Walsh, E. and Moreira, C. and Longo, Matthew (2024) Opposite size illusions for inverted faces and letters. Cognition 245 (105733), ISSN 0010-0277.

Cognition 245 (2024) 105733

Contents lists available at ScienceDirect

Cognition

journal homepage: www.elsevier.com/locate/cognit

Opposite size illusions for inverted faces and letters

Eamonn Walsh [a,b,*], Carolina Moreira [c], Matthew R. Longo [c]

[a] Department of Basic & Clinical Neuroscience, Institute of Psychiatry, Psychology & Neuroscience, King's College London, UK
[b] Cultural and Social Neuroscience Research Group, Institute of Psychiatry, Psychology & Neuroscience, King's College London, UK
[c] Department of Psychological Sciences, Birkbeck, University of London, UK

ARTICLE INFO

Keywords:
Holistic processing
Size illusion
Inversion effect
Receptive field (RF)

ABSTRACT

Words are the primary means by which we communicate meaning and ideas, while faces provide important social cues. Studying visual illusions involving faces and words can elucidate the hierarchical processing of information as different regions of the brain are specialised for face recognition and word processing. A size illusion has previously been demonstrated for faces, whereby an inverted face is perceived as larger than the same stimulus upright. Here, two experiments replicate the face size illusion, and investigate whether the illusion is also present for individual letters (Experiment 1), and visual words and pseudowords (Experiment 2). Results confirm a robust size Illusion for faces. Letters, words and pseudowords and unfamiliar letters all show a reverse size illusion, as we previously demonstrated for human bodies. Overall, results indicate the illusion occurs in early perceptual stages upstream of semantic processing. Results are consistent with the idea of a general-purpose mechanism that encodes curvilinear shapes found in both scripts and our environment. Word and face perception rely on specialised, independent cognitive processes. The underestimation of the size of upright stimuli is specific to faces. Opposite size illusions may reflect differences in how size information is encoded and represented in stimulus-specialised neural networks, resulting in contrasting perceptual effects. Though words and faces differ visually, there is both symmetry and asymmetry in how the brain 'reads' them.

1. Introduction

At first glance, words and faces are very different. Faces typically involve the same general structure (2 eyes, above a nose, above a mouth) while the spatial relationships among individual parts help inform identity (K. R. Brooks & Kemp, 2007; Mondloch, Le Grand, & Maurer, 2002). The relative size of each facial feature can differ and the distances between these features can also vary. For the recognition of words, the number, and order of letters can vary and the distances between letters are less relevant (Wong & Gauthier, 2007). Words in the English language borrow from an alphabet of 26 letters, which allows for a huge combination of different ways to order the letters, making for many different possible words. As with letters, permutations in the configuration of facial features produce a great range of human facial diversity, and with very few genetically identical exceptions; no two faces in a population are quite the same (Sheehan & Nachman, 2014).

The evolutionary time-course of faces and words is different. The fossil record shows that the face we would recognise as characteristically human first appeared in Africa around 160,000 years ago (Lacruz et al., 2019). It is thought that humans developed some capacity for words at least 50,000 years ago (Dediu & Levinson, 2013; Perreault & Mathew, 2012). The first written lexicographic letters and words were Sumerian and were introduced only about 5500 years ago (Perreault & Mathew, 2012), and were confined to a small minority of intellectual elites (Glassner & Herron, 2003). In this sense, reading is considered an "artificial ability" (Sereno & Rayner, 2003) or "cognitive gadget" (Heyes, 2018); a recent product of cultural rather than of genetic evolution. In an individual's life, faces are first encountered at birth. Though the spoken word is usually also heard at birth, letters and the written word are learnt later and through more formal educational instruction. Taken together, these factors imply that words and faces ought to be processed differently in the human brain.

Studying visual illusions can elucidate the hierarchical processing of visual information in the brain as different regions of the brain are specialised for different stimuli. An example is the face size illusion where an inverted face is perceived by most individuals as larger than when the same stimulus is rotated upright (Araragi, Aotani, & Kitaoka, 2012; Walsh, Vormberg, Hannaford, & Longo, 2018; Zhang, Wang, & Jiang, 2021). A reverse size illusion is observed for bodies, so that upright bodies are perceived as larger than inverted bodies, and no such

---

* Corresponding author at: Department of Basic & Clinical Neuroscience, Institute of Psychiatry Psychology & Neuroscience, King's College London, UK.
E-mail address: eamonn.walsh@kcl.ac.uk (E. Walsh).

https://doi.org/10.1016/j.cognition.2024.105733
Received 5 June 2023; Received in revised form 11 December 2023; Accepted 21 January 2024

size illusion is observed for hands or everyday objects (Walsh et al., 2018). We previously suggested a possible neural mechanism underlying the illusion (namely, differing activation of neural populations with different receptive field (RF) size for upright and inverted stimuli) (Walsh et al., 2018). It is known that receptive fields are smallest in primary visual cortex (V1) and then progressively increase in size anteriorly through V2, V3/VP, V3A to areas V4 (Smith et al., 2001; Zeki, 1978). Upright stimuli can differentially activate the same, separate or partially overlapping high-level visual areas relative to their inverted counterparts. For example, in an fMRI study Yovel and Kanwisher (2005) showed that the fusiform face area (FFA) showed a greater neural response to upright than inverted faces. Further, this activation correlated with the behavioural face inversion effect, i.e., the disproportionate drop in recognition of upside-down (inverted) stimuli relative to upright faces observed in participants (D. Maurer, Le Grand, & Mondloch, 2002; Yin, 1969). This finding suggests a possible neural mechanism underlying the illusion of size for faces. Activation of different neuronal populations with different receptive field sizes, one for the upright and one for the inverted stimulus, and of differing activation strength, could give rise to a conscious percept of size difference (Moutsiana et al., 2016). We argued (Walsh et al., 2018) that the size illusion and its direction, may depend on the ratio of the RF sizes of the neuronal populations simultaneously activated for upright and inverted stimuli. Contrasting size illusions may reflect differences in how size information is encoded and represented in specialised neural networks, resulting in different perceptual effects.

Such visual illusions are a fruitful way to study perception as, while they deceive us, they also help us tease apart conscious perception from physical stimulation (Schwarzkopf, Song, & Rees, 2011). Featural information perception refers to the characteristics of the individual parts of a face. Configural information relates to the relative spatial arrangements or configurations of these parts and the metric distances between them. Inverting a face results in a substantial performance decrement (Carey & Diamond, 1977; Farah, Tanaka, & Drain, 1995), perhaps due to a failure in spatial integration of information across face features due to stimulus inversion (Poltoratski, Kay, Finzi, & Grill-Spector, 2021), and as demonstrated in face illusions such as the composite illusion (Murphy, Gray, & Cook, 2017), the part-whole illusion (Tanaka & Farah, 1993), the “fat face thin” illusion (Thompson & Wilson, 2012), and the inverted face size illusion (Araragi et al., 2012; Walsh et al., 2018). All these face illusions are thought to derive from holistic processing as higher-order, configural face information (Diamond & Carey, 1986) is disrupted when inverted faces are processed (Leder & Bruce, 2010; Searcy & Bartlett, 1996; however, for a counterargument see Burton, Schweinberger, Jenkins, & Kaufmann, 2015). Holistic processing is a concept central to the study of faces and yet its definition is controversial and can have many meanings (Richler, Palmeri, & Gauthier, 2012). Here, holistic processing refers to the theoretical construct of how individuals perceive a stimulus as a whole, rather than breaking it down into its individual parts. When the stimulus is a face, holistic processing involves integrating information from all parts of the face, emphasising the overall structure, patterns, and relationships between features (Rossion, 2008, 2009). An inverted face is perceived more as a collection of separate features like eyes, nose, or mouth and is less well recognised (Farah et al., 1995). The face inversion effect is thought to be related to holistic processing and suggests that face processing relies on internal representations derived from visual experience (D. Maurer et al., 2002; Rossion, 2009). Notably, face illusions, which are thought to mark holistic processing, occur for the upright but not for the inverted face stimulus (Araragi et al., 2012; McKone et al., 2013; Thompson & Wilson, 2012). Illusions that rely on inversion effects enable us to explore the strength of holistic coding and the processes underlying the various illusions.

While words and faces differ qualitatively, there is evidence that both stimulus categories may involve similar perceptual processes (Martelli, Majaj, & Pelli, 2005). Faces are associated with holistic processing which emphasises detailed spatial relationships between all parts of a face (D. Maurer et al., 2002). Holistic processing is associated with (expert) face perception. However, the relationship between word perception and holistic processing is less clear (Farah, Wilson, Drain, & Tanaka, 1998; Martelli et al., 2005). Holistic processing may have evolved to facilitate recognition and could occur for other objects such as words (Bukach, Gauthier, & Tarr, 2006). Part-based processing could therefore facilitate word perception rather than holistic processes. Some evidence supports part-based processing for words (Farah et al., 1998; Martelli et al., 2005). Farah et al. (1998) proposed that words fall at the part-based processing end of a continuum, with the holistic processing of faces at the other extreme. Other everyday objects, including hands, and bodies, lie somewhere in between. However, expert word perception is also thought to involve, at least partially, holistic processing, as demonstrated by the word superiority effect, whereby people better recognise letters presented within the context of words as compared to isolated letters, or letters presented within nonwords (McClelland & Rumelhart, 1981; Reicher, 1969).

Some theorists have argued for a visual word form area (VWFA), a specific brain region hypothesised for processing alphabetical writing and located in left occipital-temporal cortex (e.g., Nakamura, Dehaene, Jobert, Le Bihan, & Kouider, 2005), including fusiform gyrus, and proposed to process letter shapes and words *prior* to phonology or semantics (Dehaene & Cohen, 2011). Behavioural and neuroimaging studies have found evidence for left lateralised word specificity. Studies have established a left hemisphere dominance for visual word processing (Mercure, Cohen Kadosh, & Johnson, 2011). fMRI studies have shown that the VWFA shows greater selectivity for letters and words (Cohen et al., 2000; Price & Devlin, 2011; Puce, Allison, Asgari, Gore, & McCarthy, 1996). Lesions in the VWFA are associated with reading disorders, but not with face or object recognition, or even general language abilities (Gaillard et al., 2006). The VWFA is thought to be ‘pre-lexical’, i.e., prior to any association with phonology or semantics (Dehaene & Cohen, 2011).

EEG evidence also supports lateralisation effects for words and faces. The N170 component event-related potential (ERP) that signals the neural processing of faces, familiar objects, or words, is consistently left-lateralised for words, while bilateral or right-lateralised for faces (U. Maurer, Rossion, & McCandliss, 2008). The face N170 reduces in amplitude across trials, but this habituation effect does not occur for visual word stimuli (U. Maurer et al., 2008), though a reduced N400 is associated with word repetition (Rugg, 1985), further supporting independent processes for the early perception for words and faces. Behavioural and neuroimaging studies have found evidence for right lateralised face specificity (Dundas, Plaut, & Behrmann, 2013). ERP studies have recorded a stronger N170 component in the RH relative to the LH in response to faces (Scott & Nelson, 2006). The “Fusiform Face Area” (FFA) in right inferior temporal cortex of adults – an area that is located mirror-symmetrical to the Visual Word Form Area, shows greater activation to upright faces compared with other objects (Kanwisher, McDermott, & Chun, 1997).

Face and word recognition can be selectively affected by brain injury or developmental disorders, again suggesting separate processes (Robotham & Starrfelt, 2017). While reading and face recognition deficits can co-occur, in some patients face recognition can be preserved in dyslexia while reading may be preserved in prosopagnosia, signifying independent face and word recognition processes (Robotham & Starrfelt, 2017). Even when there are no differences in reaction time between the dyslexic and control participants for inverted stimuli, people with dyslexia were slower than controls for the recognition of upright word stimuli suggesting that holistic word recognition is impaired in dyslexia (Conway, Brady, & Misra, 2017).

Overall, the evidence suggests that word and face recognition rely on specialised, independent cognitive processes. The human brain allocates mirror symmetrical, largely independent neural systems for the processing of visual words and faces. The processing network for visual words is largely in the left while conversely, the processing of faces shows greater selectivity in the right hemisphere e.g., (Kanwisher et al., 1997). Competition for cortical ‘real estate’ in mid-fusiform regions, may have resulted in faces becoming predominantly right-lateralised (Cantlon, Pinel, Dehaene, & Pelphrey, 2011; Dundas et al., 2013). Thus, left-lateralised visual representation for words may have arisen from right hemispheric specialisation for face processing.

Words are the primary means by which we communicate meaning and ideas, while faces provide important social and emotional cues. By investigating the similarities and differences in how the brain processes these stimuli in the same participants, we can come to better understand the neural mechanisms underlying these cognitive processes. We adopt the size illusion (Araragi et al., 2012) with the method of constant stimuli to measure the bias to perceive inverted stimuli as different in size than upright stimuli. We have used this task previously to demonstrate processing for faces, objects, hands, and bodies. Two experiments are presented. In Experiment 1, as well as replicating the face size illusion, we investigated whether a size illusion is also present for individual familiar Roman and unfamiliar Tamil letters. Experiment 2 extends this investigation to simple visual word and pseudoword stimuli. In both cases, we find clear evidence that the underestimation of the size of upright stimuli is specific to faces.

2. Experiment 1

2.1. Methods

2.1.1. Participants

Forty individuals in the United Kingdom, recruited from the Prolific service (https://www.prolific.co/), participated after giving informed consent. Procedures were approved by the Birkbeck Department of Psychological Sciences Research Ethics Committee. Data from two participants were excluded due to low model fit (see below). The remaining 38 participants (12 men, 36 women) ranged from 18 to 71 years of age (*M*: 35.7 years, *SD*: 13.3 years). Thirty-one were right-handed, and seven left-handed, by self-report. All participants were native English speakers physically located in the United Kingdom, and confirmed they were unfamiliar with the Tamil alphabet. An a priori power analysis using G*Power 3.1 (Faul, Erdfelder, Land, & Buchner, 2007), assuming a medium effect size ($d_z = 0.5$) for the *t*-tests comparing the face condition with the two letter conditions, with power of 0.80 and an alpha of $p =$ .05, indicated that 35 participants were needed. The present study is thus appropriately powered to detect such an effect.

2.1.2. Visual Stimuli

Fig. 1 shows examples of the stimuli used in this experiment. Across blocks, we presented stimuli of three categories: faces, Latin letters, and Tamil letters. Face stimuli were taken from the Karolinska Directed Emotional Faces (Lundqvist, Flykt, & Öhman, 1998). The neutral facial expression for four men and four women was used. Tamil letters were selected as they are also curvilinear and belong to an unrelated family of languages than English.

Letter stimuli were drawn using the *Catamaran* font, by Indian designer Pria Ravichandran, a sans-serif font which is designed to work equally well in both the Latin and Tamil alphabets. Thus, despite their historical differences, letters from both alphabets were presented in a consistent typographic style. We used eight letters of the Latin alphabet (A, B, D, G, K, R, V, and Y) and eight letters of the Tamil alphabet (Latin transliterations from the ISO 15919 standard: e, j, c, n, r, t, ai, ka). In choosing these specific letters, we aimed to ensure that the upright and inverted versions of the letter were distinct and unambiguous.

2.1.3. Procedures

The experiment was conducted online using the Gorilla platform (Anwyl-Irvine, Massonnié, Flitton, Kirkham, & Evershed, 2020). Participants were physically located in the UK and could complete the study using a desktop, laptop, or tablet computer, but were restricted by the platform from using their phone. Other aspects of the study were closely modelled on our previous study using this paradigm (Walsh et al., 2018).

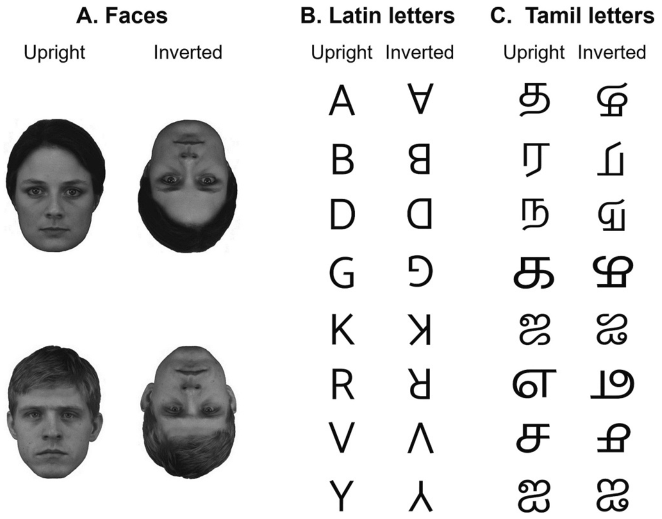

Fig. 1. Examples of visual stimuli used in Experiment 1. There were three categories of stimuli: A. faces (there were 4 male and 4 female faces); (Lundqvist et al., 1998); B. 8 Latin letters, and C. 8 Tamil letters. Stimuli are shown in their upright and inverted orientations.

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 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 remained a constant size. This is a change from our previous study (Walsh et al., 2018) in which the sizes of both upright and inverted stimuli differed across trials, motivated by keeping the experimental design as simple as possible.

Each trial started with a fixation cross which lasted for 500 ms, followed by the simultaneous appearance of the upright and inverted stimuli on either side of the fixation cross. Both stimuli were vertically centred and remained on the screen until the participant responded. The participant's task was to click on the stimulus that appeared physically larger, either using the mouse cursor or their finger (if using a touchscreen or tablet computer).

There were six blocks of trials, two of each of the three stimulus types. The first three blocks included one repetition of each stimulus type, counterbalanced across participants according to a Latin square. The final three blocks were in the reverse order as the first three.

Each block consisted of 56 trials, 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. There was a total of 336 trials and the mean time to complete the task was 12:05 min:*sec* (SD = 04:08). The trials within each block were presented in random order.

Before the start of the experiment, participants completed a short demographics questionnaire, where they were asked to indicate their gender, age, handedness (as reported above), native language, and to indicate whether they had any experience with the Tamil alphabet. All our participants were native English speakers, and all reported having no familiarity with Tamil.

2.1.4. Analysis

The analysis was similar to that in our recent paper (Walsh et al., 2018). For each participant we used the Palamedes toolbox (Prins & Kingdom, 2009) for MATLAB (Mathworks, Natick, MA) to estimate the parameters of the best-fitting cumulative Gaussian curve using maximum-likelihood estimation. For each curve, we obtained: (1) the point of subjective equality (PSE; i.e., the mean of the best-fitting cumulative Gaussian function), (2) the slope of the curve (i.e., the inverse of the standard deviation), and (3) the R^2 value as a measure of goodness-of-fit. The key parameter for assessing the illusion is the PSE, which estimates the difference in size between the inverted and upright stimuli such that the two stimuli are perceived as being the same size. Positive PSEs indicate a bias to perceive upright stimuli as larger than inverted stimuli, and negative PSEs indicate the opposite. The presence of bias for each condition was assessed using one-sample *t*-tests comparing PSEs to 0.

We set goodness-of-fit criteria for inclusion, excluding any participants who had an R^2 value less than 0.50 in any of the three conditions. As mentioned above, data from two participants were excluded from analyses based on this criterion. 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. Stimuli from both experiments and raw data are available on the Open Science Framework: https://osf.io/xg2a7/

2.2. Results and discussion

Results are shown in Fig. 2. There was good overall fit of the psychometric functions to the data, with mean R^2 values of 0.927, 0.939, and 0.926, in the face, Latin, and Tamil conditions, respectively. The R^2 values did not differ significantly across conditions, $F(1.68, 62.03) = 0.548$, $p > .20$, $\eta_p^2 = 0.015$.

To assess the presence of the basic size illusion, we compared PSEs in each condition to 0 (Fig. 2A). In the face condition, there was a clear bias to judge upright faces as smaller than inverted faces (*M*: −1.24%), $t(37) = -4.69$, $p < .0001$, $d = 0.761$ (Fig. 2B, blue bar). This replicates the illusion described in previous studies (Araragi et al., 2012; Walsh et al., 2018). Interestingly, there was a bias in the opposite direction for both the Latin letters (*M*: 2.19%; Fig. 2B, orange bar), $t(37) = 8.71$, $p < .0001$, $d = 1.412$, and the Tamil letters (*M*: 1.81%; Fig. 2B, green bar), $t(37) = 5.38$, $p < .0001$, $d = 0.873$.

An ANOVA on PSEs showed clear differences across conditions, $F(2, 74) = 44.93$, $p < .0001$, $\eta_p^2 = 0.548$. Follow-up *t*-tests showed that the magnitude of bias was significantly different for faces compared to either Latin, $t(37) = 8.96$, $p < .0001$, $d_z = 1.453$, or Tamil, $t(37) = 7.37$, $p < .0001$, $d_z = 1.195$, letters. There was no significant difference between the two alphabets, $t(37) = 0.96$, $p = .343$, $d_z = 0.156$. An ANOVA on slopes revealed a significant difference across conditions, $F(1.61, 59.67) = 5.23$, $p < .05$, $\eta_p^2 = 0.124$. Slopes were significantly higher for faces (*M*: 0.258) than for either Latin letters (*M*: 0.209), $t(37) = 2.46$, $p < .05$, $d_z = 0.399$, or Tamil letters (*M*: 0.207), $t(37) = 2.55$, $p < .05$, $d_z = 0.414$.

These results replicate the basic face size illusion reported in previous research (Araragi et al., 2012; Walsh et al., 2018; Zhang et al., 2021). In our previous study (Walsh et al., 2018), we showed that this illusion was highly-selective for faces, not occurring for other categories including bodies, hands, and everyday objects. The present results provide further evidence for the face-specificity of this effect, showing a reverse illusion for letters.

3. Experiment 2

The *sec*ond experiment used the same logic and design as the first experiment, but compared faces with familiar 3-letter English words and matched pseudowords.

3.1. Methods

3.1.1. Participants

Forty additional individuals recruited from the Prolific service participated after giving informed consent. Data from eight participants were excluded (see below). The remaining 32 participants (13 men, 18 women, 1 who preferred not to indicate sex) ranged from 18 to 74 years of age (*M*: 34.5 years, *SD*: 13.4 years). Twenty-seven were right-handed, three left-handed, and two ambidextrous by self-report. As in Experiment 1, all participants confirmed that they understood written English fluently and had no familiarity with the Tamil alphabet.

The sample size for Experiment 2 was based on the same a priori power analysis as in Experiment 1. However, given the larger number of excluded participants, the final usable sample size is slightly smaller than indicated by that power analysis. Nevertheless, we believe that in

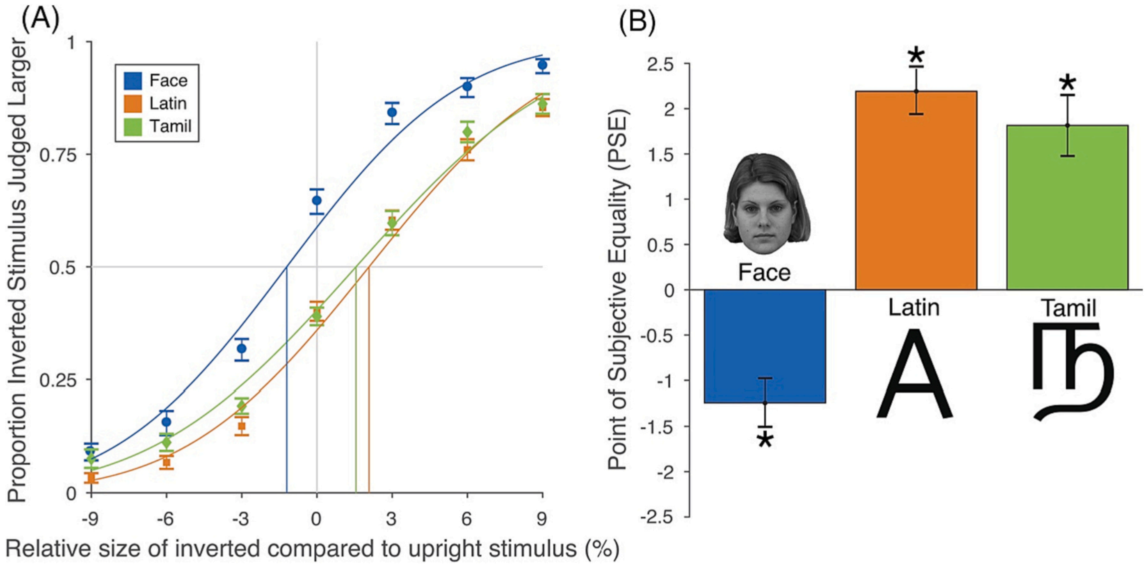

Fig. 2. Results from Experiment 1. A: psychometric functions showing the proportion of trials in which the inverted stimulus was judged as larger as a function of the relative size of the inverted stimulus compared to the upright stimulus. B: mean PSE values for the three stimulus categories. For faces, the PSEs were significantly negative, indicating a bias to perceive faces as smaller when upright than inverted, consistent with previous studies. For both types of letters, in contrast, there was a bias in the opposite direction. Error bars are one standard error. * indicates significant difference from 0 by one-sample t-test.

light of the results of Experiment 1, this experiment remains well-powered. Specifically, the obtained effect sizes for the *t*-tests comparing faces to Latin and Tamil letters in Experiment 1 were more than twice as large as assumed by the power analysis ($d_z = 1.453$ and 1.195, respectively vs. 0.5 for the power analysis). A power analysis using G*Power 3.1 based on the smaller of these effect sizes from Experiment 1 showed that with our final sample size of 32 participants we have power of over 0.999 to detect a comparably sized effect of faces versus words.

3.1.2. Visual Stimuli

Examples of stimuli are shown in Fig. 3. The face stimuli were identical to those used in Experiment 1. For word stimuli, we identified the most frequent 3-letter words in British English as identified by the British National Corpus (Leech, Rayson, & Wilson, 2001). This includes a wide range of both written and spoken sources of British English. We used 8 of the 9 most frequent 3-letter words (the, and, was, for, you, are, not, had). The 8th most frequent word, 'but', was not used as no suitable matched pseudoword was identified.) For each word, we generated a matched English pseudoword (tha, ang, wam, fom, yie, ank, nat, veb) using the Wuggy toolbox (Keuleers & Brysbaert, 2010). As for the letter stimuli in Experiment 1, we presented word and pseudoword stimuli as capital letters in the Catamaran font. The mean time to complete the task (336 trials) was 11:42 min:sec (SD = 02:48).

3.1.3. Procedures

Experimental procedures were the same as in Experiment 1, except that the Latin and Tamil conditions were replaced with the word and pseudoword conditions.

3.1.4. Analysis

Analyses were identical to those in Experiment 1. Six participants were excluded due to R^2 values below 0.5 in at least one condition, and an additional two because PSE values were outside the range of stimulus values presented. This exclusion rate is notably higher than in Experiment 1, for unclear reasons. Two of the excluded participants clicked the upright face on nearly every trial, which presumably reflects a failure to understand (or to have read) the instructions.

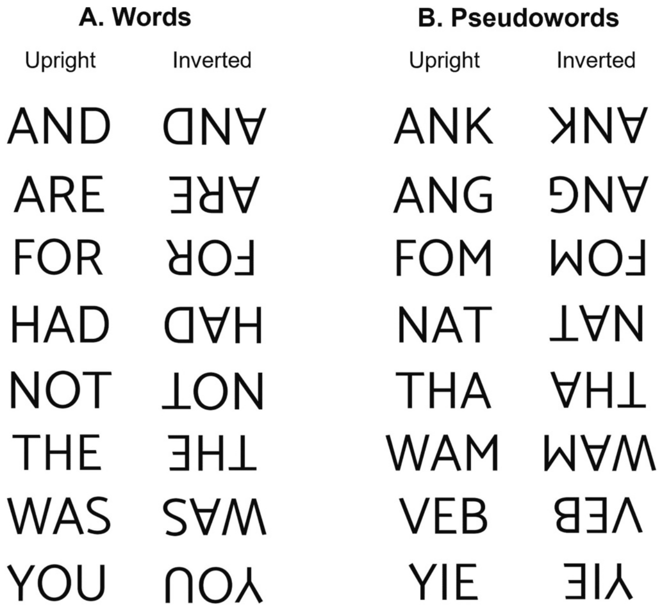

Fig. 3. Experiment 2 stimuli. A. The 8 words and B. The 8 pseudowords. Stimuli are shown in their upright and inverted orientations. The same 8 face stimuli (not shown) were used for both Experiments 1 and 2.

3.2. Results and discussion

Results are shown in Fig. 4. The mean R^2 was 0.940, 0.931, and 0.937, in the face, word, and pseudoword conditions, respectively, indicating good overall fit to the data. The R^2 values did not differ significantly across the three conditions, $F(1.37, 42.30) = 2.07$, $p = .152$, $\eta_p^2 = 0.063$.

In the face condition, there was again a clear bias to judge upright faces as smaller than inverted ones (*M*: −2.00%), $t(31) = -6.50$, $p < .0001$, $d = 1.149$, consistent with Exp 1 and previous studies. There were significant effects in the opposite direction for both the words (*M*: 1.23%), $t(31) = 2.97$, $p < .01$, $d = 0.525$, and the pseudowords (*M*: 0.93%), $t(31) = 3.01$, $p < .01$, $d = 0.532$.

An ANOVA on PSEs showed a clear effect of condition, $F(2, 62) = 30.07$, $p < .0001$, $\eta_p^2 = 0.492$. There were clear differences in PSEs between faces and words, $t(31) = 5.98$, $p < .0001$, $d_z = 1.056$, and between faces and pseudowords, $t(31) = 7.28$, $p < .0001$, $d_z = 1.286$. There was, however, no difference between words and pseudowords, $t(31) = 0.70$, $p = .489$, $d_z = 0.124$.

An ANOVA on slopes revealed significant differences across conditions, $F(2, 62) = 6.81$, $p < .005$, $\eta_p^2 = 0.180$. Slopes were greater for faces (*M*: 0.281%) than either words (*M*: 0.229%), $t(31) = 2.37$, $p < .05$, $d_z = 0.419$, or pseudowords (*M*: 0.213%), $t(31) = 3.83$, $p < .001$, $d_z = 0.677$. There was, however, no difference in slope between words and pseudowords, $t(31) = 0.89$, $p < .380$, $d_z = 0.156$.

4. General discussion

The current study presents two experiments which investigated holistic processing for letters, words and faces using a powerful size illusion. We show robust holistic processing for faces as measured by the size Illusion, replicating previous work (Araragi et al., 2012; Walsh et al., 2018). Inverted faces were perceived as larger relative to the same upright face. In Exp 1, a size illusion was also observed for familiar Roman letters, and unfamiliar Tamil letters, though the direction of the illusion for both was opposite to that for faces i.e., upside-down letters were perceived as smaller (not larger) than the same letter upright. Exp 2 extended this finding to whole words. Both experiments showed that other orthographic stimuli i.e., pseudowords (which have phonology but little or no meaning) and unfamiliar Tamil letters (with no phonology and no meaning) all show the same reverse illusion, as shown previously for bodies (Walsh et al., 2018).

The processing of a single letter by the human brain is complex. First, basic features of the letter, such as its shape, orientation, and contrast are processed (Dehaene & Cohen, 2011; Pegado, Nakamura, Cohen, & Dehaene, 2011). This information is then relayed to higher-order mainly left-hemispheric language centres, which identify the letter as a specific alphabetical symbol (Dehaene & Cohen, 2011; Henry et al., 2005). When we read a word, the brain first analyses its individual letters which are stripped of any font or size characteristics into their abstract form (Cohen et al., 2003; Dehaene et al., 2010; Petersen, Fox, Posner, Mintun, & Raichle, 1988). According to one model of reading, the dual route model (Friedmann & Coltheart, 2018), this processing takes place in the orthographic-visual analysis system. Next, the position of each letter within the word is encoded, and the letters are bound into a word (Coltheart, 1981; Peressotti & Grainger, 1995). The structural output of the orthographic-visual analysis process is held in the orthographic input buffer, which acts as a short-term memory store. The information is subsequently processed via one or two routes: the lexical and sublexical routes. Both the lexical and sublexical routes process known written words. However, the sublexical route only allows the reading of random letter strings, and pseudowords (such as the stimuli presented in Exp. 2). Pseudowords refer to letter strings which though pronounceable, do not actually exist in any specific language and therefore do not activate semantic representations (Mechelli, Gorno-Tempini, & Price, 2003). Similarly, unfamiliar characters such as Tamil (Exp. 2), do not

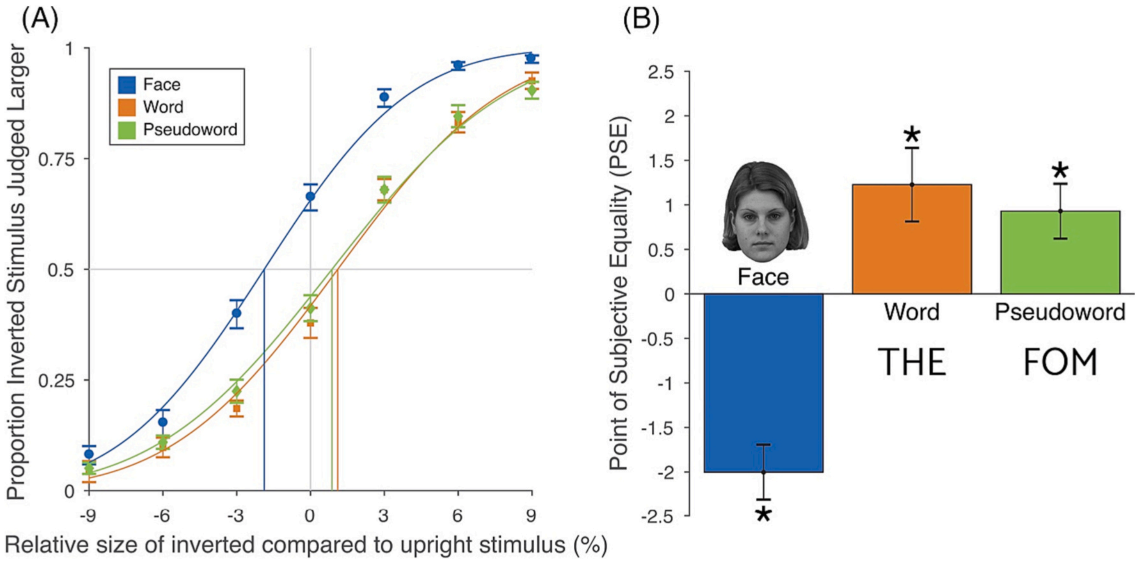

Fig. 4. Results from Experiment 2. A: psychometric functions showing the proportion of trials in which the inverted stimulus was judged as larger as a function of the relative size of the inverted stimulus compared to the upright stimulus. B: mean PSE values for the three stimulus categories. For faces, there was again a bias to perceive faces as smaller when upright than inverted. As for letters in Exp 1, there was a bias in the opposite direction for both words and pseudowords. Error bars are one standard error. * indicates significant difference from 0 by one-sample *t*-test.

activate meaning. In contrast, familiar words activate the lexical route which connects the orthographic input lexicon to the conceptual-semantic system. Once a word has been identified, its meaning is then automatically activated from memory (semantic processing), allowing for the comprehension of the written word. For the expert reader, letter and word recognition of a learned language occur rapidly and automatically.

Known words are processed both lexically and sublexically (Marshall & Newcombe, 1973), while reading pseudowords, which resemble real words orthographically and phonologically (Taylor, Rastle, & Davis, 2013), but possess little or no semantics, are processed sub-lexically (Cassani, Chuang, & Baayen, 2020). Meanwhile, Tamil stimuli have no phonology and little or no meaning for our Tamil-naïve participants. The finding that the illusion is present for both words and pseudowords indicates that the perceptual processes engaged during the illusion are upstream of lexical and sublexical pathways. Thus, the illusion seems to be driven purely by stimulus-driven visual form, rather than higher-level semantics. The illusion does not discriminate between letters, non-familiar Tamil characters (Exp 1), words, and pseudowords (Exp 2), and responds to these visual stimuli which share similar statistical visual regularities, similarly. Our results are consistent with the idea of a general-purpose mechanism that recognises curvilinear shapes found in writing. Writing systems may take advantage of a general human visual system which evolved to scan the environment by selecting and developing characters that fit its own constraints (Morin, 2018). Vertical and horizontal orientation strokes, which mirror basic topological shapes and cardinal orientations found in the natural environment, and which are easier to process by basic visual mechanisms such as neural edge detectors, are evident in letters (Changizi & Shimojo, 2005; Changizi, Zhang, Ye, & Shimojo, 2006). Overall, results indicate cortical processing driving the illusion occurs in early perceptual stages concerned with word form processing, and upstream of brain networks involved in semantics and meaning (Moseley, Pulvermüller, & Shtyrov, 2013).

Alternative models of reading to the dual route model, are provided by "parallel distributed processing" (PDP) models, such as the connectionist triangle model of reading (Harm & Seidenberg, 2004; Plaut, McClelland, Seidenberg, & Patterson, 1996; Woollams, Lambon Ralph, Plaut, & Patterson, 2007). These models incorporate groups of units for the distributed representation of orthography, semantics, and phonology, the three systems which form the 'triangle'. Here, reading is considered a distributed process involving multiple brain regions working in parallel. Visual information about a letter or word is processed by neural networks that represent different aspects of the stimulus' features, including its sound and meaning, which then functionally converge to provide a representation of the whole word, thereby facilitating rapid recognition. Our findings inform such models as the presence of the illusion for both words and non-words is not compatible with a full implementation of the triangle model that includes parallel operation of orthography, semantics and phonology processes (Plaut et al., 1996), as the illusion still generates in the absence of the semantic system.

How can the reverse illusion for faces and orthographic (and body) stimuli be explained? We previously proposed one possible explanation based on receptive field (RF) size (Walsh et al., 2018). Recent evidence for this account is provided by Poltoratski and colleagues (Poltoratski et al., 2021), who used fMRI and population receptive fields (pRF) modelling to quantify spatial processing in face-selective regions. They observed that visual field coverage is smaller and shifted downward in population receptive fields (pRFs), specifically in face-selective regions in response to face inversion. Face inversion disrupts spatial processing of faces (Rossion, 2009; Yin, 1969) through altering pRF outputs and visual field coverage in face-selective regions, resulting in the face inversion effect, the deficit in face recognition for faces presented upside down. Holistic processes rely on spatial processing by receptive fields in face-selective regions (but not primary visual cortex) of the visual system. The finding that face inversion generates smaller pRFs specifically in face-selective regions is important. Larger receptive fields cause the size of an object to appear smaller (Moutsiana et al., 2016), while smaller RFs result in a visual stimulus as appearing larger. Therefore, processing by smaller pRFs should result in the conscious perception of an inverted face stimulus appearing larger.

A similar link between the size of cortical RFs and conscious size perception has been made in the somatosensory system. In primary somatosensory cortex (S1), RFs are smallest on regions of high tactile spatial acuity and larger on less sensitive regions (Sur, Merzenich, & Kaas, 1980), directly analogous to comparisons of the fovea and periphery in primary visual cortex (V1) (Hubel & Wiesel, 1974). Psychophysical studies have found corresponding perceptual biases for the distance between two touches to scale inversely with RF size, an effect generally known as *Weber's illusion* (Cholewiak, 1999; Miller, Longo, &

Saygin, 2016; Weber, 1834). Similarly, tactile RFs on the limbs tend to be oval-shaped rather than circular with the longer axis aligned with the proximo-distal limb axis (Alloway, Rosenthal, & Burton, 1989; V. B. Brooks, Rudomin, & Slayman, 1961; Mountcastle, 1957). This anisotropy in RF geometry is again paralleled by perceptual biases to perceive tactile distances as bigger when oriented across body width than length on several body regions, including the hands (Fiori & Longo, 2018; Longo & Haggard, 2011), arms (Green, 1982; Le Cornu Knight, Longo, & Bremner, 2014), legs (Green, 1982; Tosi & Romano, 2020), feet (Manser-Smith, Tamè, & Longo, 2021), face (Longo, Ghosh, & Yahya, 2015, Longo, Amoruso et al, 2020), and back (Nicula & Longo, 2021). There thus appears to be a systematic link between RF size and tactile distance perception, with smaller RFs associated with larger perceived size. Longo and colleagues (Longo, 2017, 2022; Longo & Haggard, 2011) have proposed a ‘pixel’ model of these effects, suggesting that perceived distance involves a process of counting the number of RF widths between two activation foci within a somatotopic map.

Given the finding that upright visual face stimuli produce larger pRF responses than inverted face stimuli (Poltoratski et al., 2021), extrapolating the logic of the pixel model from the tactile to the visual system would account for the perceptual illusion we find in this study, and which has been described previously (Araragi et al., 2012; Walsh et al., 2018). The human brain may employ similar neural RF anisotropic principles across the different sensory modalities to facilitate multisensory integration. Our theoretical account is parsimonious in that it is based on known findings, i.e., that upright and inverted visual stimuli activate different patterns of cortical neuronal populations, that these have differing RF sizes, and that these in turn can give rise to the conscious experience of differences in relative size.

Face illusions such as the composite, Thatcher and size illusions are commonly used to assess holistic processing (D. Maurer et al., 2002; Tanaka & Farah, 1993; Walsh et al., 2018) and provide a powerful method for investigating general principles of visual perception. In the Thatcher illusion, when the eyes and the mouth are turned upside-down in an upright face, the facial expression is perceived as grotesque. The grotesque effect largely disappears when the composition is inverted and holistic processing is disrupted (Thompson, 1980). Like the “Thatcher illusion, when words are rotated, configural processing is disrupted, arguing for holistic processing for words (Barnhart & Goldinger, 2013). The composite face illusion (Murphy et al., 2017) describes how when the top half of one face is aligned with the bottom half of another face, the resulting composite face results in the perception of a new face due to an illusory fusion of the two aligned halves. The composite face illusion has influenced holistic theories of face perception, which argue that facial features are integrated into a unified, and efficient for recognition purposes, holistic representation (Farah et al., 1998; D. Maurer et al., 2002; Murphy et al., 2017). Composite effects have also been observed for everyday objects such as cars (Bukach, Phillips, & Gauthier, 2010), words (Wong et al., 2011), and Chinese characters (Wong et al., 2012). Wong and colleagues (Wong et al., 2011) found clear evidence of holistic processing for English words, which was proportional to the participant's expertise with the English language. Each illusion may capture different aspects of holistic processing – the Thatcher illusion tells us about configural processing, the composite illusion may better capture attention to multiple parts, while aspects for the face size illusion remain as yet unclear. Testing the same participants on the above three face illusions may allow direct comparison and reveal more about the perceptual expertise required in different aspects of holistic processing.

A strength of our study is that we tested the same participants on both words and faces and used the same size illusion paradigm to compare category-specific holistic processing effects. While results show evidence of holistic processing (as measured by the illusion) for each category, the reverse directionality of these effects suggest different neural loci for each category and also the processing of upright and inverted stimuli (Poltoratski et al., 2021). Future neuroimaging studies could map the neural loci of the size illusion for faces and words, and this may elucidate why and where the illusion and its reverse directionality is occurring. Word recognition engages left occipitotemporal networks (Cohen et al., 2000; Dehaene & Cohen, 2011; Lerma-Usabiaga, Carreiras, & Paz-Alonso, 2018). The left posterior fusiform gyrus responds to words and non-words, regardless of their semantic context (Nobre, Allison, & Mccarthy, 1994). Higher cognitive processing for word meaning occurs in anterior fusiform gyrus. Neuronal populations in left OT cortex respond to complex, high–spatial frequency stimuli such as letters and words, and stimuli with the statistical regularities of words (Vogel, Petersen, & Schlaggar, 2012). The right hemisphere fusiform gyrus is proposed to be more involved in face processing than the left hemisphere. The Visual Word Form Area (VWFA), traditionally viewed as specialised for reading (Dehaene & Cohen, 2011), may also process letters and forms before meaning, as well as lines and shapes found in our environment (Latto, Brain, & Kelly, 2000; Latto & Russell-Duff, 2002; Morin, 2018). While VWFA/left hemisphere specialises for words and the right hemisphere for faces, one hypothesis is that the word and face size illusions are processed in left and right homologues of OT respectively. Identifying the neural loci of the size illusion could be an important step for advancing our understanding of how both categories are processed by the human brain.

Our study also has limitations. In Experiment 1, we selected eight letters from the Roman alphabet, consisting of 1 vowel and 7 consonants (vowel: consonant ratio = 1:7). In English, there are 26 letters, consisting of 5 vowels and 21 consonants (vowel: consonant ratio = 5:21). Also, consonants and vowels differ linguistically, with some theorists arguing both belong to separate linguistic categories and serve different linguistic functions (Mehler, Peña, Nespor, & Bonatti, 2006). Letters such as “O” appear the same even when inverted, and therefore we avoided using them as stimuli here (Exp. 1). Similarly, letters such as “W” resemble a different letter “M” (with a different phonology) when inverted and were also not selected. Our selection was designed to ensure that upright and inverted versions of our stimuli letter were always distinct and unambiguous. However, this selection necessarily means that our letters represent only a subset of the Roman alphabet. Similarly, stimuli in Experiment 2 consisted of monosyllabic 3-letter high frequency words with little semantic meaning. Future experiments should consider polysyllabic words with richer meaning which may potentiate semantic network activation.

Here in both experiments, we replicate a size illusion for faces in which upright stimuli appear *smaller* than inverted ones (Araragi et al., 2012; Walsh et al., 2018). We observed a similar illusion for single letters and word stimuli, and previously for human bodies (Walsh et al., 2018), but interestingly the orientation of the illusion was in the opposite direction (i.e., upright stimulus appears *larger* than inverted). Words and faces are central to how humans communicate. Studying how the human brain processes them increases our understanding of the underlying perceptual and linguistic cognitive processes. Neuroimaging studies are required to further investigate how the different brain regions interact, and how these networks are affected by developmental, neurological, and psychiatric disorders. Though words and faces differ visually, and rely on specialised, independent, cognitive processes, there is both symmetry and asymmetry in how the brain ‘reads’ them.

CRediT authorship contribution statement

Eamonn Walsh: Conceptualization, Writing – original draft. **Carolina Moreira:** Project administration. **Matthew R. Longo:** Conceptualization, Data curation, Formal analysis, Supervision, Writing – review & editing.

Declaration of Competing Interest

We have no known conflict of interest to disclose.

Data availability

Stimuli from both experiments and raw data are available on the Open Science Framework: https://osf.io/xg2a7/

Acknowledgements

'For the purposes of open access, the author has applied a Creative Commons Attribution (CC BY) licence to any Accepted Author Manuscript version arising from this submission.'

References

Alloway, K. D., Rosenthal, P., & Burton, H. (1989). Quantitative measurements of receptive field changes during antagonism of GABAergic transmission in primary somatosensory cortex of cats. *Experimental Brain Research, 78*, 514–532. https://doi.org/10.1007/BF00230239

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*(7), 840–853. https://doi.org/10.1068/p7058

Barnhart, A. S., & Goldinger, S. D. (2013). Rotation reveals the importance of configural cues in handwritten word perception. *Psychonomic Bulletin and Review, 20*, 1319–1326. https://doi.org/10.3758/s13423-013-0435-y

Brooks, K. R., & Kemp, R. I. (2007). Sensitivity to feature displacement in familiar and unfamiliar faces: Beyond the internal/external feature distinction. *Perception, 36*, 1646–1660. https://doi.org/10.1068/p5675

Brooks, V. B., Rudomin, P., & Slayman, C. L. (1961). Peripheral receptive fields of neurons in the cat's cerebral cortex. *Journal of Neurophysiology, 24*, 302–325. https://doi.org/10.1152/jn.1961.24.3.302

Bukach, C. M., Gauthier, I., & Tarr, M. J. (2006). Beyond faces and modularity: The power of an expertise framework. *Trends in Cognitive Sciences, 10*, 159–166. https://doi.org/10.1016/j.tics.2006.02.004

Bukach, C. M., Phillips, W. S., & Gauthier, I. (2010). Limits of generalization between categories and implications for theories of category specificity. *Attention, Perception, & Psychophysics, 72*, 1865–1874. https://doi.org/10.3758/APP.72.7.1865

Burton, A. M., Schweinberger, S. R., Jenkins, R., & Kaufmann, J. M. (2015). Arguments against a configural processing account of familiar face recognition. *Perspectives on Psychological Science, 10*, 482–496.

Cantlon, J. F., Pinel, P., Dehaene, S., & Pelphrey, K. A. (2011). Cortical representations of symbols, objects, and faces are pruned back during early childhood. *Cerebral Cortex, 21*, 191–199. https://doi.org/10.1093/cercor/bhq078

Carey, S., & Diamond, R. (1977). From piecemeal to configural representation of faces. *Science, 195*, 312–314.

Cassani, G., Chuang, Y.-Y., & Baayen, R. H. (2020). On the semantics of nonwords and their lexical category. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 46*, 621–637.

Changizi, M. A., & Shimojo, S. (2005). Character complexity and redundancy in writing systems over human history. *Proceedings of the Royal Society B, 272*, 267–275. https://doi.org/10.1098/rspb.2004.2942

Changizi, M. A., Zhang, Q., Ye, H., & Shimojo, S. (2006). The structures of letters and symbols throughout human history are selected to match those found in objects in natural scenes. *American Naturalist, 167*, E117–E139. https://doi.org/10.1086/502806

Cholewiak, R. W. (1999). The perception of tactile distance: Influences of body site, space, and time. *Perception, 28*, 851–876. https://doi.org/10.1068/p2873

Cohen, L., Dehaene, S., Naccache, L., Lehéricy, S., Dehaene-Lambertz, G., Hénaff, M.-A., & 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.

Cohen, L., Martinaud, O., Lemer, C., Lehéricy, S., Samson, Y., Obadia, M., Slachevsky, A., & Dehaene, S. (2003). Visual word recognition in the left and right hemispheres: Anatomical and functional correlates of peripheral alexias. *Cerebral Cortex, 13*, 1313–1333. https://doi.org/10.1093/cercor/bhg079

Coltheart, M. (1981). Disorders of reading and their implications for models of normal reading. *Visible Language, 15*, 245–286.

Conway, A., Brady, N., & Misra, K. (2017). Holistic word processing in dyslexia. *PLoS One, 12*, Article e0187326. https://doi.org/10.6084/m9.figshare.5540014

Dediu, D., & Levinson, S. C. (2013). On the antiquity of language: The reinterpretation of Neandertal linguistic capacities and its consequences. *Frontiers in Psychology, 4*, 397. https://doi.org/10.3389/fpsyg.2013.00397

Dehaene, S., & Cohen, L. (2011). The unique role of the visual word form area in reading. *Trends in Cognitive Sciences, 15*, 254–262. https://doi.org/10.1016/j.tics.2011.04.003

Dehaene, S., Pegado, F., Braga, L. W., Ventura, P., Filho, G. N., Jobert, A., … Cohen, L. (2010). How learning to read changes the cortical networks for vision and language. *Science, 330*, 1359–1364. https://doi.org/10.1126/science.1194140

Diamond, R., & Carey, S. (1986). Why faces are and are not special: An effect of expertise. *Journal of Experimental Psychology: General, 115*, 107–117. https://doi.org/10.1037//0096-3445.115.2.107

Dundas, E. M., Plaut, D. C., & Behrmann, M. (2013). The joint development of hemispheric lateralization for words and faces. *Journal of Experimental Psychology: General, 142*, 348–358. https://doi.org/10.1037/a0029503

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.

Farah, M. J., Wilson, K. D., Drain, M., & Tanaka, J. N. (1998). What is 'special' about face perception? *Psychological Review, 105*, 482–498.

Faul, F., Erdfelder, E., Land, 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

Fiori, F., & Longo, M. R. (2018). Tactile distance illusions reflect a coherent stretch of tactile space. *Proceedings of the National Academy of Sciences, 115*, 1238–1243. https://doi.org/10.1073/pnas.1715123115

Friedmann, N., & Coltheart, M. (2018). Types of developmental dyslexia. In *Handbook of communication disorders: Theoretical, empirical, and applied linguistics perspectives* (pp. 721–751). De Gruyter Mouton.

Gaillard, R., Naccache, L., Pinel, P., Clémenceau, S., Volle, E., Hasboun, D., Dupont, S., Baulac, M., Dehaene, S., Adam, C., & Cohen, L. (2006). Direct intracranial, fMRI, and lesion evidence for the causal role of left inferotemporal cortex in reading. *Neuron, 50*, 191–204. https://doi.org/10.1016/j.neuron.2006.03.031

Glassner, J.-J., & Herron, D. M. (2003). *The invention of cuneiform: Writing in Sumer*. Johns Hopkins University Press.

Green, B. G. (1982). The perception of distance and location for dual tactile pressures. *Perception & Psychophysics, 31*, 315–323.

Harm, M. W., & Seidenberg, M. S. (2004). Computing the meanings of words in reading: Cooperative division of labor between visual and phonological processes. *Psychological Review, 111*, 662–720. https://doi.org/10.1037/0033-295X.111.3.662

Henry, C., Gaillard, R., Volle, E., Chiras, J., Ferrieux, S., Dehaene, S., & Cohen, L. (2005). Brain activations during letter-by-letter reading: A follow-up study. *Neuropsychologia, 43*, 1983–1989. https://doi.org/10.1016/j.neuropsychologia.2005.04.007

Heyes, C. (2018). *Cognitive gadgets: The cultural evolution of thinking*. Belknap Press.

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*, 295–306. https://doi.org/10.1002/cne.901580305

Kanwisher, N., McDermott, J., & Chun, M. M. (1997). The fusiform face area: A module in human extrastriate cortex specialized for face perception. *Journal of Neuroscience, 17*, 4302–4311.

Keuleers, E., & Brysbaert, M. (2010). Wuggy: A multilingual pseudoword generator. *Behavior Research Methods, 42*, 627–633. https://doi.org/10.3758/BRM.42.3.627

Lacruz, R. S., Stringer, C. B., Kimbel, W. H., Wood, B., Harvati, K., O'Higgins, P., … Arsuaga, J.-L. (2019). The evolutionary history of the human face. *Nature Ecology & Evolution, 3*, 726–736. https://doi.org/10.1038/s41559-019-0865-7

Latto, R., Brain, D., & Kelly, B. (2000). An oblique effect in aesthetics: Homage to Mondrian (1872—1944). *Perception, 29*, 981–987. https://doi.org/10.1068/p2352

Latto, R., & Russell-Duff, K. (2002). An oblique effect in the selection of line orientation by twentieth century painters. *Empirical Studies of the Arts, 20*, 49–60.

Le Cornu Knight, F., Longo, M. R., & Bremner, A. J. (2014). Categorical perception of tactile distance. *Cognition, 131*, 254–262. https://doi.org/10.1016/j.cognition.2014.01.005

Leder, H., & Bruce, V. (2010). When inverted faces are recognized: The role of configural information in face recognition. *Quarterly Journal of Experimental Psychology, 53A*, 513–536.

Leech, G., Rayson, P., & Wilson, A. (2001). *Word frequencies in written and spoken English, based on the British National Corpus*. Longman.

Lerma-Usabiaga, G., Carreiras, M., & Paz-Alonso, P. M. (2018). Converging evidence for functional and structural segregation within the left ventral occipitotemporal cortex in reading. *Proceedings of the National Academy of Sciences, 115*, E9981–E9990. https://doi.org/10.1073/pnas.1803003115

Longo, M. R. (2017). Distorted body representations in healthy cognition. *Quarterly Journal of Experimental Psychology, 70*, 378–388. https://doi.org/10.1080/17470218.2016.1143956

Longo, M. R. (2022). Distortion of mental body representations. *Trends in Cognitive Sciences, 26*, 241–254. https://doi.org/10.1016/j.tics.2021.11.005

Longo, M. R., Amoruso, E., Calzolari, E., Ben Yehuda, M., Haggard, P., & Azañón, E. (2020). Anisotropies of tactile distance perception on the face. *Attention, Perception, & Psychophysics, 82*, 3636–3647. https://doi.org/10.3758/s13414-020-02079-y

Longo, M. R., Ghosh, A., & Yahya, T. (2015). Bilateral symmetry of distortions of tactile size perception. *Perception, 44*, 1251–1262. https://doi.org/10.1177/0301006615594949

Longo, M. R., & Haggard, P. (2011). Weber's illusion and body shape: Anisotropy of tactile size perception on the hand. *Journal of Experimental Psychology: Human Perception and Performance, 37*, 720–726. https://doi.org/10.1037/a0021921

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.

Manser-Smith, K., Tamè, L., & Longo, M. R. (2021). Tactile distance anisotropy on the feet. *Attention, Perception, & Psychophysics, 83*, 3227–3239. https://doi.org/10.3758/s13414-021-02339-5

Marshall, J. C., & Newcombe, F. (1973). Patterns of paralexia: A psycholinguistic approach. *Journal of Psycholinguistic Research, 2*, 175–199.

Martelli, M., Majaj, N. J., & Pelli, D. G. (2005). Are faces processed like words? A diagnostic test for recognition by parts. *Journal of Vision, 5*, 58–70. https://doi.org/10.1167/5.1.6

Maurer, D., Le Grand, R., & 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

Maurer, U., Rossion, B., & McCandliss, B. D. (2008). Category specificity in early perception: Face and word N170 responses differ in both lateralization and habituation properties. *Frontiers in Human Neuroscience, 2*, 18. https://doi.org/10.3389/neuro.09.018.2008

McClelland, J. L., & Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception: Part 1. An account of basic findings. *Psychological Review, 88*, 375–407.

McKone, E., Aimola Davies, A., Darke, H., Crookes, K., Wickramariyaratne, T., Zappia, S., Fiorentini, C., Favelle, S., Broughton, M., & Fernando, D. (2013). Importance of the inverted control in measuring holistic face processing with the composite effect and part-whole effect. *Frontiers in Psychology, 4*, 33. https://doi.org/10.3389/fpsyg.2013.00033

Mechelli, A., Gorno-Tempini, M. L., & Price, C. J. (2003). Neuroimaging studies of word and pseudoword reading: Consistencies, inconsistencies, and limitations. *Journal of Cognitive Neuroscience, 15*, 260–271. https://doi.org/10.1162/089892903321208196

Mehler, J., Peña, M., Nespor, M., & Bonatti, L. (2006). The "soul" of language does not use statistics: Reflections on vowels and consonants. *Cortex, 42*, 846–854. https://doi.org/10.1016/S0010-9452(08)70427-1

Mercure, E., Cohen Kadosh, K., & Johnson, M. H. (2011). The N170 shows differential repetition effects for faces, objects, and orthographic stimuli. *Frontiers in Human Neuroscience, 5*, 6. https://doi.org/10.3389/fnhum.2011.00006

Miller, L. E., Longo, M. R., & Saygin, A. P. (2016). Mental body representations retain homuncular shape distortions: Evidence from Weber's illusion. *Consciousness and Cognition, 40*, 17–25. https://doi.org/10.1016/j.concog.2015.12.008

Mondloch, C. J., Le Grand, R., & Maurer, D. (2002). Configural face processing develops more slowly than featural face processing. *Perception, 31*, 553–566. https://doi.org/10.1068/p3339

Morin, O. (2018). Spontaneous emergence of legibility in writing systems: The case of orientation anisotropy. *Cognitive Science, 42*, 664–677. https://doi.org/10.1111/cogs.12550

Moseley, R. L., Pulvermüller, F., & Shtyrov, Y. (2013). Sensorimotor semantics on the spot: Brain activity dissociates between conceptual categories within 150 ms. *Scientific Reports, 3*, 1928. https://doi.org/10.1038/srep01928

Mountcastle, V. B. (1957). Modality and topographic properties of single neurons of cat's somatic sensory cortex. *Journal of Neurophysiology, 20*, 408–434. https://doi.org/10.1152/jn.1957.20.4.408

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*, 1–12. https://doi.org/10.1038/ncomms12110

Murphy, J., Gray, K. L. H., & Cook, R. (2017). The composite face illusion. *Psychonomic Bulletin and Review, 24*, 245–261. https://doi.org/10.3758/s13423-016-1131-5

Nakamura, K., Dehaene, S., Jobert, A., Le Bihan, D., & Kouider, S. (2005). Subliminal convergence of kanji and kana words: Further evidence for functional parcellation of the posterior temporal cortex in visual word perception. *Journal of Cognitive Neuroscience, 17*, 954–968.

Nicula, A., & Longo, M. R. (2021). Perception of tactile distance on the back. *Perception, 50*, 677–689. https://doi.org/10.1177/03010066211025384

Nobre, A. C., Allison, T., & Mccarthy, G. (1994). Word recognition in the human inferior temporal lobe. *Nature, 372*, 260–263.

Pegado, F., Nakamura, K., Cohen, L., & Dehaene, S. (2011). Breaking the symmetry: Mirror discrimination for single letters but not for pictures in the visual word form area. *NeuroImage, 55*, 742–749. https://doi.org/10.1016/j.neuroimage.2010.11.043

Peressotti, F., & Grainger, J. (1995). Letter-position coding in random consonant arrays. *Perception & Psychophysics, 57*, 875–890.

Perreault, C., & Mathew, S. (2012). Dating the origin of language using phonemic diversity. *PLoS One, 7*, Article e35289. https://doi.org/10.1371/journal.pone.0035289

Petersen, S. E., Fox, P. T., Posner, M. I., Mintun, M., & Raichle, M. E. (1988). Positron emission tomographic studies of the cortical anatomy of single-word processing. *Nature, 331*, 585–589. https://doi.org/10.1038/331585a0

Plaut, D. C., McClelland, J. L., Seidenberg, M. S., & Patterson, K. (1996). Understanding normal and impaired word reading: Computational principles in quasi-regular domains. *Psychological Review, 103*, 56–115. https://doi.org/10.1037/0033-295x.103.1.56

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

Price, C. J., & Devlin, J. T. (2011). The interactive account of ventral occipitotemporal contributions to reading. *Trends in Cognitive Sciences, 15*, 246–253. https://doi.org/10.1016/j.tics.2011.04.001

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

Puce, A., Allison, T., Asgari, M., Gore, J. C., & McCarthy, G. (1996). Differential sensitivity of human visual cortex to faces, letterstrings, and textures: A functional magnetic resonance imaging study. *Journal of Neuroscience, 16*, 5205–5215.

Reicher, G. M. (1969). Perceptual recognition as a function of meaningfulness of stimulus material. *Journal of Experimental Psychology, 81*, 275–280. https://doi.org/10.1037/h0027768

Richler, J. J., Palmeri, T. J., & Gauthier, I. (2012). Meanings, mechanisms, and measures of holistic processing. *Frontiers in Psychology, 3*, 553. https://doi.org/10.3389/fpsyg.2012.00553

Robotham, R. J., & Starrfelt, R. (2017). Face and word recognition can be selectively affected by brain injury or developmental disorders. *Frontiers in Psychology, 8*, 1547. https://doi.org/10.3389/fpsyg.2017.01547

Rossion, B. (2008). Picture-plane inversion leads to qualitative changes of face perception. *Acta Psychologica, 128*, 274–289. https://doi.org/10.1016/j.actpsy.2008.02.003

Rossion, B. (2009). Distinguishing the cause and consequence of face inversion: The perceptual field hypothesis. *Acta Psychologica, 132*, 300–312. https://doi.org/10.1016/j.actpsy.2009.08.002

Rugg, M. D. (1985). The effects of semantic priming and word repetition on event-related potentials. *Psychophysiology, 22*, 642–647.

Schwarzkopf, D. S., Song, C., & Rees, G. (2011). The surface area of human V1 predicts the subjective experience of object size. *Nature Neuroscience, 14*, 28–30. https://doi.org/10.1038/nn.2706

Scott, L. S., & Nelson, C. A. (2006). Featural and configural face processing in adults and infants: A behavioral and electrophysiological investigation. *Perception, 35*, 1107–1128. https://doi.org/10.1068/p5493

Searcy, J. H., & Bartlett, J. C. (1996). Inversion and processing of component and spatial-relational information in faces. *Journal of Experimental Psychology: Human Perception and Performance, 22*, 904–915.

Sereno, S. C., & Rayner, K. (2003). Measuring word recognition in reading: Eye movements and event-related potentials. *Trends in Cognitive Sciences, 7*, 489–493. https://doi.org/10.1016/j.tics.2003.09.010

Sheehan, M. J., & Nachman, M. W. (2014). Morphological and population genomic evidence that human faces have evolved to signal individual identity. *Nature Communications, 5*, 4800. https://doi.org/10.1038/ncomms5800

Sur, M., Merzenich, M. M., & Kaas, J. H. (1980). Magnification, receptive-field area, and size in areas 3b and 1 of somatosensory cortex in owl monkeys. *Journal of Neurophysiology, 44*, 295–311. https://doi.org/10.1152/jn.1980.44.2.295

Tanaka, J. W., & Farah, M. J. (1993). Parts and wholes in face recognition. *Quarterly Journal of Experimental Psychology, 46*, 225–245. https://doi.org/10.1080/14640749308401045

Taylor, J. S. H., Rastle, K., & Davis, M. H. (2013). Can cognitive models explain brain activation during word and pseudoword reading? A meta-analysis of 36 neuroimaging studies. *Psychological Bulletin, 139*, 766–791.

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

Thompson, P., & Wilson, J. (2012). Why do most faces look thinner upside down? *I-Perception, 3*, 765–774. https://doi.org/10.1068/i0554

Tosi, G., & Romano, D. (2020). The longer the reference, the shorter the legs: How response modality affects body perception. *Attention, Perception, & Psychophysics, 82*, 3737–3749. https://doi.org/10.3758/s13414-020-02074-3 The

Vogel, A. C., Petersen, S. E., & Schlaggar, B. L. (2012). The left occipitotemporal cortex does not show preferential activity for words. *Cerebral Cortex, 22*, 2715–2732. https://doi.org/10.1093/cercor/bhr295

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

Weber, E. H. (1834). De subtilitate tactus. In H. E. Ross, & D. J. Murray (Eds.), *E. H. Weber on the tactile senses* (pp. 21–128). Academic Press.

Wong, A. C., Bukach, C. M., Hsiao, J., Greenspon, E., Ahern, E., Duan, Y., & Lui, K. F. H. (2012). Holistic processing as a hallmark of perceptual expertise for nonface categories including Chinese characters. *Journal of Vision, 12*, 7. https://doi.org/10.1167/12.13.7.Introduction

Wong, A. C., Bukach, C. M., Yuen, C., Yang, L., Leung, S., & Greenspon, E. (2011). Holistic processing of words modulated by reading experience. *PLoS One, 6*, Article e20753. https://doi.org/10.1371/journal.pone.0020753

Wong, A. C., & Gauthier, I. (2007). An analysis of letter expertise in a levels-of-categorization framework. *Visual Cognition, 15*, 854–879. https://doi.org/10.1080/13506280600948350

Woollams, A. M., Lambon Ralph, M. A., Plaut, D. C., & Patterson, K. (2007). SD-squared: On the association between semantic dementia and surface dyslexia. *Psychological Review, 114*, 316–339. https://doi.org/10.1037/0033-295X.114.2.316

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*, Article 104718. https://doi.org/10.1016/j.cognition.2021.104718