

Monocular Advantage for Face Perception Implicates Subcortical Mechanisms in Adult Humans

Shai Gabay, Adrian Nestor, Eva Dundas, and Marlene Behrmann

Abstract

■ The ability to recognize faces accurately and rapidly is an evolutionarily adaptive process. Most studies examining the neural correlates of face perception in adult humans have focused on a distributed cortical network of face-selective regions. There is, however, robust evidence from phylogenetic and ontogenetic studies that implicates subcortical structures, and recently, some investigations in adult humans indicate subcortical correlates of face perception as well. The questions addressed here are whether low-level subcortical mechanisms for face perception (in the absence of changes in expression) are conserved in human adults, and if so, what is the nature of these subcortical representations. In a series of four experiments, we presented pairs of images to the same or different eyes. Participants' performance demonstrated that subcortical mechanisms, indexed by monocular portions of the visual sys-

tem, play a functional role in face perception. These mechanisms are sensitive to face-like configurations and afford a coarse representation of a face, comprised of primarily low spatial frequency information, which suffices for matching faces but not for more complex aspects of face perception such as sex differentiation. Importantly, these subcortical mechanisms are not implicated in the perception of other visual stimuli, such as cars or letter strings. These findings suggest a conservation of phylogenetically and ontogenetically lower-order systems in adult human face perception. The involvement of subcortical structures in face recognition provokes a reconsideration of current theories of face perception, which are reliant on cortical level processing, inasmuch as it bolsters the cross-species continuity of the biological system for face recognition. ■

INTRODUCTION

A distributed network of face-selective cortical regions has been uncovered in many studies examining face perception in adult humans (e.g., Avidan & Behrmann, 2009; Fairhall & Ishai, 2007; Haxby, Hoffman, & Gobbini, 2000) and in nonhuman primates (e.g., Liu et al., 2013; Bell, Hadj-Bouziane, Frihauf, Tootell, & Ungerleider, 2009; Tsao, Moeller, & Freiwald, 2008). Although most studies of the neural correlates of face recognition have focused on these cortical loci, much less attention has been devoted to exploring the potential contribution of lower-order structures (Johnson, 2005). There exists, however, phylogenetic evidence that indicates that the ability to discriminate kin from nonkin is ubiquitous even in species with rudimentary brain structures, such as wasps (Sheehan & Tibbetts, 2011; Tibbetts, 2002), and honeybees (Dyer, Neumeyer, & Chittka, 2005). And, along similar lines, some neuroimaging studies in nonhuman primates have detected activation of lower-order, subcortical structures when monkeys view images of monkey faces and bodies compared with images of their scrambled counterparts (Logothetis, Guggenberger, Peled, & Pauls, 1999). A subsequent high-resolution imaging study in nonhuman

primates has even succeeded in uncovering separable activations of subnuclei within the amygdala in response to faces (Hoffman, Gothard, Schmid, & Logothetis, 2007).

Ontogenetic evidence also indicates a contribution from more rudimentary neural structures to face perception: Even with a rather immature neural system, newborn human infants are able to discriminate faces, an ability attributed to a primitive subcortical bias to orient toward face-like patterns with relevant configural information (Johnson, Dziurawiec, Ellis, & Morton, 1991; Johnson & Morton, 1991). Consistent with this, under monocular viewing, infants preferentially orient to images resembling faces to a greater extent in the temporal compared with nasal hemifield (Simion, Valenza, Umiltà, & Barba, 1998), a result indicative of retinotectal mediation (Williams, Azzopardi, & Cowey, 1995). Despite these findings implicating more rudimentary neural structures in face perception, evidence for the contribution of such structures in adult humans is rather sparse. One possibility is that, as one ascends the phylogenetic and ontogenetic scale, cortical structures, alone, are functionally implicated in face perception. Alternatively, although cortical contributions may predominate, subcortical structures may still continue to play a role even in adulthood. Here, we examine whether adult humans engage subcortical structures for the purpose of face perception, and if so, what is the nature of this subcortical contribution.

Subcortical Structures and Face Perception in Humans

One obvious reason why there might be a paucity of evidence for the engagement of subcortical structures in adult humans is that there is a fundamental methodological limitation: Access to these structures is extremely difficult. The subcortical structures of interest are small, and their size and location make it difficult to image because of a reduction in signal-to-noise ratio relative to cortical regions (LaBar, Gitelman, Mesulam, & Parrish, 2001). Also, because functional localizers typically adopt rather stringent thresholds to identify selective ROIs, regions with lower signal, such as the subcortical areas, fail to be differentiated.

Unsurprisingly, then, rather few studies with adult humans have provided direct evidence for face selectivity in subcortical regions. The studies that have reported activation of subcortical structures have typically focused on responses of subcortical regions to affective faces (perhaps the most salient inputs), and in particular, the activation of the amygdala to emotional faces (especially those with negative valence) has been well established (Todorov, Mende-Siedlecki, & Dotsch, 2013; Troiani & Schultz, 2013; Todorov, 2012; Pessoa & Adolphs, 2010; Goossens et al., 2009). With one exception, the neuroimaging studies have not explored the response of subcortical regions to faces in the absence of emotional expression; in this study, using fMRI data from 215 participants, Mende-Siedlecki, Verosky, Turk-Browne, and Todorov (2013) were able to detect robust and reliable responses to neutral faces in the amygdala bilaterally and observed strong functional coupling between the amygdala and posterior face-selective regions (such as FFA). Although the major emphasis of this study is on the amygdala, face-selective responses were also noted in the superior colliculus and hippocampus (see also Ishai & Yago, 2006). The results from this large-scale study indicate that, when methodology permits, a substantial contribution from subcortical structures to face perception in adult humans can be uncovered. What remains uncertain from this finding is whether these activated structures contribute functionally to face perception and, moreover, what type of representation might subservise face perception in these lower structures.

There is also some indication from electrophysiological studies that subcortical regions might be involved in face perception, but again, this is primarily in response to emotional or motivational cues. It is notoriously difficult to ascribe the source of the waveform sampled at the scalp surface to an underlying neural structure, and even the temporal component cannot help resolve the question of the underlying generator. For example, the well-known ERP N170 marker (Bentin, Allison, Puce, Perez, & McCarthy, 1996; Jeffreys, 1996) is considered too late to be generated by subcortical mechanisms, and the same is likely true for ERP components manifest as early as 100 msec (Herrmann, Ehlis, Ellgring, & Fallgatter, 2005; Liu, Harris, & Kanwisher,

2002). Inferences can be made, however, regarding modulation of cortical signals by subcortical structures; for example, Sabatinelli, Keil, Frank, and Lang (2013) argue that the fact that the centroparietal late positive potential showed covariation with subcortical and corticolimbic structures suggests that these regions may contribute to discrimination of emotional state of faces (but no evidence for identity discrimination appears to exist). It is likely that, just as it is challenging to probe the subcortical signal using fMRI, it may be similarly difficult to sample electrophysiological data from subcortical structures remotely at the scalp surface.

In this study, we exploit a psychophysical technique, which circumvents the above-noted limitations, to examine the nature and role of subcortical mechanisms for face perception in human adults.

Our Approach

The technique we adopt takes advantage of the fact that visual input, once received by the retina, is propagated in an eye-specific fashion through the early stages of the visual system. This monocular segregation is retained up to layer IV of striate cortex (Menon, Ogawa, Strupp, & Uğurbil, 1997; Horton, Dagi, McCrane, & de Monasterio, 1990). Because there are relatively few monocular neurons beyond area V1 (Bi et al., 2011), activation of extrastriate areas is not eye dependent (see Figure 1). Given that

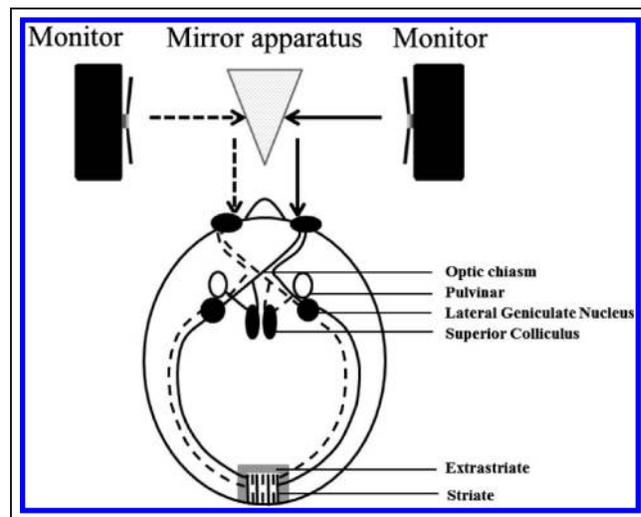
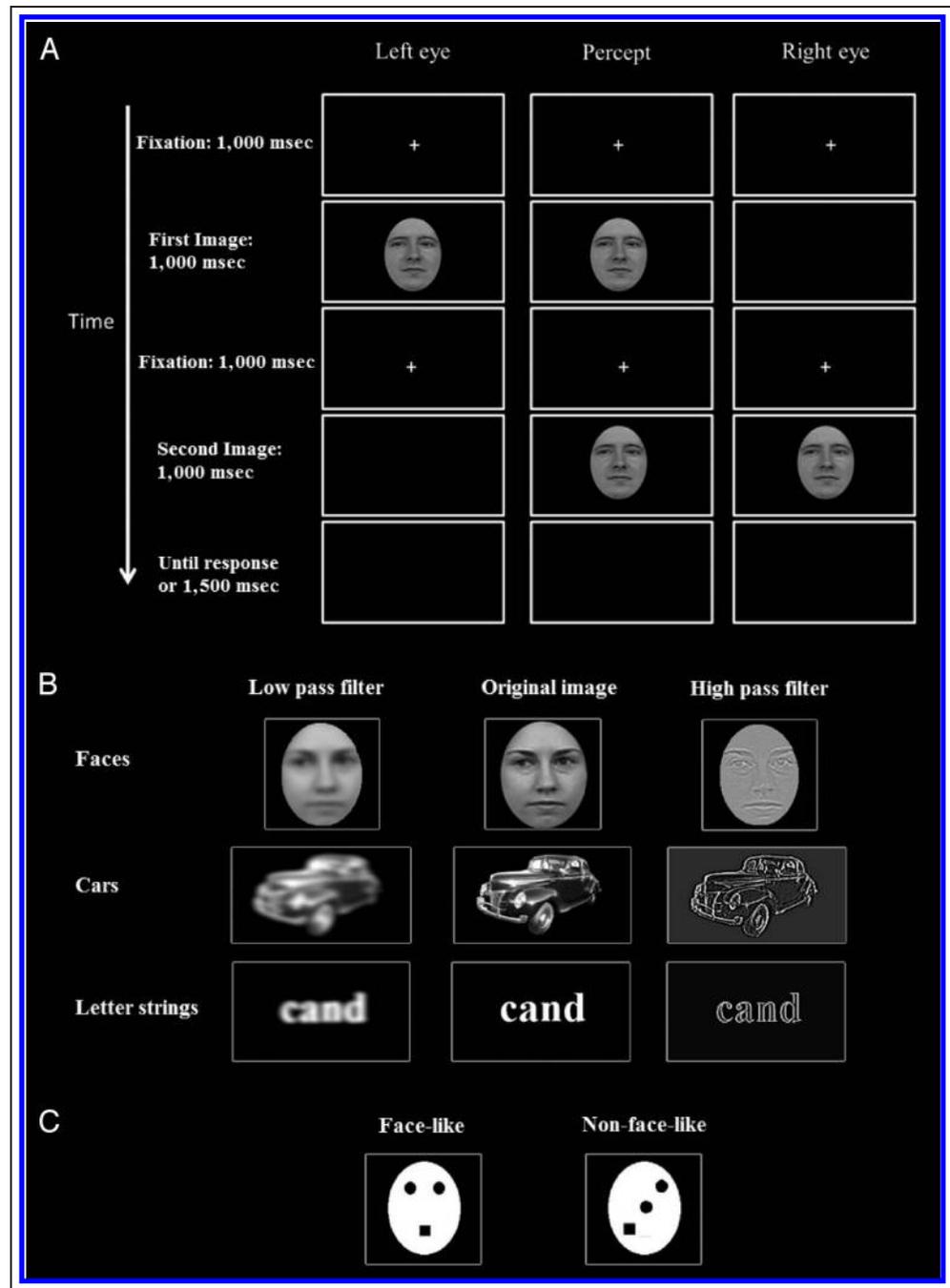


Figure 1. A schematic depiction of the experimental apparatus and visual pathways from the eyes to the brain (shown in axial plane). Each monitor provided visual information to a different eye. The visual information first passes through monocularly segregated subcortical regions (left eye, dashed lines; right eye, solid lines), which is then projected to the pulvinar, LGN, and superior colliculus en route to the striate and then binocular extrastriate regions. Note that we have excluded the amygdala from this schematic depiction as the focus is on face (and car and letter string) perception rather than on perception of facial emotional expression. For simplicity, we depict only the input from the contralateral eye to each superior colliculus.

Figure 2. (A) A typical different-eye trial in which the first image is presented to the left eye (left column) and the second image is presented to the right eye (right column). The middle column represents the participant's fused perception. A "same" response is required. (B) An example of a face, car, and letter string stimulus in the low-pass filter, original, and high-pass filter condition. (C) An example of face-like and non-face-like images.



observers are not explicitly aware of the eye to which a visual stimulus is projected (Schwarzkopf, Schindler, & Rees, 2010; Blake & Cormack, 1979) and perceive the images from different eyes as “fused” (see Figure 2), manipulating the eye of origin of the stimulus provides a useful tool for isolating monocular versus binocular neural channels. Thus, the logic of our studies is as follows: If perceptual performance is enhanced when two images are presented sequentially to a single eye versus interocularly to different eyes, we can infer that the monocular advantage is a product of neural facilitation within lower levels of the visual pathway. This technique has been used

successfully in the past to examine plasticity in transferring perceptual learning from one eye to another (Karni & Sagi, 1991), examination of spatial attention (Self & Roelfsema, 2010), and multisensory perception (Batson, Beer, Seitz, & Watanabe, 2011).

Across four studies, we examined whether there is a monocular advantage for face perception (better discrimination when both faces are presented to a single compared with two eyes) versus other visual categories and then further elaborated the type of representation that might be exploited by the subcortical structures using four separate manipulations. The first involves a comparison of the

monocular advantage for upright versus inverted faces. The behavioral and BOLD activation advantage for perceiving upright versus inverted faces is considered a signature of cortical activation (Yovel & Kanwisher, 2005; Rossion & Gauthier, 2002). Thus, if the subcortical representation is less elaborated than the cortical substrate, then we might expect to see an equivalent monocular advantage for both upright and inverted faces (i.e., a more rudimentary mechanism that is not tuned to orientation). If, however, the subcortical system represents faces in the same way as cortical systems, then the monocular advantage should be greater for upright faces than for inverted faces.

The second manipulation examines whether the monocular advantage is present for more complex face perception tasks. Determining whether a face belongs to the male or female sex requires a more sophisticated representation than that required for simply judging whether two faces are the same or not and indicating the sex of a face engages an extensive network of cortical regions (Kaul, Rees, & Ishai, 2010). A monocular advantage might not be evident under these more taxing task demands if the subcortical representations are indeed rather more elementary compared with cortical representations.

The third manipulation entails a contrast between the monocular advantage for high versus low spatial frequency face images. Subcortical structures are notoriously more sensitive to low than high spatial frequency information in the visual input (Johnson, 2005; Vuilleumier, Armony, Driver, & Dolan, 2003), given, for example, the retinal projection via the magnocellular pathway to the midbrain's superior colliculus (Schiller, Malpeli, & Schein, 1979). A monocular advantage for low- but not high-frequency faces would, therefore, serve as a reliable signature of subcortical involvement.

Finally, the fourth manipulation is driven from the evidence that newborns' visual system evinces a preference toward face-like configurations (Johnson, 2005; Johnson & Morton, 1991; Johnson et al., 1991). If this preference is guided by subcortical structures, which are involved in face perception, we would predict a monocular advantage for any stimulus with face-like configuration but not for a stimulus that violates this configural arrangement. In this task, participants compare geometrical shapes that are aligned in face-like or non-face-like configurations. A monocular benefit only for face-like configurations will provide support for the claim that the same subcortical mechanisms that guide newborns' preference toward faces are also involved in the human adult face processing network.

METHODS

Participants

Participants (105 in total, 13 left-handers), all of whom had normal or corrected-to-normal vision, consented to

participate. Participants performed same/different judgments with either upright stimuli (age 18–27 years; 7 women and 8 men) or inverted stimuli (age 20–30 years; 8 women and 7 men) in the first study. They performed same/different sex judgments (age 18–22 years; 14 women and 8 men) in the second study or made same/different judgments with low-pass (age 18–24 years; 9 women and 6 men) or high-pass filtered stimuli (age 18–26 years; 7 women and 8 men) in the third study. In the final study, participants performed same/different judgments (age 18–21 years; 14 women and 9 men) on geometric shapes aligned in face-like or non-face-like configurations. No participant completed more than one experiment. Participants volunteered to participate in exchange for payment or course credits and the protocol was approved by the Institutional Review Board of Carnegie Mellon University.

Stimuli

Twenty-four male and 24 female face images, obtained from the Face-Place Database Project (Copyright 2008, Dr. M. Tarr, wiki.cmb.cmu.edu/Face_Place), were used in the experiments. All images displayed front views of faces with neutral emotional expression (see example in Figure 2). The faces were cropped to remove hair cues and were presented in grayscale against a black background. Face stimuli were 8° in height and 6° in width. Letter string stimuli consisted of 48 four-letter strings (24 pairs), presented in white Times New Roman font against a black background, approximately 2° in height and 5.5° in width. Each pair was matched for brightness. Car stimuli consisted of 48 cars, oriented to 45° (24 pairs), approximately 8.5° in width and 6° in height. Participants responded by pressing the "P" button of a keyboard using the right index finger for "same" and "Q" button of a keyboard using the left index finger for "different." In all image discrimination experiments, faces, cars, and letter strings were presented in different blocks of trials.

In the last experiment, face-like configured images and the non-face-like configured images, 8° in height and 6° in width, were presented randomly within the experimental blocks. Both image types were constructed of three geometrical shapes (circles or squares) 0.6° in height and 0.6° in width that appeared inside a white oval identical in size and shape to the faces in the face images. Images could contain either two circles and one square or two squares and one circle. The two images to be compared always shared the same overall configuration, and when they differed, the specific arrangement of the geometric shapes was changed (e.g., a square and a circle switched locations). For the face-like configuration images, two geometric shapes appeared at the upper part of the oval (1° above the center) equidistantly of the horizontal plane (0.5° to the right and left of the center) indicating the eyes. Another geometric shape appeared at the lower part of the image (1° below the center)

indicating a mouth. For the non-face-like configuration images, the three geometric shapes appeared equidistant from top right to bottom left or from top left to bottom right (see Figure 2C).

Procedure

The same procedure was employed across all experiments. The participant's head was stabilized with the aid of a chin rest. Two mirrors, one at 45° and one at 135°, each reflecting one of two monitors (50 cm from left or right side of observer), were placed in front of the participant (see Figure 1). Two cardboard dividers were attached to the chin rest, blocking the participant's direct view of the monitors, so that the display was only visible in the mirror. A single trial started with the appearance of a fixation cross (0.5°) for 1000 msec on both monitors (see Figure 2A). Participants were instructed to maintain fixation throughout the experiment. The first image appeared for 1000 msec followed by 1000 msec fixation and then by the second image for 1000 msec. Participants were instructed to respond after the appearance of the second image. If no response (by 2500 msec) or a wrong response was delivered, three red Xs appeared on the screen providing feedback for 1500 msec. If a correct response was given, a blank screen ensued for 1500 msec before the next trial.

Discrimination of Faces, Cars, and Words: Upright and Inverted Stimuli

In this first experiment, a trial consisted of a pair of faces (front views, neutral expressions; see Figure 2B), or letter strings or cars. Half of the trials contained the identical image presented twice (the "same" condition), whereas the remaining half contained two different images (the "different" condition). Participants were randomly assigned to the upright or inverted condition. In the inverted condition, all stimuli (cars, letter string, and faces) were rotated 180° in plane.

On half of the trials, both images were presented to the same eye, and on the other half, each image was presented to a different eye; these trial types were randomized in a block. For each visual category (faces, cars, strings) and each orientation, participants completed three blocks of trials with each block comprising 96 trials (24 trials for same/different response × same-/different-eye presentation). The order of the blocks was counterbalanced across participants. Responses were made via button presses, and accuracy and RT were measured.

Same/Different Sex Experiment

In this sex discrimination experiment, participants had to determine whether the two faces were of individuals of the same or different sex. Participants completed one block of 96 face trials using the same images of

the upright face condition (see above). Each block began with 16 practice trials. Participants were instructed to respond to the presentation of the second image as quickly and as accurately as possible.

Spatial Frequency Experiment

The availability of low/high spatial frequency information was manipulated by applying Gaussian filters to the original, upright images of faces, cars, and words. This was done by (i) applying fast Fourier transformations to the original images, (ii) multiplying Fourier energy profiles with Gaussian filters, (iii) bringing the results back into the image domain via inverse Fourier transforms, and (iv) normalizing the resulting images with the same luminance mean and root mean square contrast as the original images. A broadband band-pass filter was used initially to restrict frequencies between 0.3° and 10.6° cycles per degree of visual angle (cpd). Then, two types of filters, low pass (<1.3° cpd) and high pass (>5.3° cpd), were used to generate stimuli preserving low and high spatial frequency information. The difference between low and high spatial frequency stimuli was maximized by selecting cutoffs separated by two octaves. Image filtering and normalization were performed separately for each category of stimuli. For this experiment, the background color was gray, and the pairs of letter strings differed only in a single letter.

Participants were randomly assigned to the low or high spatial frequency condition. They then performed three blocks of trials for different types of stimuli (faces, letter strings, and cars). A block comprised 96 trials (24 trials for each eye validity condition × stimulus type × response). The order of the blocks was counterbalanced across participants.

Discrimination of Geometrical Shapes Aligned in Face-like or Non-face-like Configurations

This final experiment addresses the concern that any monocular facilitation we observe might not be face-specific but, rather, might result from differences in low-level visual attributes among the visual classes. The previous experiment examines spatial frequency as a potential property, but other possibilities also exist; for example, subcortical structures (e.g., LGN) have concentric receptive fields (RFs), whereas cortical structures (e.g., V1) have elongated RFs (Schiller & Malpeli, 1978). Because face stimuli were primarily composed of circular shapes whereas cars and letter strings were composed of more "edgy" (lines, wedges) structures, RF properties could potentially account for the monocular facilitation of faces. Therefore, to claim that subcortical regions are face-selective (rather than selective to low-level visual features) specificity for faces should be demonstrated compared with a more closely matched control stimulus. Here, we use the well-known stimuli of Johnson (2005) to examine whether there is monocular facilitation for a stimulus

comprised of three darkened shapes (e.g., reflecting two eyes and mouth) aligned as a face or aligned in a different configuration.

In this geometric shapes discrimination experiment, participants had to determine whether two successive displays were identical or not. Participants completed two blocks of 96 images (in half of the trials the images had a face-like alignment, and in the other half the images had a non-face-like alignment; see Figure 2C). The experiment began with 16 practice trials. Participants were instructed to respond to the presentation of the second image as quickly and as accurately as possible.

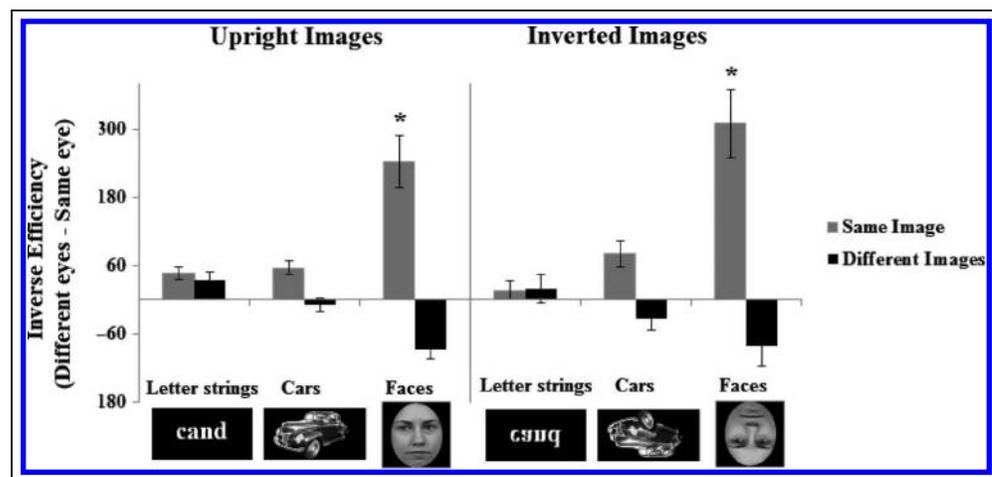
RESULTS

Because the experimental manipulation can influence both RT and accuracy, we used inverse efficiency (IE) score (RT correct responses divided by the proportion of correct responses) as the dependent measure, with lower scores reflecting better performance (Townsend & Ashby, 1983). There are a number of provisions for the use of this procedure, and we confirmed these a priori. For example, IE should not be used when accuracy rates are low; in all of our experiments accuracy exceeded 80%. In addition, IE should only be used when there is a positive correlation between RT and percent error. We calculated the linear correlation between the different conditions' RT and percent error. This correlation takes the RT values and percent error values across participants for every experimental condition, and so the number of cases on which the correlation is calculated (n) is dependent on the number of experimental conditions. We calculated the correlation for Experiments 1, 3, and 4 [$r(12) = .87$, $r(12) = .35$, $r(8) = .32$]. For Experiment 2, there were only four experimental conditions, so calculating a correlation for such a small number of cases is not informative. Yet the analyses of RT and accuracy reveals that the pattern of results was similar to that of the IE score.

Discrimination of Faces, Cars, and Words: Upright and Inverted Stimuli

Trials in which RT was longer than 1500 msec or shorter than 100 msec were excluded from the analyses (3%). On average, error rates constituted 7% and 9% of the trials for the upright and inverted groups, respectively. To explore both the category and orientation effects, we conducted an ANOVA with Stimulus Presentation (same, different eye), Image Match (same, different), and Stimulus Type (faces, letter string, and cars) as within-subject factors and Orientation (upright, and inverted) as a between-subject factor. Figure 3 presents the difference in mean IE between the same versus different eyes as a function of the three other factors listed above. There was a main effect of Stimulus Type, $F(2, 56) = 33.0$, $MSE = 56,510$, $p < .001$, with faces discriminated more poorly than letter strings, which did not differ from cars ($p < .05$ paired comparisons with Tukey correction). Performance was better for same-than different-eye presentation, $F(1, 28) = 42$, $MSE = 5,280$, $p < .001$, and was better for same than for different judgments, $F(1, 28) = 13.2$, $MSE = 20,044$, $p < .01$, and there was a significant advantage for same over different judgments for images presented to the same than different eyes (stimulus presentation \times same/different matching, $F(1, 28) = 52.6$, $MSE = 9,851$, $p < .001$). Most importantly, the Stimulus Type \times Stimulus Presentation interaction was significant,¹ $F(1, 56) = 9.1$, $MSE = 5,308$, $p < .001$, indicating a benefit for same-over different-eye presentation that held only for faces ($p < .05$ paired comparisons with Tukey correction). A significant three-way interaction between Stimulus Type, Stimulus Presentation, and Same/Different Matching was also observed, $F(2, 56) = 49.9$, $MSE = 5,205$, $p < .001$ ($p < .05$ paired comparisons with Tukey correction), as the same-versus different-eye advantage was greater when the two faces were the same than when they were different, and this enhancement was not evident for the other visual stimuli. There was neither an influence of stimulus orientation nor an interaction of

Figure 3. Difference (different eyes minus same eye) in IE (RT/proportion accuracy) for upright (left column) and inverted (right column) trials as a function of stimuli type plotted separately for same/different image trials. Positive values represent better performance for the same-eye than different-eye trials, and the error bars represent 1 SE.



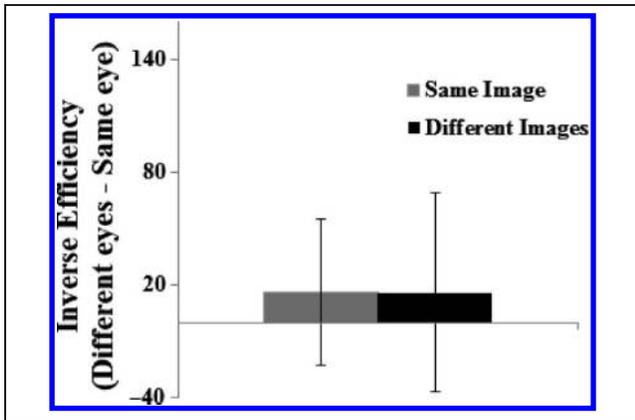


Figure 4. Inverse efficiency difference score for the sex comparison task plotted separately for same-/different-sex trials.

orientation with any other factor, all $F < 1$, ns , reflecting no difference between the upright and inverted condition. In a secondary analysis, we added as a factor the eye to which the first image was presented (right vs. left eye). This analysis showed that the same-eye advantage was present for both the right and left eyes and to an equivalent degree.

It is possible that the monocular benefit for faces may result from the fact that face matching might be harder than car or letter string matching. To address this interpretation, we reanalyzed our data to confirm that task demands cannot solely account for the differential pattern of performance on faces compared with cars and letter strings. Specifically, we extracted a subset of the 15 participants whose performance in the monocular condition (in the same image trials) was equivalently accurate for faces, cars, and letter strings (when comparing accuracy rates between faces and the other two conditions, $F(1, 14) = 1.2$, ns) and then compared the performance of these participants with the remaining 15 participants whose performance was not accurate across all stimulus types. We found that the three-way interaction demon-

strating the IE facilitation for faces in the same-eye condition compared with the different-eye condition for same images did not interact with subgroup. In this way, we rule out any potential confound of differential difficulty of the different stimulus types: Independent of overall accuracy, the advantage for face matching under the monocular condition still holds.

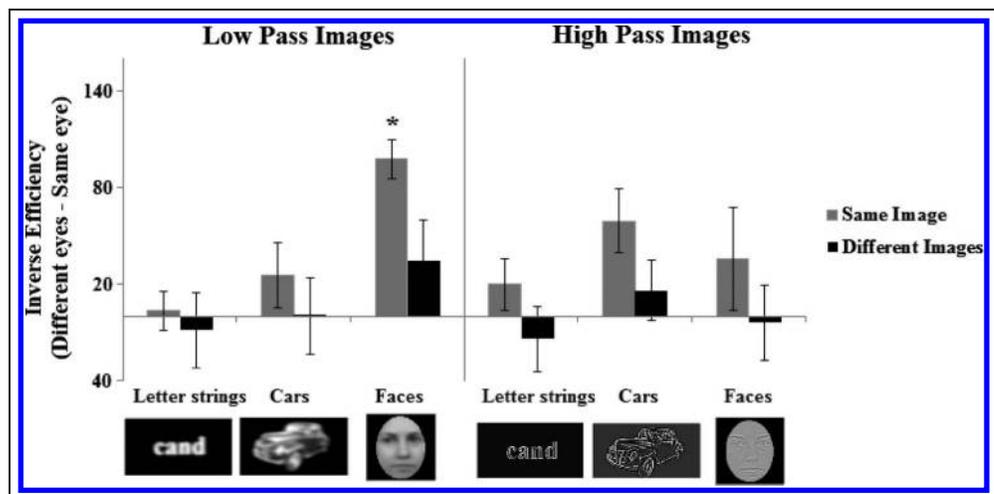
Same/Different Sex Experiment

Trials in which participants' RT was longer than 1500 msec or shorter than 100 msec were excluded from the analyses (8%). On average, participants made errors on 20% of the trials. An ANOVA with same/different Sex Judgments and Same/Different Eye Presentation as within-subject factors (see Figure 4) revealed no main effects nor significant interactions between any of the factors.

Spatial Frequency Experiment

Trials in which participants' RT was longer than 1500 msec or shorter than 100 msec were excluded from the analyses (3%). On average, error rates constituted 7% of the trials for the low-frequency group and 6% for the high-frequency group. An ANOVA was conducted with Spatial Frequency (high, low) as a between-subject factor and Stimulus Presentation (same/different eye), Same/Different Match (same, different) and Stimulus Type (faces, letter string, and cars) as within-subject factors. Figure 5 presents the difference in mean IE between same- and different-eye presentations as a function of the other three factors. In these analyses, we focused specifically on the interactions with spatial frequency. The interaction between stimulus type and spatial frequency was significant, $F(1, 56) = 3.8$, $MSE = 24,219$, $p < .05$ (none of the paired comparisons reached significance using a stringent Tukey post hoc procedure). More importantly, the Stimulus Type \times Stimulus Presentation \times Spatial Frequency interaction was significant,² $F(2, 56) = 4.5$, $MSE = 2,473$,

Figure 5. Difference (different eyes minus same eye) in IE (RT/proportion accuracy) for low-pass frequency images (left column) and high-pass frequency images (right column) trials as a function of stimuli type plotted separately for same/different trials. Positive values represent better performance for the than same-eye different-eye trials, and the error bars represent 1 SE.



$p < .05$, indicating a benefit for same- over different-eye presentation but only for the low-pass images of faces ($p < .05$ paired comparisons with Tukey HSD post hoc correction). The same-eye advantage for low-frequency face images was evident for both same and different judgment trials, but, as above, was enhanced for same trials ($p < .05$ paired comparisons, Tukey HSD post hoc correction).

The finding that there was no difference in overall performance between the two spatial frequency conditions for face stimuli (806 and 824 averaged IE score for the low and high spatial frequency, respectively) but that the same-eye (over different-eye) advantage was only observed in the low-frequency condition is a further indication that the pattern of results (differential frequency effect for faces) is not simply a product of differential task demands imposed by different stimulus types.

Discrimination of Geometrical Shapes Aligned in Face-like or Non-face-like Configurations

Trials in which participants' RT was longer than 1500 msec or shorter than 100 msec were excluded from the analyses (3.5%). On average, error rates constituted 6% of the trials. An ANOVA was conducted with Stimulus Presentation (same/different eye), Same/Different Match (same, different) and Stimulus Type (face-like, non-face-like) as within-subject factors. Figure 6 presents the difference in mean IE between Same- and Different-Eye presentations as a function of the other two factors.

In these analyses, we focus specifically on the interactions with stimulus type. Most importantly, the Stimulus

Type \times Stimulus Presentation \times Same/Different Match interaction was significant, $F(1, 22) = 4.8$, $MSE = 2,105$, $p < .05$, indicating a benefit for same-over different-eye presentation but only for the matched images in the face-like configuration³ ($p < .05$ paired comparisons, Tukey HSD post hoc correction).

DISCUSSION

In this study, we explored the nature of the functional contribution of evolutionarily rudimentary visual regions to face perception by determining whether face (but not car or letter string) matching differs depending on the eye/s to which the visual images are presented. The logic is that any facilitation afforded by two stimuli presented to the same eye (monocularly) versus to two different eyes (interocularly) would indicate involvement of the monocular portions of the visual pathway. Adult participants were significantly better at judging the likeness of two faces than the likeness of two cars or of two letter strings, when the stimuli were presented to the same eye compared with when they were presented to different eyes. Having established that this monocular enhancement was selective for faces, we then demonstrated that the monocular advantage was of equal magnitude for faces presented in the upright and inverted orientations, was not present when participants judged the sex of two consecutively presented faces to be the same or different, and was evident only for low- but not high-pass face images. Finally, the monocular benefit is only observed when the inputs are face-like in their spatial configuration.

We have argued that the presence of the monocular benefit only for the similarity judgment tasks but not for the more complex sex comparison task is consistent with the upper bounds of the subcortical contribution. That is, the subcortical contribution is limited to simple representations that can be engaged when the task is simple but when the task requires more abstract or semantic representations, subcortical representations do not suffice and cortical regions are implicated (which are indifferent to the visual information eye of origin). On the surface, this claim seems to be inconsistent with some existing findings. For example, Khalid, Finkbeiner, König, and Ansorge (2012) demonstrated that low-pass (but not high-pass) filtered face primes presented peripherally produce a congruency effect in a sex discrimination task; that is, performance was enhanced when the preceding prime and following probe were of the same gender compared with when they were not. This finding seems to suggest that information about the sex of a face may be subcortically represented, contrary to our findings. We note, however, that it is possible that the featural (low-level similarity) might be higher within versus across sex and that this featural similarity, which may well be represented subcortically, can enhance the processing of the target image and facilitate its sex classification. Our task required that participants decide whether two sequentially presented faces are of the

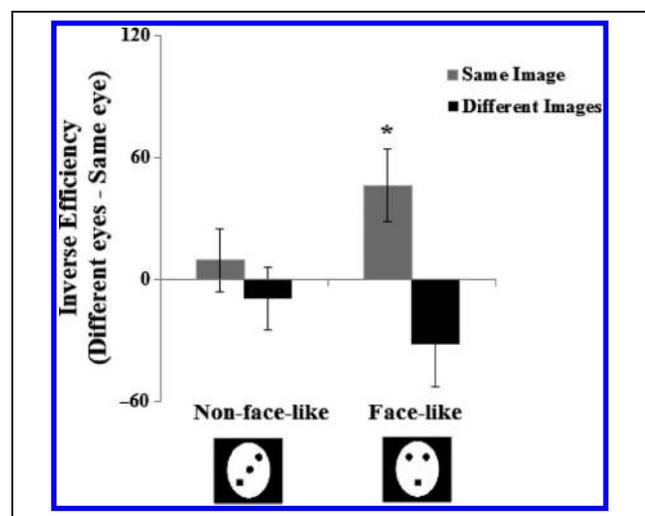


Figure 6. Difference (two eye minus same eye) in IE (RT/proportion accuracy) for non-face-like configuration images (left column) and face-like configuration images (right column) trials plotted separately for same/different trials. Positive values represent better performance for the same-eye than different-eye trials, and the error bars represent 1 SE.

same sex or not: This requires a more abstract comparison across different images, and subcortical mechanisms may not suffice under this more challenging condition. Also, in contrast with our study, the authors did not compare the facilitation afforded by the congruent primed face to other stimulus types, and so it remains an open question whether the observed facilitation by a preceding prime was specific to faces or not. Finally, the authors examined the influence of face processing without awareness while our task examined overt face comparison, and these might involve different neural substrates. It is not so clear, therefore, that our findings are at odds with those of Khalid et al. (2012), and further consideration for the apparent differences is warranted.

Taken together, our findings implicate subcortical visual regions in face perception: by virtue of sharing the same monocular pathway in the same-eye condition, the second face capitalizes on the activation of the visual pathway triggered by the first presentation of the face. That these prestriate face representations license only simple, possibly feature-based comparison of faces, as reflected by the equivalent magnitude, same-eye advantage for upright and inverted faces further suggests that the mechanism at play is not cortical. Finally, given that the same-eye advantage was observed for low- but not high-pass face images and is observed for any stimuli with a face-like configuration is also consistent with a subcortical mechanism (Johnson, 2005) and suggests that the subcortical mechanisms that are involved in the newborns' preference for face-like configurations are present in adults and may play a functional role in face perception.

Arguably, the information derived by the subcortical system is rather impoverished; these prestriate face representations are coarse, and although able to support direct face image matching, they are unlikely to contribute significantly to more complex tasks such as sex or identity recognition. To the extent that data exist to support these findings, we note that habituation to faces has been observed in subcortical structures as reported in a functional imaging study in which BOLD adaptation was observed in a comparison of second versus first run of images (Mende-Siedlecki et al., 2013). It was also recently demonstrated that the pulvinar nuclei in monkeys is responsive to face-like stimuli (Nguyen et al., 2013). This adaptation, however, might occur as a function of featural/geometric rather than of identity similarity per se.

We have suggested that the mechanism that supports the finding of a monocular advantage is prestriate. However, layer IV of area V1 is composed almost entirely of monocular neurons, as well, and so, contrary to our claim, one might propose that the results originate in V1 itself rather than in prestriate cortex. We argue, however, that the presence of the same-eye advantage only for low-pass face images is entirely consistent with a subcortical system (Johnson, 2005), which receives direct input from the retina (Benevento & Standage, 1983; Schiller et al., 1979). Because striate cortex is less spatial frequency

dependent, the differential advantage for low-pass faces likely rules out cortex as a mediating mechanism, and so a more parsimonious account is one that implicates a prestriate system. Yet a further interpretation of our data is possible: One might argue that, although the cortex receives a signal that is mixed from the two eyes, a cortical mechanism may still be biased to faster detection of a match of the input comes from the same eye. This claim may well be true; however, given that the facilitation we observe is limited to the repetition of an identical face, but not of an identical car or letter string, explanations in terms of general neural repetition or low-level feature matching are not obviously supported. Instead, our findings are consistent with the view, referred to as the "mid-brain hypothesis" that face perception can be mediated by subcortical structures. Although the midbrain hypothesis is specifically designated to refer to subliminal face perception (and is focused more on the analysis of emotional/motivational aspects of faces), the data from those studies support the engagement of the subcortical system in face perception (see Finkbeiner & Palermo, 2009; see also Khalid et al., 2012, for review of this hypothesis). Although we are not able to identify which particular substructure of the subcortical system is implicated here, the most likely candidates are the amygdala and superior colliculus based on existing evidence. On the basis of their analysis of the large group of participants, Mende-Siedlecki et al. (2013) observed robust and individually defined (for the most part) face-selective activation in these two structures. It is also the case that these structures (and particularly, the amygdala) are functionally coupled to posterior face-selective regions as well as to parts of visual cortex, and the amygdala itself has many subdivisions, some of which may be differentially implicated (Aggleton, 2000). We note that our findings cannot be accounted for by this coupled connectivity; if that were so and the behavioral findings we report were subserved by cortical structures, we should have obtained a difference in performance for upright versus inverted faces. At this stage, then, we can assert that, in adult humans, subcortical structures appear to participate in face perception and the representations that are engaged are somewhat coarse.

As noted in the Introduction, phylogenetic and ontogenetic evidence point to a rudimentary structure of set of structures that may be conserved in human adults. Our data, along with the emerging imaging data, are compatible with this claim. Moreover, there is some further, supporting evidence from studies of human patients with epilepsy in which intracranial field potentials in the amygdala of such patients showed stronger gamma band activity to faces than houses or scrambled faces (Sato et al., 2012; Pourtois, Spinelli, Seeck, & Vuilleumier, 2010). It should be noted that commonly used imaging techniques suffer from many limitations and artifacts when adopted to study the involvement of subcortical structures and so obtaining neural evidence for subcortical involvement in humans is not trivial (LaBar et al., 2001). The approach

used in this study, which capitalizes on neural pathway differences for monocular versus binocular visual projections, is the first to indicate the functional contribution of subcortical regions in face perception in adults. Face perception has been extensively studied using functional imaging, and to date, few studies have implicated subcortical mechanisms, perhaps a reflection that the method may not have sufficient sensitivity. Future work using more sensitive imaging methods and analytic techniques may help shed light on the involvement of more primitive structures in higher cognitive abilities such as face perception.

Our findings provide evidence that subcortical face representations play a role in perception in adults and contribute beyond bootstrapping face recognition in infancy or aiding automatic emotional processing (Kleinhans et al., 2011). Subcortical visual regions are conserved across species and developmental scale. Indeed, these structures are neither simply a vestige of phylogeny nor of ontogeny; rather, they contribute by deriving a coarse first-pass representation of a face that may assist accurate and rapid face perception, perhaps propagating the signals to cortex for more detailed and sophisticated computation.

Acknowledgments

We would like to thank the anonymous reviewers for their thoughtful comments and helpful suggestions and, in particular, the idea for replicating the Johnson experiment, as we have done in the last experiment.

Reprint requests should be sent to Shai Gabay or Marlene Behrmann, Department of Psychology, Carnegie Mellon University, 5000 Forbes Ave., Pittsburgh, PA 15213, or via e-mail: shaigaba@andrew.cmu.edu, behrmann@cmu.edu.

Notes

1. This interaction was also significant when examining either just RT or just accuracy as dependent measurements.
2. This interaction was also significant when using accuracy as a dependent measurement.
3. This comparison was also significant when using just RT as the dependent measurement.

REFERENCES

Aggleton, J. P. (2000). *The amygdala: A functional analysis*. Oxford, UK: Oxford University Press.

Avidan, G., & Behrmann, M. (2009). Functional MRI reveals compromised neural integrity of the face processing network in congenital prosopagnosia. *Current Biology*, *19*, 1146.

Batson, M. A., Beer, A. L., Seitz, A. R., & Watanabe, T. (2011). Spatial shifts of audio-visual interactions by perceptual learning are specific to the trained orientation and eye. *Seeing and Perceiving*, *24*, 579–594.

Bell, A. H., Hadj-Bouziane, F., Frihauf, J. B., Tootell, R. B., & Ungerleider, L. G. (2009). Object representations in the temporal cortex of monkeys and humans as revealed

by functional magnetic resonance imaging. *Journal of Neurophysiology*, *101*, 688–700.

Benevento, L. A., & Standage, G. P. (1983). The organization of projections of the retinorecipient and nonretinorecipient nuclei of the pretectal complex and layers of the superior colliculus to the lateral pulvinar and medial pulvinar in the macaque monkey. *The Journal of Comparative Neurology*, *217*, 307–336.

Bentin, S., Allison, T., Puce, A., Perez, E., & McCarthy, G. (1996). Electrophysiological studies of face perception in humans. *Journal of Cognitive Neuroscience*, *8*, 551–565.

Bi, H., Zhang, B., Tao, X., Harwerth, R., Smith, E., & Chino, Y. (2011). Neuronal responses in visual area V2 (V2) of macaque monkeys with strabismic amblyopia. *Cerebral Cortex*, *21*, 2033–2045.

Blake, R., & Cormack, R. H. (1979). Psychophysical evidence for a monocular visual cortex in stereoblind humans. *Science*, *203*, 274.

Dyer, A. G., Neumeier, C., & Chittka, L. (2005). Honeybee (*Apis mellifera*) vision can discriminate between and recognise images of human faces. *Journal of Experimental Biology*, *208*, 4709–4714.

Fairhall, S. L., & Ishai, A. (2007). Effective connectivity within the distributed cortical network for face perception. *Cerebral Cortex*, *17*, 2400–2406.

Finkbeiner, M., & Palermo, R. (2009). The role of spatial attention in nonconscious processing a comparison of face and nonface stimuli. *Psychological Science*, *20*, 42–51.

Goossens, L., Kukulja, J., Onur, O. A., Fink, G. R., Maier, W., Griez, E., et al. (2009). Selective processing of social stimuli in the superficial amygdala. *Human Brain Mapping*, *30*, 3332–3338.

Haxby, J. V., Hoffman, E. A., & Gobbini, M. I. (2000). The distributed human neural system for face perception. *Trends in Cognitive Sciences*, *4*, 223–233.

Herrmann, M., Ehlis, A.-C., Ellgring, H., & Fallgatter, A. (2005). Early stages (P100) of face perception in humans as measured with event-related potentials (ERPs). *Journal of Neural Transmission*, *112*, 1073–1081.

Hoffman, K. L., Gothard, K. M., Schmid, M. C., & Logothetis, N. K. (2007). Facial-expression and gaze-selective responses in the monkey amygdala. *Current Biology*, *17*, 766–772.

Horton, J. C., Dagi, L. R., McCrane, E. P., & de Monasterio, F. M. (1990). Arrangement of ocular dominance columns in human visual cortex. *Archives of Ophthalmology*, *108*, 1025.

Ishai, A., & Yago, E. (2006). Recognition memory of newly learned faces. *Brain Research Bulletin*, *71*, 167–173.

Jeffreys, D. A. (1996). Evoked potential studies of face and object processing. *Visual Cognition*, *3*, 1–38.

Johnson, M. H. (2005). Subcortical face processing. *Nature Reviews Neuroscience*, *6*, 766–774.

Johnson, M. H., Dziurawiec, S., Ellis, H., & Morton, J. (1991). Newborns' preferential tracking of face-like stimuli and its subsequent decline. *Cognition*, *40*, 1–19.

Johnson, M. H., & Morton, J. (1991). *Biology and cognitive development: The case of face recognition*. Oxford, UK: Blackwell.

Karni, A., & Sagi, D. (1991). Where practice makes perfect in texture discrimination: Evidence for primary visual cortex plasticity. *Proceedings of the National Academy of Sciences*, *88*, 4966–4970.

Kaul, C., Rees, G., & Ishai, A. (2010). The gender of face stimuli is represented in multiple regions in the human brain. *Frontiers in Human Neuroscience*, *4*.

Khalid, S., Finkbeiner, M., König, P., & Ansorge, U. (2012). Subcortical human face processing? Evidence from masked

- priming. *Journal of Experimental Psychology: Human Perception and Performance*, *39*, 989–1002.
- Kleinmans, N. M., Richards, T., Johnson, L. C., Weaver, K. E., Greenson, J., Dawson, G., et al. (2011). fMRI evidence of neural abnormalities in the subcortical face processing system in ASD. *NeuroImage*, *54*, 697.
- LaBar, K. S., Gitelman, D. R., Mesulam, M.-M., & Parrish, T. B. (2001). Impact of signal-to-noise on functional MRI of the human amygdala. *NeuroReport*, *12*, 3461–3464.
- Liu, J., Harris, A., & Kanwisher, N. (2002). Stages of processing in face perception: An MEG study. *Nature Neuroscience*, *5*, 910–916.
- Liu, N., Kriegeskorte, N., Mur, M., Hadj-Bouziane, F., Luh, W.-M., Tootell, R. B., et al. (2013). Intrinsic structure of visual exemplar and category representations in macaque brain. *The Journal of Neuroscience*, *33*, 11346–11360.
- Logothetis, N. K., Guggenberger, H., Peled, S., & Pauls, J. (1999). Functional imaging of the monkey brain. *Nature Neuroscience*, *2*, 555–562.
- Mende-Siedlecki, P., Verosky, S. C., Turk-Browne, N. B., & Todorov, A. (2013). Robust selectivity for faces in the human amygdala in the absence of expressions. *Journal of Cognitive Neuroscience*, *25*, 2086–2106.
- Menon, R. S., Ogawa, S., Strupp, J. P., & Uğurbil, K. (1997). Ocular dominance in human V1 demonstrated by functional magnetic resonance imaging. *Journal of Neurophysiology*, *77*, 2780–2787.
- Nguyen, M. N., Hori, E., Matsumoto, J., Tran, A. H., Ono, T., & Nishijo, H. (2013). Neuronal responses to face-like stimuli in the monkey pulvinar. *European Journal of Neuroscience*, *37*, 35–51.
- Pessoa, L., & Adolphs, R. (2010). Emotion processing and the amygdala: From a “low road” to “many roads” of evaluating biological significance. *Nature Reviews Neuroscience*, *11*, 773–783.
- Pourtois, G., Spinelli, L., Seeck, M., & Vuilleumier, P. (2010). Temporal precedence of emotion over attention modulations in the lateral amygdala: Intracranial ERP evidence from a patient with temporal lobe epilepsy. *Cognitive, Affective, & Behavioral Neuroscience*, *10*, 83–93.
- Rossion, B., & Gauthier, I. (2002). How does the brain process upright and inverted faces? *Behavioral and Cognitive Neuroscience Reviews*, *1*, 63–75.
- Sabatinelli, D., Keil, A., Frank, D. W., & Lang, P. J. (2013). Emotional perception: Correspondence of early and late event-related potentials with cortical and subcortical functional MRI. *Biological Psychology*, *92*, 513–519.
- Sato, W., Kochiyama, T., Uono, S., Matsuda, K., Usui, K., Inoue, Y., et al. (2012). Temporal profile of amygdala gamma oscillations in response to faces. *Journal of Cognitive Neuroscience*, *24*, 1420–1433.
- Schiller, P., & Malpeli, J. G. (1978). Functional specificity of lateral geniculate nucleus laminae of the rhesus monkey. *Journal of Neurophysiology*, *41*, 788–797.
- Schiller, P., Malpeli, J. G., & Schein, S. (1979). Composition of geniculostriate input of superior colliculus of the rhesus monkey. *Journal of Neurophysiology*, *42*, 1124–1133.
- Schwarzkopf, D. S., Schindler, A., & Rees, G. (2010). Knowing with which eye we see: Utrocular discrimination and eye-specific signals in human visual cortex. *PLoS One*, *5*, e13775.
- Self, M. W., & Roelfsema, P. R. (2010). A monocular, unconscious form of visual attention. *Journal of Vision*, *10*, article 17.
- Sheehan, M. J., & Tibbetts, E. A. (2011). Specialized face learning is associated with individual recognition in paper wasps. *Science*, *334*, 1272–1275.
- Simion, F., Valenza, E., Umiltà, C., & Barba, B. D. (1998). Preferential orienting to faces in newborns: A temporal-nasal asymmetry. *Journal of Experimental Psychology: Human Perception and Performance*, *24*, 1399.
- Tibbetts, E. A. (2002). Visual signals of individual identity in the wasp *Polistes fuscatus*. *Proceedings of the Royal Society of London, Series B, Biological Sciences*, *269*, 1423–1428.
- Todorov, A. (2012). The role of the amygdala in face perception and evaluation. *Motivation and Emotion*, *36*, 16–26.
- Todorov, A., Mende-Siedlecki, P., & Dotsch, R. (2013). Social judgments from faces. *Current Opinion in Neurobiology*.
- Townsend, J., & Ashby, F. (1983). *The stochastic modelling of elementary psychological processes*. Cambridge: Cambridge University Press.
- Troiani, V., & Schultz, R. T. (2013). Amygdala, pulvinar, and inferior parietal cortex contribute to early processing of faces without awareness. *Frontiers in Human Neuroscience*, *7*, article 241.
- Tsao, D. Y., Moeller, S., & Freiwald, W. A. (2008). Comparing face patch systems in macaques and humans. *Proceedings of the National Academy of Sciences, U.S.A.*, *105*, 19514–19519.
- Vuilleumier, P., Armony, J. L., Driver, J., & Dolan, R. J. (2003). Distinct spatial frequency sensitivities for processing faces and emotional expressions. *Nature Neuroscience*, *6*, 624–631.
- Williams, C., Azzopardi, P., & Cowey, A. (1995). Nasal and temporal retinal ganglion cells projecting to the midbrain: Implications for “blindsight”. *Neuroscience*, *65*, 577–586.
- Yovel, G., & Kanwisher, N. (2005). The neural basis of the behavioral face-inversion effect. *Current Biology*, *15*, 2256–2262.

This article has been cited by:

1. Shai Gabay, Charles Burlingham, Marlene Behrmann. 2014. The nature of face representations in subcortical regions. *Neuropsychologia* 59, 35-46. [[CrossRef](#)]