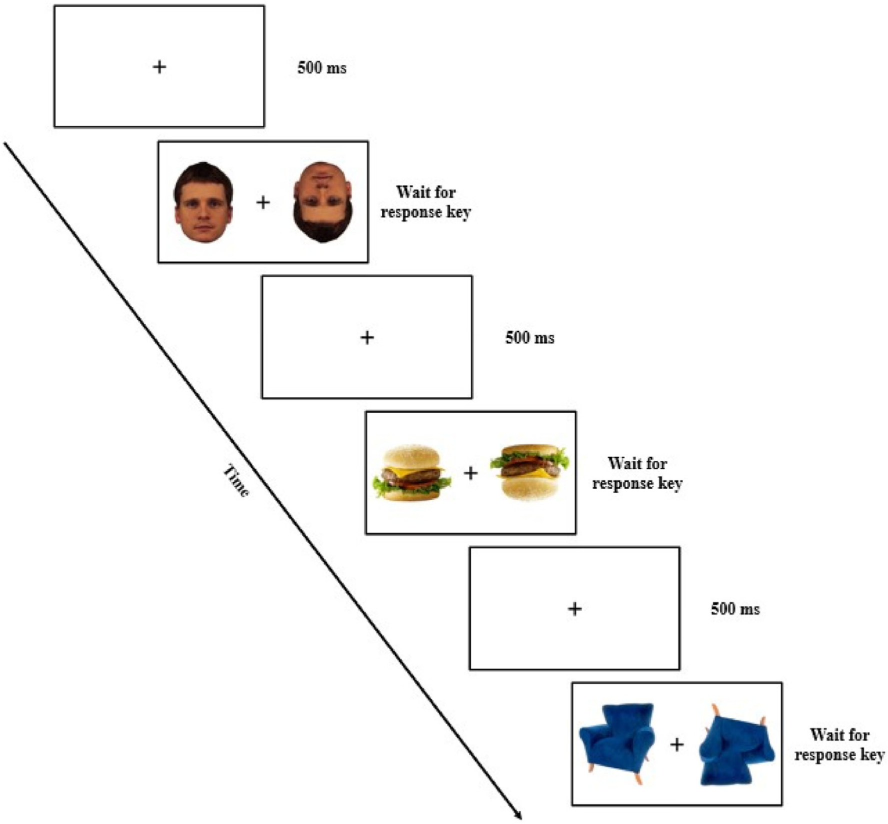

Figure 2. Schematic showing typical trials. For presentation purposes, one trial of each stimulus category is shown, but contrary to above, each stimulus category was in fact presented in separate blocks of trials.

size of the inverted stimulus was manipulated across trials according to the method of constant stimuli. Seven sizes of the inverted stimulus were used, corresponding to an increase of the linear dimensions of the image of −9, −6, −3, 0, 3, 6, or 9% (i.e., 455, 470, 485, 500, 515, 530, 545 pixels). Thus, across trials only the size of the inverted stimulus changed, while the upright stimulus which could appear on the left or right of the screen, remained a constant size.

There were 6 blocks of trials consisting of 56 trials each, including 8 repetitions of each of the 7 inverted sizes. Of these 8 repetitions, 4 had the upright stimulus on the left and the inverted stimulus on the right, and the remaining 4 the reverse. This resulted in a total of 336 trials. Each trial block contained stimuli from a single stimulus category. The first 3 blocks consisted of one repetition of all three stimulus categories, counterbalanced across participants according to a Latin square. The last 3 blocks were the reverse order of the first 3. The trials within each block were presented in random order.

Prior to the experiment proper, participants were able to complete a practice trial block consisting of 6 trials using an example of each stimulus type. The total duration of the experiment was approximately 20 min.

Following the main experiment, we collected information about participants' hunger levels and the appeal of the specific food items shown. Hunger levels were measured using four questions, broadly based on the Grand Hunger Scale (Grand, 1968). These questions include the following, each based on a 1–5 scale: "time since last meal" (0–30, 30–60, 60–120, 120–180, 180+ minutes); "how would you describe your hunger right now" (totally satisfied, not uncomfortable, hunger is gone for a while; satisfied, slightly full stomach but could eat a little more; neutral, neither hungry or full; slightly empty stomach, starting to feel hungry, could wait if needed; very hungry, eager to eat something, stomach growling); "how much of your favourite food subject could you imagine eating at this time" (none at all; not much; indifferent - take it or leave it; considerable amount; as much as I can get); and "time until next meal" (0–30, 30–60, 60–120, 120–180, 180+ minutes). A hunger composite score was calculated using the mean score of all four individual hunger questions ($M = 2.9$, $SD = 0.7$). This reflects a moderate level of hunger for all food items across the sample.

Finally, at the end of the study the appeal of each food item was measured using two questions. Appeal questions were assigned to each food image employed and when presented on screen, the food image appeared centrally aligned, with the question and responses below it. For each food image, appeal was measured using two questions measuring *liking* (likeability) and *wanting* (palatability), the responses for which were also quantified, and average composite score calculated per food item and across the whole set of food stimuli. The questions include the following: "how pleasant would it be to experience a mouthful of this food now" (none at all; not much; indifferent; neither pleasant or distasteful; somewhat; extremely); and, "how much do you want to eat some of this food now" (none at all; not much; indifferent - take it or leave it; somewhat; extremely). Appeal composite scores were calculated per food item using the mean score of both appeal questions. These individual appeal composite scores were then averaged to produce an overall appeal score (1–5 scale) for all food items employed ($M = 3.2$, $SD = 0.8$). This reflects a moderate level of appeal for all food items across the sample.

Analysis

Analyses were similar to those used in our previous studies using this paradigm (Walsh et al., 2018; Walsh, Moreira et al., 2024; Walsh, Whitby et al., 2024). Separate psychometric curves were constructed for each stimuli condition (face, food, object) per participant. The proportion of responses where the inverted stimulus was judged larger was modelled as a function of the difference in size between inverted and upright stimuli by constructing a cumulative Gaussian curve. This was conducted utilising maximum likelihood estimation via the Palamedes toolbox (Prins & Kingdom, 2009) in MATLAB (Mathworks, Natick, MA). The point of subjective equality (PSE, i.e., the mean of the best-fitting Gaussian), the slope of the psychometric function (i.e., inverse standard deviation), and the goodness of fit (R^2) were subsequently calculated for each curve.

The PSE estimates the size difference between upright and inverted stimuli where participants perceived these to be the same size. The difference in linear dimensions of the inverted stimuli is quantified as a percentage of the standard size. If there is no size difference between upright and inverted stimuli and no perceptual bias is present, stimuli should be perceived the same and PSEs should on average equal 0. Positive PSEs represent participants judging inverted stimulus as smaller than its upright counterpart; negative PSEs reflect the opposite. For each condition separately and to confirm the presence or not of a size illusion, we used one-sample t-tests to compare mean PSE to 0, where 0 indicates no size illusion (i.e., the upright and inverted stimuli are judged as equal in size when they really are equal in size).

Any participant with data measuring below predesignated exclusion criteria threshold ($R^2 < .50$) for any stimuli condition (face, food, object) was removed. This criterion is the same that we have used in previous studies (Walsh, Moreira et al., 2024; Walsh, Whitby et al., 2024). In total, seven participants were excluded, and seven new replacement participants recruited to match our *a priori* target sample size. Potential differences between conditions for R^2, PSEs, and slopes were assessed using repeated-measures analysis of variance (ANOVA). Where Mauchley's test indicated a violation of the sphericity assumption, the Greenhouse-Geisser correction was applied.

The complete set of images used and raw data are available on the OSF website: https://osf.io/dvkq9/

Results

Results are shown in Figure 3. There was good overall fit of the psychometric functions to the data. Mean R^2 values were .963, .926, and .933, in the face, food, and object conditions, respectively. An ANOVA showed that R^2 values differed significantly across conditions, $F(2, 106) = 6.26$, $p < .005$, $\eta_p^2 = .106$. Follow-up tests showed that R^2 values were significantly higher for faces than for food, $t(53) = 3.13$, $p < .005$, $d_z = 0.425$, and objects, $t(53) = 2.83$, $p < .01$, $d_z = 0.385$. R^2 values for food and objects did not differ significantly, $t(53) = 0.71$, $p = .48$, $d_z = 0.097$. Nevertheless, all values indicated strong overall fit to the data.

PSEs for each category were compared to 0 using one-sample *t*-tests (two-tailed) to measure potential overall biases. There was a significant bias for faces to be perceived as smaller when upright than when inverted (*M*: −2.26%; *SD*: 1.97), $t(53) = -8.43$, $p < .0001$, Cohen's $d = 1.148$. This demonstrates a successful replication of the inverted face illusion reported by previous studies (Araragi et al., 2012; Walsh et al., 2018).

A significant bias in the opposite direction was observed for both food (*M*: 1.04%; *SD*: 2.63), $t(53) = 2.92$, $p = .005$, $d = 0.398$, and objects (*M*: 0.75%; *SD*: 2.29), $t(53) = 2.42$, $p < .02$, $d = 0.329$. This indicates inverted food and objects were perceived smaller than their upright equivalents.

An ANOVA on PSEs showed significant differences across conditions, $F(1.45, 76.97) = 43.05$, $p < 0.0001$, $\eta_p^2 = .448$. Follow-up *t*-tests with Holm-Bonferroni correction for multiple comparisons revealed significantly differences between faces and food, $t(53) = 6.85$, $p < .0001$, $d_z = 0.932$, and between faces and objects, $t(53) = 7.41$, $p < .0001$, $d_z = 1.008$. There was no significant difference between PSEs for food and objects, $t(53) = 1.11$, $p = 0.274$, $d_z = 0.151$.

An ANOVA on slopes revealed significant differences across conditions, $F(1.61, 85.3) = 35.87$, $p < .0001$, $\eta_p^2 = .404$. Follow-up *t*-tests with Holm-Bonferroni correction indicated that slopes were steeper for faces (*M*: 0.304; *SD*: 0.145) than for food (*M*: 0.206; *SD*: 0.091), $t(53) = 5.92$, $p < .0001$, $d_z = 0.805$, or for objects (*M*: 0.187; *SD*: 0.081), $t(53) = 7.07$, $p < .0001$, $d_z = 0.962$. There was no

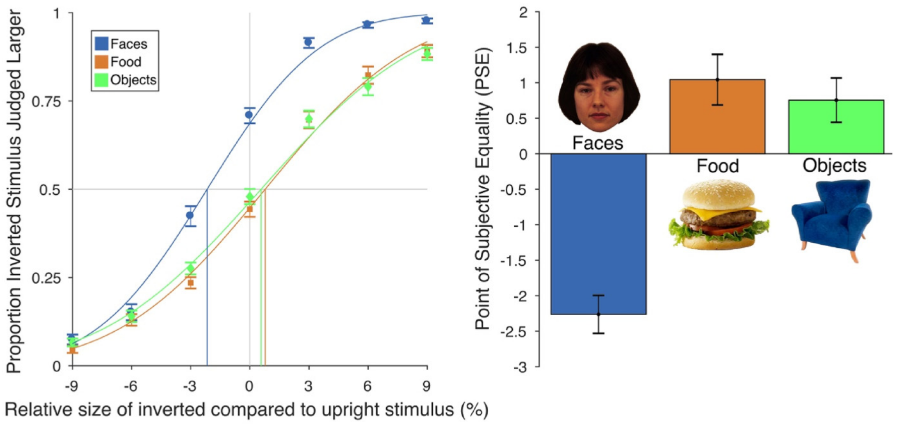

Figure 3. *Left panel*: mean probability of trials where the inverted stimulus (face, food, object) was perceived larger than the same-sized upright equivalent for experiment proper. 0% size comparison on the horizontal axis represents both inverted and upright stimuli being of objective equal size to each other. 0% on the vertical axis represents standard image size (500 square pixels). *Right panel*: Mean PSEs for each stimulus category. Positive PSE values indicate that the inverted stimulus was judged smaller than the equivalent-sized upright stimulus, while negative PSE values indicate the reverse. For faces, the PSEs were significantly negative, indicating a bias to perceive faces as smaller when upright than inverted, consistent with previous findings. For food and objects, in contrast, there was an effect in the opposite direction. Error bars represent one standard error.

significant difference in slope between food and objects, $t(53) = 1.81$, $p = .075$, $d_z = 0.247$. The steeper slopes for faces indicate that participants are more precise as distinguishing the relative size of faces than objects or food, despite the systematic bias described above.

To investigate whether hunger influenced size judgments about food, we calculated correlations between hunger composite scores and parameters of psychometric functions for food. There were no significant correlations between food PSEs and either hunger composite scores, $r(52) = -0.158$, $p = 0.254$, or appeal composite scores, $r(52) = 0.246$, $p = 0.072$ (see Figure 4).

Discussion

These results provide a clear replication of the face size illusion (Araragi et al., 2012; Walsh et al., 2018; Walsh, Moreira et al., 2024; Walsh, Whitby et al., 2024; Zhang et al., 2021). Upright faces were perceived as physically smaller than inverted faces. The key question was whether this effect is specific to faces, or would generalise to another important class of stimulus, namely food. Importantly, for food, there was a significant illusion in the opposite direction as for faces: upright food items were perceived as larger than inverted ones. Similarly, everyday upright objects were perceived as larger than inverted ones. The underestimation of upright stimuli, therefore, appears highly specific to faces.

The present results are consistent with other findings in showing a high degree of specificity of this illusion to faces. In one study, Walsh and colleagues (2018) showed that there was no comparable effect for hands and an opposite effect for headless bodies. Walsh, Moreira et al. (2024) found that there was also an effect in the opposite direction as faces for letters and words. Thus, while upright faces are consistently perceived as smaller than inverted ones, no analogous effect has been found for other types of stimulus tested thus far, including those known to rely-on category-selective mechanisms in the ventral visual pathway, such as bodies, hands, letters, and now food.

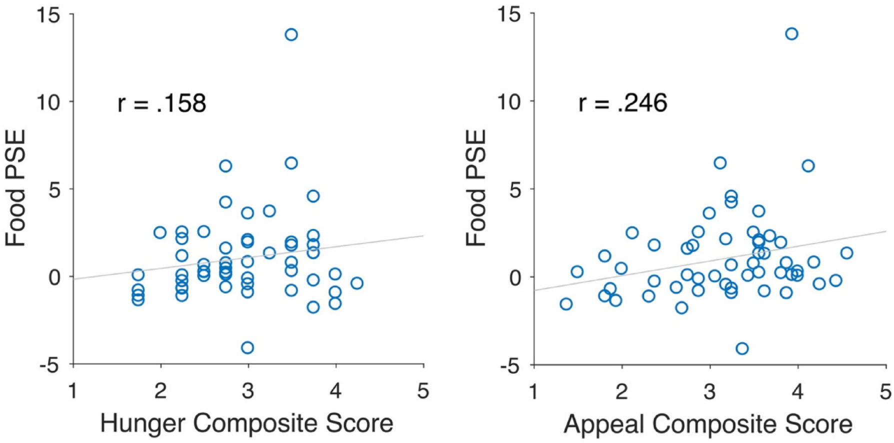

Figure 4. Scatterplots showing the relation between PSEs for food items and composite hunger scores (left panel) and composite food appeal scores (right panel).

The face size illusion may be a signature of the configural processing used by the visual system for processing faces. A recent fMRI study used population receptive field (pRF) mapping to measure the effects of inversion on spatial integration in face-selective regions of the ventral visual pathway (Poltoratski et al., 2021). Notably, pRFs were larger in response to upright than to inverted faces. This result is striking given that, in general, high levels of spatial sensitivity are associated with smaller RFs, not larger ones (Hubel & Wiesel, 1974). The relation between spatial sensitivity and RF size is also true of pRFs in fMRI, which are smaller in voxels representing the fovea than in those representing the periphery (Amano et al., 2009; Harvey & Dumoulin, 2011).

From this perspective, one might expect that a stimulus that humans are highly expert in, such as an upright face, should be associated with smaller, sharper RFs. Poltoratski and colleagues suggest that this increase in pRF size may reflect the operation of configural processing of faces which focuses on the broader spatial relationship among different features rather than individual face parts. Such spatial integration across the entire face may recruit neural populations with larger RFs. As there is known to be an inverse relation between pRF size and perceived object size (Moutsiana et al., 2016), the inversion effect on pRF size described by Poltoratski and colleagues provides a candidate neural mechanism for the face size illusion.

Similarly, there is evidence that direction of attention to stimuli produces systematic modulations of the location and size of RFs in the visual system (Anton-Erxleben & Carrasco, 2013; Treue & Martinez-Trujillo, 2012; Womelsdorf et al., 2006, 2008). Such attentional effects may underlie the pRF results described in the previous paragraph, given that there is evidence that faces attract attention compared to other stimuli (Langton et al., 2008; Theeuwes & Van Der Stigchel, 2006). As such, direction of attention to faces could underlie the face size illusion. Similarly, there are some difficulties, however, for an attentional interpretation of the illusion. First, there is evidence that the RF modulations from increased attention result in attended stimuli appearing *larger* than unattended stimuli (Anton-Erxleben et al., 2007), the opposite of the effect in the face size illusion. Second, though faces may attract attention, they do not show the “pop out” in visual search that would suggest strong automatic attentional orientation to faces (Nothdurft, 1993; VanRullen, 2006). Third, while faces may attract attention, there is no evidence that upright faces attract attention more than inverted faces (Bindemann & Burton, 2008; Kuehn & Jolicoeur, 1994). Finally, there is evidence that emotional faces attract attention differently from neutral faces (Bradley et al., 1997), but no such modulation is seen in the face size illusion (Walsh, Whitby et al., 2024).

Attention provides a potentially more promising mechanism for understanding why everyday objects and food items are perceived as larger when upright than when inverted. Objects seen in their conventional upright orientation may very well receive greater levels of attention than inverted stimuli, which could result in them being seen as larger, consistent with previous results showing that attended stimuli are perceived as larger than unattended stimuli (Anton-Erxleben et al., 2007). This could also provide a potential explanation for the finding that both letters and words are also perceived as larger when presented upright than when inverted (Walsh, Moreira et al., 2024). At the same time, hunger is known to direct attention to desired food items (Piech et al., 2010), so the absence of a relation between the size illusion and hunger ratings argues against an attentional interpretation.

The stimuli used for the different categories of object in this study came from different databases and were created by different researchers under different circumstances. There may thus be differences between stimuli beyond the semantic category to which they belong which could potentially have affected results. For example, some of the food images have highlights and shadows which could provide cues to the orientation of the object independent of knowledge of the object's canonical orientation. The shadows present in two of the images (the nigiri sushi and the chocolate bars) raise a further complication for size estimates if inversion modulates the extent to which the shadow is segmented from the object proper and not included in the estimate of stimulus size. It is most natural to suppose that it would be easier for the visual system to segment shadows from the object itself when the stimulus is in its canonical orientation and consistent with the light-from-above prior. But this would lead to the object being perceived as smaller when upright (since it should be easier to exclude the shadows from perceived object size), exactly the opposite of the size illusion that in fact occurred here for food items.

In conclusion, the present study found that items of food are perceived as physically larger when in their canonical upright orientation than when inverted. This effect was similar in magnitude to that seen for other, non-food, objects which also have a canonical orientation. In line with the lack of difference between food and non-food items, there was no apparent relation between the size illusion for food and either hunger or the appeal of food items. Notably, this effect is opposite in direction from that seen for faces which, consistent with other recent studies on the face-size illusion, were perceived as larger when inverted.

Author Contribution(s)

Henry Millbank: Conceptualization; Formal analysis; Investigation; Methodology; Software; Writing – original draft.

Eamonn Walsh: Conceptualization; Methodology; Writing – review & editing.

Matthew R. Longo: Conceptualization; Methodology; Project administration; Supervision; Visualization; Writing – review & editing.

Declaration of Conflicting Interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

ORCID iD

Matthew R. Longo https://orcid.org/0000-0002-2450-4903

References

Adamson, K., & Troiani, V. (2018). Distinct and overlapping fusiform activation to faces and food. *NeuroImage*, *174*, 393–406. https://doi.org/10.1016/j.neuroimage.2018.02.064

Allen, E. J., St-Yves, G., Wu, Y., Breedlove, J. L., Prince, J. S., Dowdle, L. T., Nau, M., Caron, B., Pestilli, F., Charest, I., Hutchinson, J. B., Naselaris, T., & Kay, K. (2022). A massive 7T fMRI dataset to bridge cognitive neuroscience and artificial intelligence. *Nature Neuroscience*, *25*, 116–126. https://doi.org/10.1038/s41593-021-00962-x

Amano, K., Wandell, B. A., & Dumoulin, S. O. (2009). Visual field maps, population receptive field sizes, and visual field coverage in the human MT+ complex. *Journal of Neurophysiology*, *102*, 2704–2718. https://doi.org/10.1152/jn.00102.2009

Anton-Erxleben, K., & Carrasco, M. (2013). Attentional enhancement of spatial resolution: Linking behavioural and neurophysiological evidence. *Nature Reviews Neuroscience*, *14*, 188–200. https://doi.org/10.1038/nrn3443

Anton-Erxleben, K., Henrich, C., & Treue, S. (2007). Attention changes perceived size of moving visual patterns. *Journal of Vision*, *7*, 5. https://doi.org/10.1167/7.11.5

Anwyl-Irvine, A. L., Massonnié, J., Flitton, A., Kirkham, N., & Evershed, J. K. (2020). Gorilla in our midst: An online behavioral experiment builder. *Behavior Research Methods*, *52*, 388–407. https://doi.org/10.3758/s13428-019-01237-x

Araragi, Y., Aotani, T., & Kitaoka, A. (2012). Evidence for a size underestimation of upright faces. *Perception*, *41*, 840–853. https://doi.org/10.1068/p7058

Avery, J. A., Liu, A. G., Ingeholm, J. E., Gotts, S. J., & Martin, A. (2021). Viewing images of foods evokes taste quality-specific activity in gustatory insular cortex. *Proceedings of the National Academy of Sciences*, *118*, e2010932118. https://doi.org/10.1073/pnas.2010932118

Bindemann, M., & Burton, A. M. (2008). Attention to upside-down faces: An exception to the inversion effect. *Vision Research*, *48*, 2555–2561. https://doi.org/10.1016/j.visres.2008.09.001

Blechert, J., Lender, A., Polk, S., Busch, N. A., & Ohla, K. (2019). Food-pics_extended—An image database for experimental research on eating and appetite: Additional images, normative ratings and an updated review. *Frontiers in Psychology*, *10*, 307. https://doi.org/10.3389/fpsyg.2019.00307

Bombari, D., Schmid, P. C., Schmid Mast, M., Birri, S., Mast, F. W., & Lobmaier, J. S. (2013). Emotion recognition: The role of featural and configural face information. *Quarterly Journal of Experimental Psychology*, *66*, 2426–2442. https://doi.org/10.1080/17470218.2013.789065

Bracci, S., Cavina-Pratesi, C., Ietswaart, M., Caramazza, A., & Peelen, M. V. (2012). Closely overlapping responses to tools and hands in left lateral occipitotemporal cortex. *Journal of Neurophysiology*, *107*, 1443–1456. https://doi.org/10.1152/jn.00619.2011

Bracci, S., Ietswaart, M., Peelen, M. V., & Cavina-Pratesi, C. (2010). Dissociable neural responses to hands and non-hand body parts in human left extrastriate visual cortex. *Journal of Neurophysiology*, *103*, 3389–3397. https://doi.org/10.1152/jn.00215.2010

Bradley, B. P., Mogg, K., Millar, N., Bonham-Carter, C., Fergusson, E., Jenkins, J., & Parr, M. (1997). Attentional biases for emotional faces. *Cognition & Emotion*, *11*, 25–42. https://doi.org/10.1080/026999397380014

Carey, S., & Diamond, R. (1977). From piecemeal to configurational representation of faces. *Science*, *195*, 312–314. https://doi.org/10.1126/science.831281

Cohen, L., Dehaene, S., Naccache, L., Lehéricy, S., Dehaene-Lambertz, G., Hénaff, F., & Michel, F. (2000). The visual word form area: Spatial and temporal characterization of an initial stage of reading in normal subjects and posterior split-brain patients. *Brain*, *123*, 291–307. https://doi.org/10.1093/brain/123.2.291

Downing, P. E., Jiang, Y., Shuman, M., & Kanwisher, N. (2001). A cortical area selective for visual processing of the human body. *Science*, *293*, 2470–2473. https://doi.org/10.1126/science.1063414

Epstein, R., & Kanwisher, N. (1998). A cortical representation of the local visual environment. *Nature*, *392*, 598–601. https://doi.org/10.1038/33402

Farah, M. J., Tanaka, J. W., & Drain, H. M. (1995). What causes the face inversion effect? *Journal of Experimental Psychology: Human Perception and Performance*, *21*, 628–634. https://doi.org/10.1037//0096-1523.21.3.628

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*, 175–191. https://doi.org/10.3758/BF03193146

Frank, S., Laharnar, N., Kullmann, S., Veit, R., Canova, C., Hegner, Y. L., Fritsche, A., & Preissl, H. (2010). Processing of food pictures: Influence of hunger, gender and calorie content. *Brain Research*, *1350*, 159–166. https://doi.org/10.1016/j.brainres.2010.04.030

Frankort, A., Roefs, A., Siep, N., Roebroeck, A., Havermans, R., & Jansen, A. (2012). Reward activity in satiated overweight women is decreased during unbiased viewing but increased when imagining taste: An event-related fMRI study. *International Journal of Obesity*, *36*, 627–637. https://doi.org/10.1038/ijo.2011.213

Führer, D., Zysset, S., & Stumvoll, M. (2008). Brain activity in hunger and satiety: An exploratory visually stimulated fMRI study. *Obesity*, *16*, 945–950. https://doi.org/10.1038/oby.2008.33

Goldstone, A. P., Prechtl De Hernandez, C. G., Beaver, J. D., Muhammed, K., Croese, C., Bell, G., Durighel, G., Hughes, E., Waldman, A. D., Frost, G., & Bell, J. D. (2009). Fasting biases brain reward systems towards high-calorie foods. *European Journal of Neuroscience*, *30*, 1625–1635. https://doi.org/10.1111/j.1460-9568.2009.06949.x

Grand, S. (1968). Color-word interference: II. An investigation of the role of vocal conflict and hunger in associative priming. *Journal of Experimental Psychology*, *77*(1), 31–40. http://doi.org/10.1037/h0025759

Harvey, B. M., & Dumoulin, S. O. (2011). The relationship between cortical magnification factor and population receptive field size in human visual cortex: Constancies in cortical architecture. *The Journal of Neuroscience*, *31*, 13604–13612. https://doi.org/10.1523/JNEUROSCI.2572-11.2011

Hubel, D. H., & Wiesel, T. N. (1974). Uniformity of monkey striate cortex: A parallel relationship between field size, scatter, and magnification factor. *Journal of Comparative Neurology*, *158*, 295–305. https://doi.org/10.1002/cne.901580305

Jain, N., Wang, A., Henderson, M. M., Lin, R., Prince, J. S., Tarr, M. J., & Wehbe, L. (2023). Selectivity for food in human ventral visual cortex. *Communications Biology*, *6*, 175. https://doi.org/10.1038/s42003-023-04546-2

Kanwisher, N. (2010). Functional specificity in the human brain: A window into the functional architecture of the mind. *Proceedings of the National Academy of Sciences*, *107*, 11163–11170. https://doi.org/10.1073/pnas.1005062107

Khosla, M., Ratan Murty, N. A., & Kanwisher, N. (2022). A highly selective response to food in human visual cortex revealed by hypothesis-free voxel decomposition. *Current Biology*, *32*, 4159–4171.e9. https://doi.org/10.1016/j.cub.2022.08.009

Killgore, W. D. S., & Yurgelun-Todd, D. A. (2007). Positive affect modulates activity in the visual cortex to images of high calorie foods. *International Journal of Neuroscience*, *117*, 643–653. https://doi.org/10.1080/00207450600773848

Kuehn, S. M., & Jolicoeur, P. (1994). Impact of quality of the image, orientation, and similarity of the stimuli on visual search for faces. *Perception*, *23*, 95–122. https://doi.org/10.1068/p230095

Langton, S. R. H., Law, A. S., Burton, A. M., & Schweinberger, S. R. (2008). Attention capture by faces. *Cognition*, *107*, 330–342. https://doi.org/10.1016/j.cognition.2007.07.012

Lundqvist, D., Flykt, A., & Öhman, A. (1998). *The Karolinska Directed Emotional Faces—KDEF, CD ROM from Department of Clinical Neuroscience, Psychology section, Karolinska Institutet, ISBN 91-630-7164-9.*

Maurer, D., Grand, R. L., & Mondloch, C. J. (2002). The many faces of configural processing. *Trends in Cognitive Sciences*, *6*, 255–260. https://doi.org/10.1016/S1364-6613(02)01903-4

McKelvie, S. J. (1995). Emotional expression in upside-down faces: Evidence for configurational and componential processing. *British Journal of Social Psychology*, *34*, 325–334. https://doi.org/10.1111/j.2044-8309.1995.tb01067.x

Mengotti, P., Foroni, F., & Rumiati, R. I. (2019). Neural correlates of the energetic value of food during visual processing and response inhibition. *NeuroImage*, *184*, 130–139. https://doi.org/10.1016/j.neuroimage.2018.09.017

Moutsiana, C., de Haas, B., Papageorgiou, A., van Dijk, J. A., Balraj, A., Greenwood, J. A., & Schwarzkopf, D. S. (2016). Cortical idiosyncrasies predict the perception of object size. *Nature Communications*, *7*, 12110. https://doi.org/10.1038/ncomms12110

Nothdurft, H.-C. (1993). Faces and facial expressions do not pop out. *Perception*, *22*, 1287–1298. https://doi.org/10.1068/p221287

Nummenmaa, L., Hirvonen, J., Hannukainen, J. C., Immonen, H., Lindroos, M. M., Salminen, P., & Nuutila, P. (2012). Dorsal striatum and its limbic connectivity mediate abnormal anticipatory reward processing in obesity. *PLoS One*, *7*, e31089. https://doi.org/10.1371/journal.pone.0031089

Peelen, M. V., & Downing, P. E. (2005). Selectivity for the human body in the fusiform gyrus. *Journal of Neurophysiology*, *93*, 603–608. https://doi.org/10.1152/jn.00513.2004

Pennock, I. M. L., Racey, C., Allen, E. J., Wu, Y., Naselaris, T., Kay, K. N., Franklin, A., & Bosten, J. M. (2023). Color-biased regions in the ventral visual pathway are food selective. *Current Biology*, *33*, 134–146.e4. https://doi.org/10.1016/j.cub.2022.11.063

Piech, R. M., Pastorino, M. T., & Zald, D. H. (2010). All I saw was the cake. Hunger effects on attentional capture by visual food cues. *Appetite*, *54*, 579–582. https://doi.org/10.1016/j.appet.2009.11.003

Piepers, D. W., & Robbins, R. A. (2012). A review and clarification of the terms “holistic,” “configural,” and “relational” in the face perception literature. *Frontiers in Psychology*, *3*, 559. https://doi.org/10.3389/fpsyg.2012.00559

Pitcher, D., Walsh, V., & Duchaine, B. (2011). The role of the occipital face area in the cortical face perception network. *Experimental Brain Research*, *209*, 481–493. https://doi.org/10.1007/s00221-011-2579-1

Plassmann, H., O’Doherty, J. P., & Rangel, A. (2010). Appetitive and aversive goal values are encoded in the medial orbitofrontal cortex at the time of decision making. *The Journal of Neuroscience*, *30*, 10799–10808. https://doi.org/10.1523/JNEUROSCI.0788-10.2010

Poltoratski, S., Kay, K., Finzi, D., & Grill-Spector, K. (2021). Holistic face recognition is an emergent phenomenon of spatial processing in face-selective regions. *Nature Communications*, *12*, 4745. https://doi.org/10.1038/s41467-021-24806-1

Prins, N., & Kingdom, F. A. A. (2009). *Palamedes: Matlab routines for analyzing psychophysical data*. http://www.palamedestoolbox.org

Prkachin, G. C. (2003). The effects of orientation on detection and identification of facial expressions of emotion. *British Journal of Psychology*, *94*, 45–62. https://doi.org/10.1348/000712603762842093

Ritchie, J. B., Andrews, S. T., Vaziri-Pashkam, M., & Baker, C. I. (2024). Graspable foods and tools elicit similar responses in visual cortex. *Cerebral Cortex*, *34*, bhae383. https://doi.org/10.1093/cercor/bhae383

Rumiati, R. I., & Foroni, F. (2016). We are what we eat: How food is represented in our mind/brain. *Psychonomic Bulletin & Review*, *23*, 1043–1054. https://doi.org/10.3758/s13423-015-0908-2

Siep, N., Roefs, A., Roebroeck, A., Havermans, R., Bonte, M. L., & Jansen, A. (2009). Hunger is the best spice: An fMRI study of the effects of attention, hunger and calorie content on food reward processing in the amygdala and orbitofrontal cortex. *Behavioural Brain Research*, *198*, 149–158. https://doi.org/10.1016/j.bbr.2008.10.035

Simmons, W. K., Martin, A., & Barsalou, L. W. (2005). Pictures of appetizing foods activate gustatory cortices for taste and reward. *Cerebral Cortex*, *15*, 1602–1608. https://doi.org/10.1093/cercor/bhi038

Simmons, W. K., Rapuano, K. M., Kallman, S. J., Ingeholm, J. E., Miller, B., Gotts, S. J., Avery, J. A., Hall, K. D., & Martin, A. (2013). Category-specific integration of homeostatic signals in caudal but not rostral human insula. *Nature Neuroscience*, *16*, 1551–1552. https://doi.org/10.1038/nn.3535

Theeuwes, J., & Van Der Stigchel, S. (2006). Faces capture attention: Evidence from inhibition of return. *Visual Cognition*, *13*, 657–665. https://doi.org/10.1080/13506280500410949

Thompson, P. (1980). Margaret thatcher: A new illusion. *Perception*, *9*, 483–484. https://doi.org/10.1068/p090483

Treue, S., & Martinez-Trujillo, J. C. (2012). The spotlight of attention: Shifting, resizing and splitting receptive fields when processing visual motion. *E-Neuroforum*, *18*, 74–79. https://doi.org/10.1007/s13295-012-0034-9

Tsantani, M., Kriegeskorte, N., Storrs, K., Williams, A. L., McGettigan, C., & Garrido, L. (2021). FFA And OFA encode distinct types of face identity information. *The Journal of Neuroscience*, *41*, 1952–1969. https://doi.org/10.1523/JNEUROSCI.1449-20.2020

Uher, R., Treasure, J., Heining, M., Brammer, M. J., & Campbell, I. C. (2006). Cerebral processing of food-related stimuli: Effects of fasting and gender. *Behavioural Brain Research*, *169*, 111–119. https://doi.org/10.1016/j.bbr.2005.12.008

Van Der Laan, L. N., De Ridder, D. T. D., Viergever, M. A., & Smeets, P. A. M. (2011). The first taste is always with the eyes: A meta-analysis on the neural correlates of processing visual food cues. *NeuroImage*, *55*, 296–303. https://doi.org/10.1016/j.neuroimage.2010.11.055

Van Meer, F., Van Der Laan, L. N., Adan, R. A. H., Viergever, M. A., & Smeets, P. A. M. (2015). What you see is what you eat: An ALE meta-analysis of the neural correlates of food viewing in children and adolescents. *NeuroImage*, *104*, 35–43. https://doi.org/10.1016/j.neuroimage.2014.09.069

VanRullen, R. (2006). On second glance: Still no high-level pop-out effect for faces. *Vision Research*, *46*, 3017–3027. https://doi.org/10.1016/j.visres.2005.07.009

Walsh, E., Moreira, C., & Longo, M. R. (2024). Opposite size illusions for inverted faces and letters. *Cognition*, *245*, 105733. https://doi.org/10.1016/j.cognition.2024.105733

Walsh, E., Vormberg, A., Hannaford, J., & Longo, M. R. (2018). Inversion produces opposite size illusions for faces and bodies. *Acta Psychologica*, *191*, 15–24. https://doi.org/10.1016/j.actpsy.2018.08.017

Walsh, E., Whitby, J., Chen, Y.-Y., & Longo, M. R. (2024). No influence of emotional expression on size underestimation of upright faces. *PLoS One*, *19*, e0293920. https://doi.org/10.1371/journal.pone.0293920

Womelsdorf, T., Anton-Erxleben, K., Pieper, F., & Treue, S. (2006). Dynamic shifts of visual receptive fields in cortical area MT by spatial attention. *Nature Neuroscience*, *9*, 1156–1160. https://doi.org/10.1038/nn1748

Womelsdorf, T., Anton-Erxleben, K., & Treue, S. (2008). Receptive field shift and shrinkage in macaque middle temporal area through attentional gain modulation. *The Journal of Neuroscience*, *28*, 8934–8944. https://doi.org/10.1523/JNEUROSCI.4030-07.2008

Yin, R. K. (1969). Looking at upside-down faces. *Journal of Experimental Psychology*, *81*, 141–145. https://doi.org/10.1037/h0027474

Yovel, G., & Kanwisher, N. (2005). The neural basis of the behavioral face-inversion effect. *Current Biology*, *15*, 2256–2262. https://doi.org/10.1016/j.cub.2005.10.072

Zhang, Y., Wang, L., & Jiang, Y. (2021). My own face looks larger than yours: A self-induced illusory size perception. *Cognition*, *212*, 104718. https://doi.org/10.1016/j.cognition.2021.104718

Zheng, L., Miao, M., & Gan, Y. (2022). A systematic and meta-analytic review on the neural correlates of viewing high- and low-calorie foods among normal-weight adults. *Neuroscience & Biobehavioral Reviews*, *138*, 104721. https://doi.org/10.1016/j.neubiorev.2022.104721