

Subcortical regions of the human visual system do not process faces holistically

Rebeka C. Almasi^{*}, Marlene Behrmann

Department of Psychology and Neuroscience Institute, Carnegie Mellon University, Pittsburgh, PA 15213, USA

ARTICLE INFO

Keywords:

Neuroscience
Visual cognition
Subcortical
Holistic processing
Composite face effect

ABSTRACT

Face perception is considered to be evolutionarily adaptive and conserved across species. While subcortical visual brain areas are implicated in face perception based on existing evidence from phylogenetic and ontogenetic studies, whether these subcortical structures contribute to more complex visual computations such as the holistic processing (HP) of faces in humans is unknown. To address this issue, we used a well-established marker of HP, the composite face effect (CFE), with a group of adult human observers, and presented two sequential faces in a trial monocularly or interocularly using a Wheatstone stereoscope. HP refers to the finding that two identical top (or bottom) halves of a face are judged to be different when their task-irrelevant bottom (or top) halves belong to different faces. Because humans process faces holistically, they are unable to ignore the information from the irrelevant half of the composite face, and this is true to an even greater extent when the two halves of the faces are aligned compared with when they are misaligned ('Alignment effect'). The results revealed the HP effect and also uncovered the Alignment effect, a key marker of the CFE. The findings also indicated a monocular advantage, replicating the known subcortical contribution to face perception. There was, however, no statistically significant difference in the CFE when the images were presented in the monocular versus interocular conditions. These findings indicate that HP is not necessarily mediated by the subcortical visual pathway, and suggest that further investigation of cortical, rather than subcortical, structures might advance our understanding of HP and its role in face processing.

1. Introduction

Recent studies have confirmed that, in addition to the role played by cortical regions, subcortical structures also play a functional role in face perception. The question we address here is whether, in the course of this functional role, the subcortical structures do so by engaging in complex computations, such as holistic processing (HP), that are considered critical for face perception. We note that the definition of HP is itself somewhat controversial: some define it as the coding of a face as a single non-decomposable unit or template (Farah, Wilson, Drain, & Tanaka, 1998), some characterize it as the tendency to represent or selectively attend to all parts of the face together (Richler & Gauthier, 2014), or, as noted by others, HP is the simultaneous integration of all parts into a single perceptual representation (Rossion, 2008). Yet others consider HP to be the processing of the spatial relations between parts of the face, a process also referred to as 'configural processing' (Maurer, Grand, & Mondloch, 2002; Maurer et al., 2007). The exact operational definition of HP is not critical for the current paper as all studies

acknowledge HP as a key component of face recognition.

Below, first, we review ontogenetic and phylogenetic evidence implicating a subcortical contribution to HP and face perception. We then examine evidence for the role of subcortical structures in HP and face perception in adult humans. Thereafter, we describe a well-established paradigm that assays HP and is considered a 'cornerstone' of face recognition research (Richler & Gauthier, 2014), the composite face paradigm, and, last, using this paradigm, we report data pertaining to HP in the subcortical visual system of adult humans. We conclude by arguing that there is no compelling evidence of HP by the subcortical visual system.

1.1. Ontogenetic and phylogenetic evidence for subcortical face perception

Existing ontogenetic and phylogenetic evidence has implicated subcortical neural structures in face perception, suggesting that some aspects of face perception may be evolutionarily conserved (Miller et al., 2019).

^{*} Corresponding author at: Department of Psychology, George Washington University, 2125 G St NW. Washington, District of Columbia 20052, USA.
E-mail address: almasi@gwu.edu (R.C. Almasi).

With respect to ontogeny, there is consensus that newborns show a bias to orient to stimuli that possess characteristics of faces (Cassia, Turati, & Simion, 2004), and converging evidence supports the view that this ability is largely mediated by a subcortical pathway (Atkinson, 2000; Farroni et al., 2005; Johnson, 2005; Johnson, Senju, & Tomalski, 2015). For example, many studies have shown that newborn babies have some bias to orient to face-like stimuli even when cortex is underdeveloped (Johnson, 2005) and some have argued, more specifically, that the superior colliculus itself is engaged in face perception (Simion, Valenza, Umiltà, & Dalla Barba, 1998). The subcortical pathway is thought to bias visual input to the developing cortical circuitry over the first weeks and months of life and, in so doing, bootstrap the relevant cortical mechanisms (Johnson et al., 2015; Morton & Johnson, 1991). One recent study has reported an eye-specific advantage in 3-4 month old infants' preferred face looking behavior, further implicating the monocular pathway (Dalrymple, Khan, Duchaine, & Elison, 2020). Relevant for the current study, 4-month old infants are sensitive to the orientation of the face, with poorer performance for inverted over upright faces (Turati, Sangrigoli, Ruel, & de Schonen, 2004), a pattern considered a benchmark of configural or holistic perception. This inversion decrement mirrors the well-known drop in face recognition performance in adults when faces are inverted (Yin, 1969), and its presence in very young children is intriguing given that infants do not have fully functional cortex. As such, the presence of the inversion effect in infancy may reflect the contribution of the subcortical visual pathway given the immaturity of cortex.

A similar argument is made with respect to phylogeny: lower-order species such as paperwasps are able to discriminate faces (Miller et al., 2019; Sheehan & Tibbetts, 2011; Tibbetts, 2002) and it has been suggested that this is accomplished through HP (Avargues-Weber, Portelli, Benard, Dyer, & Giurfa, 2010; Dyer, Neumeier, & Chittka, 2005). For example, free flying wasps are able to discriminate learned faces from novel similar faces (Avargues-Weber et al., 2018). Critically, following learning, the wasps' performance dropped markedly when the faces were inverted. That insects are sensitive to the relational structure of the face is provocative, as this relational or holistic process is considered to be a product of cortex (Rossion et al., 2000; Schiltz & Rossion, 2006), which they do not possess. By default, then, these inversion effect in insects must be a product of their subcortical system.

1.2. Subcortical contribution to face perception in adult humans

An obvious question, then, is whether subcortical contributions to face perception described above are functional in human adults: once the cortical face system is bootstrapped over development, reliance on subcortical contributions may be less relevant and perhaps even eliminated. With sufficient data and statistical power, however, face-selective activation of deep and small subcortical structures in adults can be discovered. For example, in one study, neuroimaging data from a large group ($N = 215$) revealed robust and replicable selectivity in response to face stimuli, independent of emotional expression (Mende-Siedlecki, Verosky, Turk-Browne, & Todorov, 2013; Stein, Seymour, Hebart, & Sterzer, 2014). The precise subcortical regions are generally not made explicit given the difficulty in segregating the structures but, at least in humans, likely involve the superior colliculus, pulvinar and lateral geniculate nucleus (Tamietto et al., 2010). In many non-human species, this (tectal-pulvinar) subcortical pathway is assumed to play the critical role of triggering fear responses in the amygdala and is sometimes referred to as 'the low road', to contrast with the geniculostriate 'high road' (LeDoux, 1994). In patients with cortical blindness, the amygdala has been found to activate for passive viewing of human faces and bodies, along with other non-cortical areas like the insula and cerebellum involved in perception of bodily states (Van den Stock et al., 2014), even without conscious awareness of the stimuli. Importantly, this object categorization could be feature-based rather than holistic, and previous theories of the "low road" being responsible for fast affective stimulus processing are presently being

challenged; some accounts of the amygdala's role propose that it is responsible for coordinating evaluation of biological significance and eschew the low-road vs high-road distinction altogether (Pessoa, Adolphs 2010). There is no doubt that the amygdala is one of the key regions implicated in face perception: in contrast with typical monkeys, monkeys with bilateral lesions to the amygdala showed no viewing preference for faces or illusory facial features over everyday objects even when no emotional content was presented (Taubert et al. 2018).

One approach for assessing subcortical engagement in human face perception is through the use of a Wheatstone stereoscope (Wheatstone, 1838) (Gabay, Nestor, Dundas, & Behrmann, 2014; Gabay, Burlingham & Behrmann, 2014). In these studies (see Fig. 1), we presented images of two faces, cars, or letter strings sequentially to either the same (monocular) or different (interocular) eyes (first stimulus to one eye, second stimulus to the same or the other eye), and participants made same/different judgments across the two images.

This method capitalizes on known properties of the visual system: the visual input is propagated in an eye-specific fashion from the retina through the visual system, and this monocular segregation is retained up to layer IV of striate cortex (Horton, Dagi, McCrane, & de Monasterio, 1990; Menon, Ogawa, Strupp, & Ugurbil, 1997). Given that observers are not explicitly aware of the eye to which the stimulus is projected (Blake & Cormack, 1979; Schwarzkopf, Schindler, & Rees, 2010), they perceive images in the monocular and interocular conditions as identical, with both images 'fused' into one in the center of the screen (see examples of percept in Fig. 2A and B). If performance in the same/different discrimination of the two stimuli presented to the same eye i.e. monocularly is better than when presented to two different eyes i.e. interocularly, then the monocular portions of the visual system are implicated, akin perhaps to a priming effect in the same processing channel. A performance advantage that is eye-dependent, then, is likely precortical (although V1 may be involved) and a product of this monocular pathway.

Our previous results revealed superior image discrimination in the same eye (monocular) vs different eye (interocular) condition for faces, but not for cars or letter-strings (Gabay, Nestor, Dundas, & Behrmann, 2014). Moreover, in a separate experiment, this monocular benefit was present only for low- but not high- frequency versions of the face images,

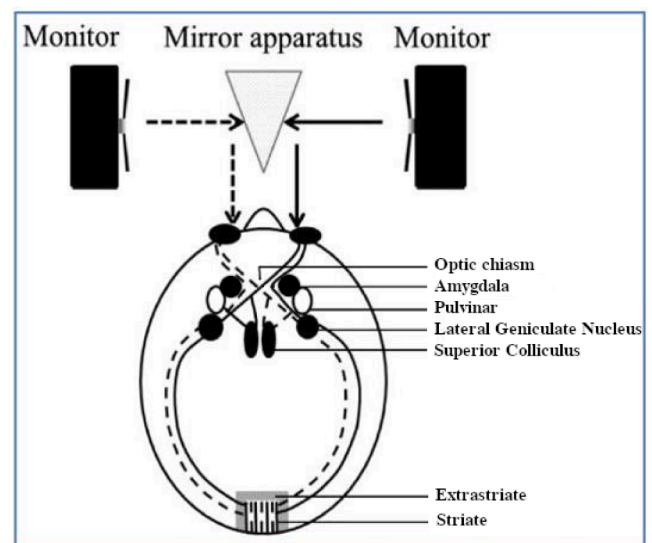


Fig. 1. Schematic representation of the Wheatstone mirror stereoscope and segregation of visual pathways. A chin rest stabilizes the participant's head, and mirrors are located at 45 and 135° along the edges of a cardboard divider that blocks the participant's direct view of the monitors. The mirrors reflect the image from a computer monitor placed 50 cm on the left and right of the observer. During calibration, the participant "fuses" the image by moving two fixation crosses until they merge. (From Gabay, Nestor, et al., 2014).

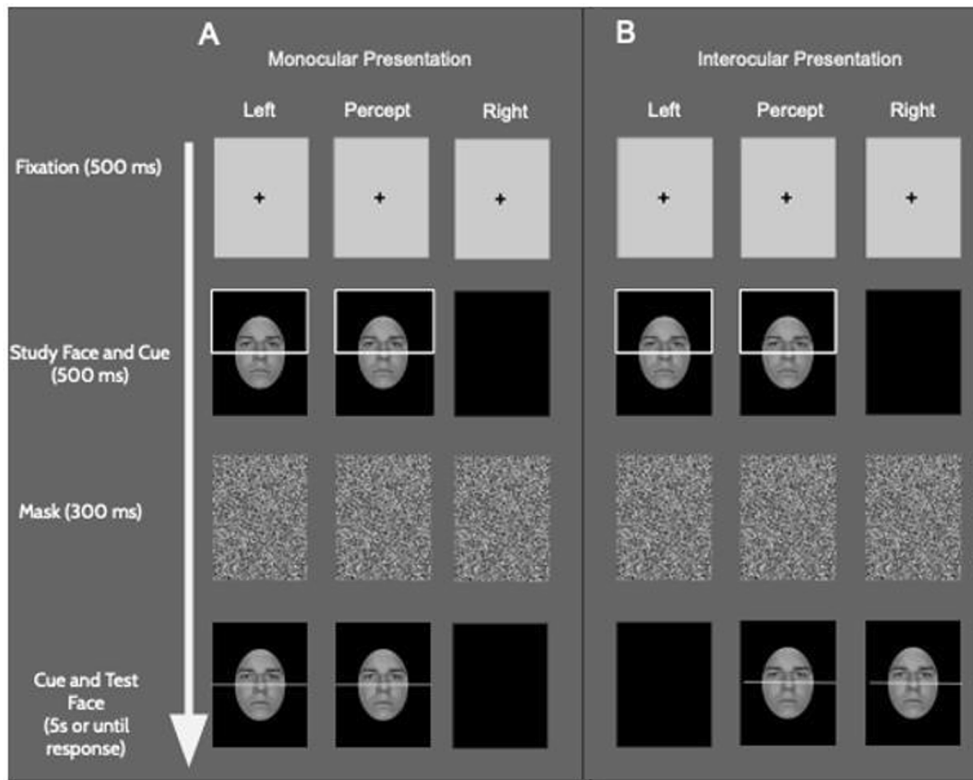


Fig. 2. A typical monocular (A) and interocular (B) trial. In both examples, following the initial fixation screen, the first ('study') image is presented to the left eye. After the mask, in the monocular condition, the second ('test') face is presented to the same (left) eye whereas in the interocular condition, the test face is presented to the right eye. The middle column shows the participant's fused perception which is identical irrespective of condition. In both types of trials shown here, a 'same' response is required.

and this effect of spatial frequency is consistent with the propagation of coarse magnocellular information through the subcortical pathways (Callaway, 2005; Kveraga, Boshyan, & Bar, 2007; Vuilleumier, Armony, Driver, & Dolan, 2003). Last, there was a monocular advantage for face-like configurations (schematic of two blobs for eyes and lower blob for mouth) but not for non-face-like configurations (three diagonal blobs)

(Gabay, Burlingham & Behrmann, 2014), further evidence of the perceptual sensitivity of faces that is compatible with subcortical computations (Johnson, 2005; Johnson, Dziurawiec, Ellis, & Morton, 1991; Schneider & Kastner, 2005).

These findings confirm that the monocular portion of the visual system is engaged by adult humans in the processing of faces, and the

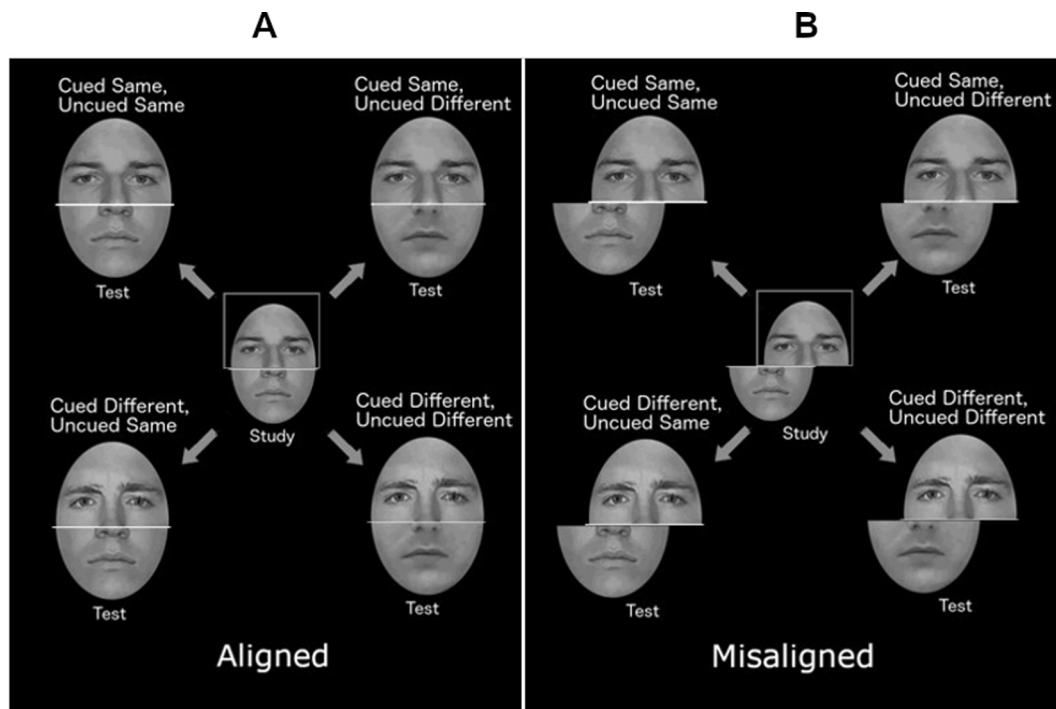


Fig. 3. Illustration of the composite face paradigm with (A) Aligned trials and (B) Misaligned trials. The "study" face is displayed in the center and contains a cue (white rectangle). It is replaced by a "test" face, and participants decide whether the cued sections of the two faces are identical or not.

question to be addressed, then, is whether this is accomplished via holistic processing. HP in humans is generally considered to arise from cortical structures (Andrews, DaviesThompson, Kingstone, & Young, 2010; Schiltz & Rossion, 2006), and thus, evidence for the participation of subcortical structures in human adults would shed further light on the computational capabilities of these more rudimentary structures.

1.3. Holistic processing paradigm

One of the best-established paradigms for evaluating HP is the ‘composite face effect’ (CFE). In this paradigm (Hole, 1994; Young, Hellawell, & Hay, 1987) (see Fig. 3), faces are typically bisected along a horizontal midline, and then upper and lower halves are recombined into a single composite image (for vertical CFE, see Liu & Behrmann, 2014). In a single trial, two composite faces are presented sequentially, with the first (‘study’) face accompanied by a cue informing participants whether to make the same/different judgement about the top or bottom halves of the two presented faces.

On half the trials, the two halves of the face are aligned (see Fig. 3A) or misaligned (see Fig. 3B). Because observers are assumed to process all parts of a face together, they cannot help but process the task-irrelevant half of the composite face (Richler et al., 2012), especially when the two halves of the faces are aligned and comprise a whole face shape (Fig. 3A). The intuition for this Alignment effect is that it takes a disproportionately long time for an observer to match two identical top halves of unfamiliar faces as ‘same’ when they are shown accompanied by two different bottom halves (or, equally, match two identical bottom halves when they are presented with two different top halves). The impact of alignment, considered a signature of HP, is also referred to as the Composite Face Illusion (Rossion, 2013).

In addition to this factor ‘Alignment’, a second factor known as ‘Congruency’ is also incorporated into the design of the CFE, making up the so-called *full composite* paradigm (Cheung, Richler, Phillips, & Gauthier, 2011; Richler, Cheung, & Gauthier, 2011; Richler, Cheung, Wong, & Gauthier, 2009), and yielding an orthogonal crossing of Alignment \times Congruency. *Congruent trials* are those in which, across study and test, both halves are the same or both are different (e.g. cued *same*, uncued *same* or cued *different*, uncued *different*). In contrast, in *incongruent trials*, across study and test, one half is the same and the other half is different, resulting in differential pairing across the two halves of the two face

displays (e.g. cued *same*, uncued *different* or cued *different*, uncued *same*).

While there is general acceptance that Alignment alone can reflect HP and the integration of face parts during face perception (Avidan, Tanzer, & Behrmann, 2011; Scherf, Whyte, Minshew, & Behrmann, 2020), Congruency alone or an Alignment \times Congruency interaction can also serve as markers of HP (Cheung et al., 2011). In the present study, we examine all main effects and interactions and discuss the relative effect of these manipulations on performance. Critically for our understanding of the subcortical contribution, crossings of all factors are presented monocularly and interocularly across the course of the experiment (see Fig. 4 for design of full design of the study).

1.4. Predictions: HP in subcortical structures

To evaluate HP and subcortical processing, we presented the full composite CFE paradigm using the Wheatstone stereoscope. The first predictions relate to the signature of HP per se and the expected outcome would be a main effect of Alignment, or a Congruency \times Alignment interaction (congruency alone generally considered too be insufficient and potentially reflects response competition). The novel hypothesis concerns HP and the role of the subcortical structures manifesting as interactions of HP with Eye. The key prediction is that we would expect to see an interaction of Eye \times Alignment or an interaction of Eye \times Alignment \times Congruency, with more obvious effects of Alignment or interaction with Alignment in the monocular than interocular pathway. If the HP markers reveal equal effects in monocular and interocular (so no interaction with Eye), we might parsimoniously conclude that the results reflect the sole cortical contribution to HP as the monocular signal is directly transmitted up to cortex. In other words, there would be no clear support for an independent role of the subcortical structures in HP.

To pre-empt our results, neither the interaction of Eye \times Alignment nor the interaction of Eye \times Alignment \times Congruency were significant. These results rule out a subcortical contribution to HP.

2. Materials and methods

2.1. Participants

We determined, using G*Power (Faul, Erdfelder, Lang, & Buchner, 2007) that 20 participants would suffice based on an effect size of 0.3

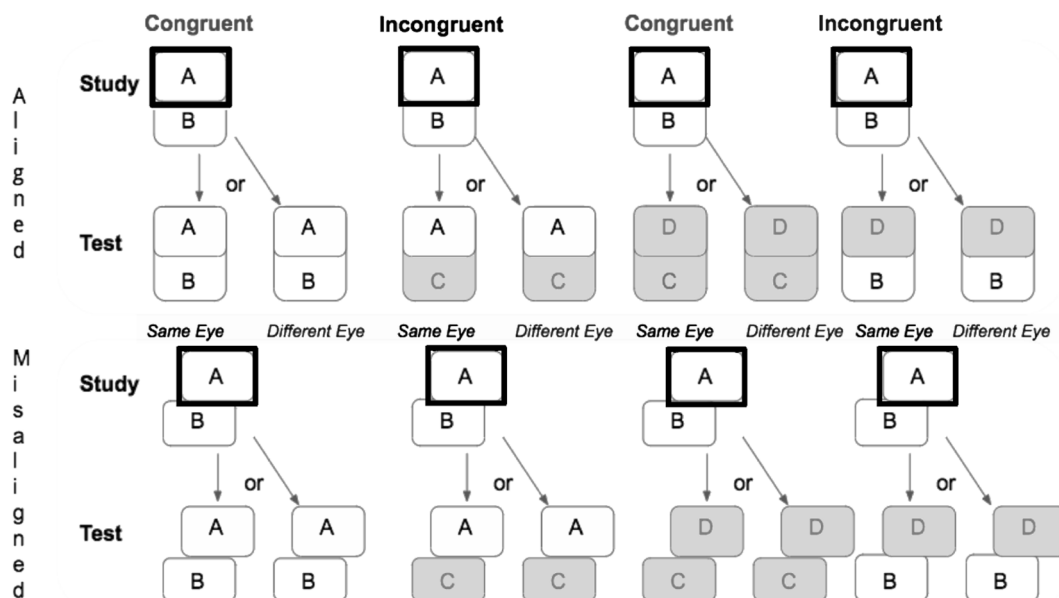


Fig. 4. Visualization of the conditions. The aligned version of faces is shown at the top and misaligned at the bottom of the figure. This diagram only displays trials where the “top” half of the face is cued, indicated by the bold black rectangle around the top half of study faces. There is an equivalent set of trials where the “bottom” half of the study face is cued. The “test” face is presented to either the same eye or the other eye compared to the study face.

with $p < .05$ difference between monocular and interocular conditions on a two-tailed t -test. Twenty-five participants (mean age 19.96, $SD = 1.93$, 14 females) were recruited from the Psychology student subject pool at Carnegie Mellon University and all received course credit for their participation. All had normal or corrected-to-normal vision and provided consent to participate in this study. The protocol was approved by the Carnegie Mellon University Institutional Review Board. Two participants had incomplete data and were excluded for any analyses. Additionally, three participants were left-handed or ambidextrous, and therefore their data were not included in the analyses either, leaving $n = 20$ for the analysis.

2.2. Stimuli

Face stimuli were created using 20 front-view Caucasian male faces (from Dr. James Tanaka database). The faces are centered in an oval so that the outer contour and hair are not visible. Each face was 8° in height and 6° in width and was displayed on a $10^\circ \times 8^\circ$ black background. The stimuli were controlled for luminance and contrast. The 20 faces were subdivided into five groups of four similar faces based on prior ratings of similarity (Liu & Behrmann, 2014). Each composite display was created by splitting each face along the horizontal midline by a white line, and then pairing the top half of one face with the bottom half of the same or another face from the same group. This ensured that more homogeneous faces were paired with each other so, when the two halves were combined, the image was still well configured. For aligned trials (see Fig. 3A), the two halves of the face shared an outer contour whereas, for misaligned trials (see Fig. 3B), there was an approximately 80 pixel offset between the top and bottom half, with the top half more rightward in half the misaligned trials and more leftward in the remaining half.

We employed a complete (or full) composite design, comprising congruent and incongruent trials crossed with whether the two halves of the faces were aligned or misaligned (see Fig. 3). Congruent trials are (i) same-congruent (cued/attended and uncued/unattended halves are both the same) and different congruent (cued/attended and uncued/unattended halves are both different), and the incongruent trials are (ii) same-incongruent (cued/attended halves are the same but uncued/unattended halves are different) and different-incongruent (cued/attended halves are different and uncued/unattended halves are different) (Richler & Gauthier, 2014).

2.3. Procedure

Each trial began with a black fixation cross on a gray screen for 500 ms (Fig. 2). Then, a composite 'study' face was shown for 500 ms with a yellow rectangle, the 'cue', indicating whether the top or bottom half was to be judged, followed by a 300-ms mask. The composite 'test' face was then shown, and remained visible for 5 s or until response. The participant responded by pressing two keys "f" or "j" for "same" and "different" judgements of the cued half, with the key assignment counterbalanced across participants.

As shown in Fig. 4, critically, the study and test displays could be presented to the same eye or to different eyes. Within each block, there was a full orthogonal crossing and randomized presentation of Study Eye (right/left) \times Test Eye (right/left) \times Cue (top/bottom) \times Congruence (congruent/incongruent). Alignment (aligned /misaligned) was blocked. Each participant completed 4 blocks of 200 trials each, for a total of 800 trials. Participants were given a break between blocks.

A chin rest was used to stabilize the participant's head in front of the Wheatstone stereoscope (see Fig. 1). Two ASUS VG248 ($24''$, 1920×1080 resolution, 60 Hz refresh rate) monitors were placed to the left and right of the center of the chin rest, perpendicular to the participant's line of sight, and facing each other. Two front surface mirrors were mounted at 45° , angled on a cardboard triangle, so that each mirror reflected the image from one of the monitors into one of the participant's eyes. The distance from the chin rest to the center of each mirror was 8 cm, and the

distance from the center of the mirror to the monitor on the same side was 55.5 cm. Two black cardboard dividers attached to the chin rest blocked the view of the monitors, so the display would only be seen through the mirror.

Before the experiment, the stereoscope was calibrated for each participant to ensure that the images presented to the two eyes separately were perceived as a single, fused image. To do so, we presented displays of a '+' to both eyes simultaneously, and asked participants how many plus signs were visible. We also participants to report on the sharpness of the image. If participants saw more than one plus sign, they were instructed to move the plus signs on the screen via key press until the images were fused into one. This shifted the position of the image/s on the two monitors, as needed. This same procedure was followed for recalibration after breaks.

Following the calibration, the experimental trials commenced. For each participant, the accuracy and reaction time (RT) of the same/different judgments were measured. The first 5 trials of each block were excluded from further analysis and, separately, blocks with accuracy levels below 75% were excluded (leading to the exclusion of a total of 13 blocks in total). This drop in performance may be related to fatigue. In the blocks for analysis, outliers in RT were defined as values more than 1.5 times the Interquartile Range (smaller than the first quartile, or larger than the third quartile) and winsorized, i.e. censored to the nearest non-outlier value to avoid deleting participants from the analysis.

Because the factors of Eye (monocular/interocular), Congruency (congruent/ incongruent) and Alignment (aligned/misaligned) were the most important variables, they were entered into a within subjects ANOVA. None of the other factors (namely, Cue location (top/bottom), Study eye (right/left) or Test eye (right/left)) are central to the hypothesis of CFE, and thus they are not considered further. We report the statistical results for accuracy, sensitivity (d'), and RT values as the dependent measures. Accuracy and d' (a measure of discriminability or sensitivity) are related in that both take the hit rate into account but d' also considers the false alarm rate, providing additional information. Posthoc Tukey comparisons at $p < .05$ and effect sizes (partial eta-squared) are included. To evaluate the weight of the evidence for a particular effect, we also calculated Bayes Factor (BF), BF_{10} , for the interactions of interest (using JASP Version 0.13.0). A BF between 0.33 and 1, 0.1 and 0.33, 0.03 and 0.1, 0.01 and 0.03, or less than 0.01 should be interpreted as anecdotal, moderate, strong, very strong, or extreme evidence for the null model, H_0 , respectively. On the other hand, a BF between 1 and 3, 3 and 10, 10 and 30, 30 and 100, or greater than 100 should be interpreted as anecdotal, moderate, strong, very strong, or extreme evidence for the alternate model, H_1 , respectively (Lee & Wagenmakers, 2014).

3. Results

The data were subjected to an Analysis of Variance (ANOVA) with Eye, Congruency and Alignment as the within subject factors. We present the results of the statistical analyses separated by dependent measures, starting with accuracy and followed by d' given their relationship, and then, last, by RT. Most studies do not report all the dependent measures, with some favoring d' and/or RT as more revealing, for example (Richler, Mack, Palmeri, & Gauthier, 2011), and others favoring accuracy and/or RT, for example (Rossion, 2013; Schiltz, Dricot, Goebel, & Rossion, 2010). As will be evident below, the same effects are not seen across all the dependent measures, complicating interpretation to some extent. Nevertheless, for full disclosure, we have opted to report the observed effects with all of the dependent measures.

3.1. Accuracy analysis

With accuracy as the dependent measure, there was a trend for performance to be significantly better for monocular (mean: 0.852) compared to interocular trials (mean: 0.838), ($F(1,19) = 3.6$, $p = .07$,

$\eta^2_p = .168$), reflecting the small benefit of two displays within the same monocular pathway ('the monocular advantage'). There was no effect of Alignment - the difference between aligned (mean: 0.846) and misaligned (mean: 0.844) trials was not significant, ($F(1,19) = 0.09$, $p > .7$, $\eta^2_p = .005$; $BF = 6.92 \times 10^{-13}$). There was, however, a statistically significant main effect of Congruency with significantly higher accuracy for congruent (mean: 0.88) than incongruent (mean: 0.81) trials, ($F(1,19) = 26.65$, $p < .001$, $\eta^2_p = .597$).

There was a significant interaction of Congruency \times Eye, ($F(1,19) = 26.1$, $p < .001$, $\eta^2_p = .592$) (see Fig. 5A, left): posthoc tests at $p < .05$ reveal an advantage for congruent over incongruent displays to a greater degree in the interocular [congruent (mean: 0.89) and incongruent (0.79)] than in the monocular [congruent (mean: 0.86) and incongruent (0.84) trials] condition. Seemingly, this results from heightened response competition when the two incongruent face images are shown in the two eyes although why this is so remains to be further investigated. There was no interaction of Eye \times Alignment, ($F(1,19) = 0.005$, $p > .9$, $\eta^2_p = .0004$, $BF = 2.93 \times 10^{-13}$), with equivalent accuracy difference between aligned and misaligned trials in both the monocular and interocular trials, and moderate evidence favoring the null model. There was, however, a significant two-way interaction of Congruency \times Alignment, ($F(1,19) = 7.38$, $p < .02$, $\eta^2_p = .291$) (see Fig. 5A, right). Posthoc tests with $p < .05$ revealed significant differences in accuracy between congruent and incongruent trials in both the aligned and the misaligned condition, but a smaller difference between congruent (mean 0.869) and incongruent (mean 0.825) trials in the aligned than between congruent (mean 0.886) and incongruent (mean 0.806) trials in the misaligned condition. Most relevant, however, is that there was no three-way interaction i.e. no modulation of Congruency \times Alignment by Eye, ($F(1,19) = 0.832$, $p > .3$, $\eta^2_p = .044$). The absence of this three-way interaction was confirmed by a $BF = 4.017 \times 10^{-5}$, providing extreme evidence and confirmation that the null model is preferred over the alternate three-way model.

The bottom line is that, for accuracy, the only interaction with the factor Eye was with Congruency. Of note, too, is the facilitation, albeit small, when two faces are presented to the same eye rather than different eyes, the hallmark of a subcortical contribution (but not one that is clearly related to HP).

3.2. d' prime analysis

The results with d' (sensitivity) as the dependent measure revealed no

significant difference for trials with displays presented monocularly (mean 2.345) versus interocularly (mean 2.385), ($F(1,19) = 0.49$, $p = .493$, $\eta^2_p = .026$; $BF = 1.78 \times 10^{-10}$). There was also no significant main effect of Alignment between aligned (mean 2.367), and misaligned (mean 2.363), trials, ($F(1,19) = 0.04$, $p > .9$, $\eta^2_p = .000$; $BF = 1.625 \times 10^{-10}$). There was, however, a significant main effect of Congruency with higher d' for congruent (mean 2.613) than incongruent (mean 2.117) trials, ($F(1,19) = 29.45$, $p < .001$, $\eta^2_p = .621$), again reflecting the interference for the incongruent half on judgments of the cued half of the face.

The two-way interaction of Congruency \times Eye condition was significant, ($F(1,19) = 17.17$, $p < .001$, $\eta^2_p = .488$), with a smaller difference (hence less interference) for the monocular condition between d' for congruent (mean 2.51) than incongruent (mean 2.25) trials than for the interocular condition between d' for congruent (mean 2.71) than incongruent (mean 1.97) trials (see Fig. 5B, left). This result parallels the same Congruency \times Eye interaction in the accuracy data above. The interaction of Alignment \times Eye was not significant, ($F(1,19) = 0.238$, $p > .6$, $\eta^2_p = .013$; $BF = 3.42 \times 10^{-11}$), with no difference between the d' for aligned (mean 2.40) versus misaligned (mean 2.37) in the monocular and the d' for aligned (mean 2.33) versus misaligned (mean 2.36) in the interocular condition. The two-way interaction of Congruency \times Alignment was significant, however, ($F(1,19) = 6.57$, $p < .02$, $\eta^2_p = .267$) (see Fig. 5B, right). Posthoc tests at $p < .05$ revealed a significantly lower d' for incongruent (mean 2.2) than congruent (mean 2.53) trials when the displays were aligned. This was also so for the misaligned trials with an even larger difference in d' for incongruent (2.032) than congruent trials (mean 2.69) than aligned trials. Most relevant, however, is, with d' as the dependent measure, there is no three-way interaction i.e. no modulation of Congruency \times Alignment by Eye, ($F(1,19) = 2.35$, $p > .1$, $\eta^2_p = .116$). The absence of this interaction was further confirmed by a $BF = 9.42 \times 10^{-10}$, supporting the claim that the null model, which does not include Eye, is preferred.

3.3. RT analysis

With RT as the dependent measure, there was no significant difference for trials with displays presented in the monocular (mean 0.793) versus interocular (mean 0.776) conditions, ($F(1,19) = 0.2$, $p = .6$, $\eta^2_p = .012$; $BF = 0.306$). There was, however a significant difference between Alignment conditions, with faster RT for the aligned (mean 0.776) than misaligned (mean 0.793) conditions, ($F(1,19) = 12.749$, $p < .005$, $\eta^2_p = .415$). The main effect of Congruency was also significant with faster RTs

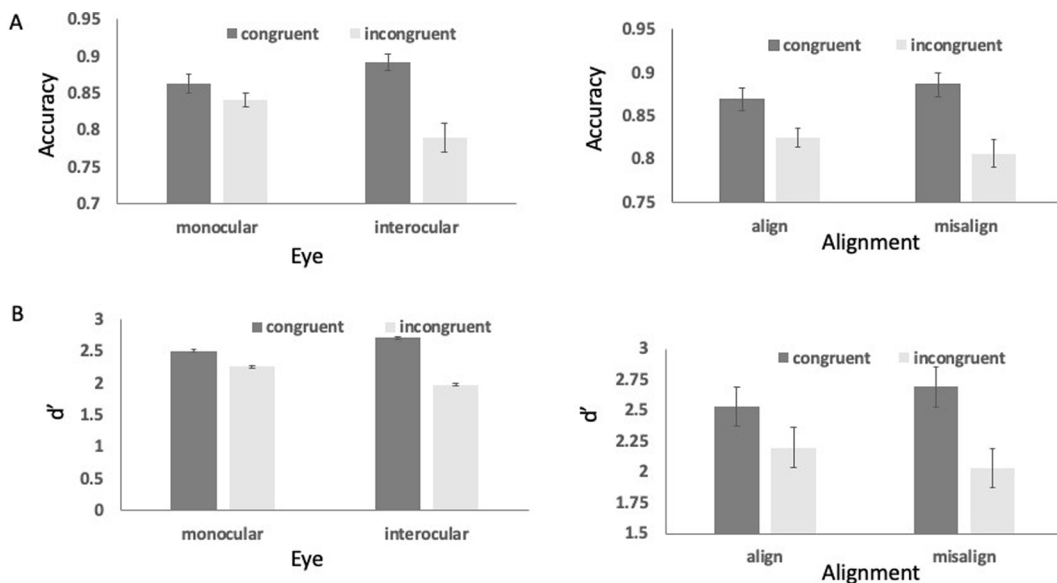


Fig. 5. The signature of holistic processing (HP). A. Accuracy for congruent and incongruent trials as a function of Eye and as a function of Alignment. B. d' for congruent and incongruent trials as a function of Eye and as a function of Alignment.

Table 1
Schematic of statistical effects for the three dependent measures*

	accuracy	d'	RT
Eye	✓	χ	χ
Alignment	χ	χ	✓
Congruency	✓	✓	✓
Eye × Congruency	✓	✓	✓
Eye × Alignment	χ	χ	χ
Congruency × Alignment	✓	✓	χ
Congruency × Alignment × Eye	χ	χ	χ

* ✓ indicates a statistically significant effect; χ indicates no statistically significant effect.

for the congruent (mean 0.766) and incongruent (mean 0.793) trials, ($F(1,19) = 20.599$, $p < .001$, $\eta^2_p = .534$).

There was a significant two-way interaction of Eye × Congruency, with better performance for congruent trials than incongruent trials for the same than for different eyes, ($F(1,19) = 5.86$, $p < .026$, $\eta^2_p = .246$). Posthoc testing at $p < .05$ revealed no difference between congruent and incongruent trials in the monocular condition (both mean 0.793) but a reduction in RT for congruent (mean 0.758) over incongruent (mean 0.794) trials in the interocular condition (see Fig. 5). The interaction of Eye × Alignment was not significant, ($F(1,19) = 0.451$, $p > .5$, $\eta^2_p = .024$; $BF = 0.146$), with equivalent faster RTs for congruent (mean 0.763) over incongruent (mean 0.784) trials in the aligned and also for congruent (mean 0.788) over incongruent (0.803) trials in the misaligned condition. There was no significant two-way interaction of Congruency × Alignment, ($F(1,19) = 0.16$, $p > .6$, $\eta^2_p = .009$) (but, for plot and comparability with accuracy and d', see Fig. 5, right). The absence of this interaction was further confirmed by a $BF = 0.001$, providing very strong confirmation that the null model is preferred. Most relevant, however, is that there is no three-way interaction i.e. no modulation of Congruency × Alignment by Eye, ($F(1,19) = 0.9$, $p > .3$, $\eta^2_p = .048$), and $BF = 3.97 \times 10^{-4}$, favoring the conclusion of extreme evidence for the null model.

To synthesize the results and summarize the findings, we show the results of the F table for all three dependent measures with ✓ indicating a statistically significant effect and χ its absence in Table 1.

4. Discussion

Recent findings have indicated that the monocular portion of the visual system (subcortical structures up to and including the monocular neurons in V1) is engaged during face perception (Gabay, Burlingham, & Behrmann, 2014; Gabay, Nestor, et al., 2014; Mende-Siedlecki et al., 2013). The involvement of this more primitive or evolutionarily older part of the visual system, including structures such as the amygdala, lateral geniculate nucleus, pulvinar and superior colliculus, is consistent with evidence gleaned from both ontogenetic and phylogenetic studies. In light of this evidence, here, we examined specifically whether, in the adult human visual system, these more primitive structures are engaged during the holistic processing (HP) of faces, a computation considered to be critical to face perception (Maurer et al., 2002). Hypothetically, at one extreme, there may be conservation over evolution and over human development, in which case subcortical structures in adult humans may process faces holistically. At the other extreme, as cortex becomes increasingly competent in adulthood, the contribution of subcortical structures could be minimal or perhaps even eliminated. Many other predictions lie between these two extremes hypotheses.

4.1. Assaying HP in the human monocular visual pathway

To examine whether and to what extent monocular portions of the visual system might contribute to HP, we adopted the well-known Composite Face Effect (CFE) paradigm, a gold standard of HP (Richler, Palmeri, & Gauthier, 2012; Rossion, 2013), with adult participants. The

experiment entailed sequential presentation of a pair of displays ('study' and 'test'), each containing two halves (top and bottom) of a face which are either aligned or misaligned. A cue on either the top or bottom half of the study face indicates the half of the face on which the same/different judgement is made across the study and test faces (see Fig. 3). The relationship between the top and bottom halves of the study and test face across the two displays also varies, being congruent or not. Critically, and the novel contribution of this study, was that the CFE paradigm was conducted using a Wheatstone stereoscope that permitted us to present the pair of displays sequentially to the same eye (equally balanced right or left) i.e. monocularly or to a different eye (equally first display to the right or left eye and second display to other eye) i.e. interocularly (see Fig. 4). Importantly, in both of these cases, the participant perceives the displays as being presented centrally (ensured by pre-experimental calibration of fusion) and has no insight into which eye is receiving the input. Although different studies have used different dependent measures (d', accuracy and/or RT) to explore HP in the CFE, we lay out the statistical results in all dependent measures.

The first major result is that there was no evidence for HP in the monocular visual pathway: neither the three-way interaction of Eye × Congruency × Alignment nor the Eye × Alignment were significant in any of the dependent measures, and the absence of these interaction is supported by the Bayes Factor calculated for each measure. The second key result was that there was a robust interaction of Eye × Congruency in all three dependent measures, as well as a main effect of Congruency in all dependent measures. As we discuss below, the effect of Congruency is not considered to be a sufficient marker for HP by many researchers.

Last, the absence of the Alignment effects and interactions with Eye cannot be explained away as a function of an unreliable measure of subcortical processing as the main effect of Eye, with better performance in monocular over interocular pathway, is present, albeit small. The absence of the statistical interactions was also not a result of the failure to evoke HP as there is either an interaction of Alignment × Congruency (accuracy and d') or a main effect of Alignment (RT) across all dependent measures. These Alignment effects are considered clear markers of HP (Richler & Gauthier, 2014; Richler, Mack, Gauthier, & Palmeri, 2009; Richler et al., 2019; Young et al., 1987) and so we were able to demonstrate HP but independent of eye condition.

4.2. What does the effect of Congruency indicate?

An important aspect of our findings that warrants further consideration is the presence of the Congruency effect and the Congruency × Eye interaction. If these effects are insufficient to claim HP in subcortical regions, what do these findings mean? The interpretation of these Congruency effects have been considered in detail by Rossion (2013) in his comprehensive examination of the CFE or Composite Face Illusion, as he calls it, and by Gauthier and her colleagues (for example, (Richler, Wong, & Gauthier, 2011)).

The crux of the argument is that an effect of Congruency is not tantamount to concluding that HP is involved. According to Rossion (2013), Alignment and Congruency tap two different mechanisms: while Alignment does measure perceptual integration and HP, Congruency measures failures of selective attention or response conflict and interference. On Rossion's account, Congruency is more like Stroop interference or congruency effects as in Navon hierarchical or compound letters (for example, a global 'H' made of small 's's'). Intriguingly, a recent study has provided evidence for the global-to-local interference using hierarchical stimuli, with the typical global advantage and global-to-local interference observed for the same-eye but not different-eye condition (Soloveichick, Kimchi, & Gabay, 2021).

The presence of Congruency effects but not Alignment effects is compatible with the notion that different aspects of subcortical function are tapped, with Alignment explaining HP when it is present and Congruency indexing attentional selection and/or interference or response

conflict. Note that Gauthier and colleagues (e.g. Richler, Wong & Gauthier, 2011) suggest that the interference that characterizes the Congruency effects, although not an index of HP initially, with the acquisition of perceptual expertise, can become more indicative of HP. As is evident, the term 'HP' is a multi-splendored construct, as demonstrated by the many operational definitions in the introduction and the debate about the empirical manifestations. A last cautionary comment is that many of the measures purported to analyze HP do not correlate with each other and may be independent (Rezlescu, Susilo, Wilmer, & Caramazza, 2017).

4.3. Reconciling these results with previous studies

At first glance, the absence of HP for faces presented in the monocular visual pathway might be at odds with the findings of face-selective activation in subcortical regions. A straightforward way to reconcile these results would be one in which face-selective activation and behavioral engagement do, indeed, occur subcortically, but this is not achieved by virtue of HP. Rather, a more feature-based strategy in which, for example, matching eyes or mouth across images may be within the purview of subcortical regions. Another strategy, one based on 'first order' properties of faces i.e., two eyes, above a nose, above a mouth might be in play, with representation of the face parts independently is also a possibility. Such a computation may suffice for matching and for driving subcortical regions without invoking HP. Potentially, a new experiment in which the faces are familiar to the observer or are faces of well-known celebrities, as in the initial Young et al. (1987) study might be illuminating. Alternatively, a study in which identity recognition is required might similarly shed light on the nature and limitations of the subcortical computations.

Rejecting subcortical accounts of holistic face processing initially seems in disagreement with present research on the "blindsight" phenomenon and cortical blindness. Such patients are able to respond to or show activation for visual stimuli while denying seeing them (Celesia, 2010). A cortically blind patient with bilateral visual cortex destruction showed amygdala activation for directed over averted gaze stimuli though he was at chance in guessing the direction, and his amygdala responded to the low but not high spatial frequency components of fearful faces (Burra et al. 2013, Burra et al., 2019). In this case, however, feature-by-feature processing of the eyes, and not HP, would be consistent with the results, as only perception of the eyes was necessary for this task.

Another cortically blind patient with Anton's syndrome maintained the ability to distinguish between jumbled and normal faces at over 80% accuracy while being unable to explicitly recognize or describe them, and visual contrast could not account for the distinction. He further was able to maintain eye contact with his examiners. Again, a feature-based or first-order subcortical representation may have also sufficed in this case assuming features were preserved in the process of jumbling (Solcà, Guggisberg, Schnider, & Leemann, 2015).

That we do not obviously observe HP in the monocular pathway may fit with findings of studies of phylogeny in which face-specific effects are sometimes presented as a face detector or even an eye detector (Pascalis & Kelly, 2009). This may also apply to the inversion effects in infancy and in low-order species that do not have a cortex with 'first-order' information supporting subcortical face activation, in adult humans. Neither first- nor second-order (or relational) information necessarily reflect the perceptual integration of all parts of the face, and it is this integration that may ultimately be the imprimatur of HP, and tapped by the composite face paradigm.

5. Conclusions

Despite the apparent engagement of evolutionarily older parts of the human visual system in face perception, the findings from the current investigation do not support the role of these subcortical structures in

holistic processing, HP, as computed by the human visual system. The findings do support the claim that there is a subcortical contribution to face processing, consistent with previous demonstrations (Gabay, Burlingham, et al., 2014; Gabay, Nestor, et al., 2014; Mende-Siedlecki et al., 2013), but do not indicate that holistic processing of faces per se is necessarily subcortically-based. Although newborns and insects without cortex can distinguish between face-like stimuli and even evince the inversion decrement when faces are rotated away from upright, these perceptual feats are likely not equivalent to the HP subserved by the mature human face system. Furthermore, the fact that individuals with blindsight are able to process faces (and show selective activation in imaging) may be a product of feature-based processing or some sensitivity to the first-order position, for example of the eyes above the nose, and not the result of HP.

The present results, therefore, suggest that HP is probably a property of cortical rather than subcortical (and/or V1) computation. Consistent with this are results showing that, when participants viewed natural or schematic faces with internal parts that were either normally configured or randomly rearranged, the activation profile of the amygdala and superior temporal sulcus, but not the fusiform face area, were reduced, likely reflecting the tuning of these areas to the features rather than the whole (Golarai, Ghahremani, Eberhardt, & Gabrieli, 2015). Responses in the fusiform face area were only reduced when both the parts and the configuration of the face were perturbed, specifically implicating this region in holistic processing. Similar results were obtained using electroencephalographic frequency tagging showing a dissociation between part-based and holistic responses to faces with the holistic representations confined to high-level visual regions (Boremanse, Norcia, & Rossion, 2014). The absence of engagement of early visual cortex in holistic processing is consistent with our findings of no differential composite effect in the monocular versus interocular portions of the visual system.

CRedit authorship contribution statement

Rebeka C. Almasi: Conceptualization, Methodology, Software, Investigation, Data curation, Writing - original draft, Writing - review & editing. **Marlene Behrmann:** Conceptualization, Methodology, Supervision, Writing - review & editing.

Acknowledgements and Funding

This research was supported by a grant from National Institutes of Health (NEI grant RO1EY027018 to M.B.) and was conducted by R. A. in partial fulfillment of the Honor's thesis for Dietrich College, Carnegie Mellon University. The authors thank Dr James Tanaka for providing us with the face stimuli, Dr Tina Liu for sharing her code and Drs Shai Gabay and Isabel Gauthier for the discussion of the work.

References

- Andrews, T. J., Davies-Thompson, J., Kingstone, A., & Young, A. W. (2010). Internal and external features of the face are represented holistically in face-selective regions of visual cortex. *Journal of Neuroscience*, 30(9), 3544–3552. <https://doi.org/10.1523/JNEUROSCI.4863-09.2010>.
- Atkinson, J. (2000). *The Developing Visual Brain*. Oxford, UK: Oxford University Press.
- Avargues-Weber, A., d'Amaro, D., Metzler, M., Finke, V., Baracchi, D., & Dyer, A. G. (2018). Does holistic processing require a large brain? Insights from honeybees and wasps in fine visual recognition tasks. *Frontiers in Psychology*, 9, 1313.
- Avargues-Weber, A., Portelli, G., Benard, J., Dyer, A., & Giurfa, M. (2010). Configural processing enables discrimination and categorization of face-like stimuli in honeybees. *Journal of Experimental Biology*, 213(4), 593–601. <https://doi.org/10.1242/jeb.039263>.
- Avidan, G., Tanzer, M., & Behrmann, M. (2011). Impaired holistic processing in congenital prosopagnosia. *Neuropsychologia*, 49(9), 2541–2552. <https://doi.org/10.1016/j.neuropsychologia.2011.05.002>.
- Blake, R., & Cormack, R. H. (1979). Psychophysical evidence for a monocular visual cortex in stereoblind humans. *Science*, 203(4377), 274.
- Boremanse, A., Norcia, A. M., & Rossion, B. (2014). Dissociation of part-based and integrated neural responses to faces by means of electroencephalographic frequency tagging. *European Journal of Neuroscience*, 40(6), 2987–2997. <https://doi.org/10.1111/ejn.12663>.

- Burra, N., Hervais-Adelman, A., Celeghin, A., De Gelder, B., & Pegna, A. J. (2019). Affective blindsight relies on low spatial frequencies. *Neuropsychologia*, 128, 44–49.
- Burra, N., Hervais-Adelman, A., Kerzel, D., Tamietto, M., De Gelder, B., & Pegna, A. J. (2013). Amygdala activation for eye contact despite complete cortical blindness. *Journal of Neuroscience*, 33(25), 10483–10489.
- Callaway, E. M. (2005). Structure and function of parallel pathways in the primate early visual system. *Journal of Physiology*, 566(Pt 1), 13–19. <https://doi.org/10.1113/jphysiol.2005.088047>.
- Cassia, V. M., Turati, C., & Simion, F. (2004). Can a nonspecific bias toward top-heavy patterns explain newborns' face preference? *Psychological Science*, 15(6), 379–383.
- Cheung, O. S., Richler, J. J., Phillips, W. S., & Gauthier, I. (2011). Does temporal integration of face parts reflect holistic processing? *Psychonomic Bulletin & Review*, 18(3), 476–483. <https://doi.org/10.3758/s13423-011-0051-7>.
- Celesia, G. G. (2010). Visual perception and awareness: A modular system. *Journal of Psychophysiology*, 24(2), 62–67. <https://doi.org/10.1027/0269-8803/a000014>.
- Dalrymple, K. A., Khan, A. F., Duchaine, B., & Elison, J. T. (2020). Visual input to the left versus right eye yields differences in face preferences in 3-month-old infants. *Developmental Science*, e13029. <https://doi.org/10.1111/desc.13029>.
- Dyer, A. G., Neumeyer, C., & Chittka, L. (2005). Honeybee (apis mellifera) vision can discriminate between and recognise images of human faces. *The Journal of Experimental Biology*, 208, 4709–4717.
- Farah, M. J., Wilson, K. D., Drain, M., & Tanaka, J. N. (1998). What is “special” about face perception? *Psychological Review*, 105(3), 482–498.
- Farroni, T., Johnson, M. H., Menon, E., Zulian, L., Faraguna, D., & Csibra, G. (2005). Newborns' preference for face-relevant stimuli: Effects of contrast polarity. *Proceedings of the National Academy of Sciences of the United States of America*, 102(47), 17245–17250. <https://doi.org/10.1073/pnas.0502205102>.
- Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, 39(2), 175–191. <https://doi.org/10.3758/bf03193146>.
- Gabay, S., Burlingham, C., & Behrmann, M. (2014). The nature of face representations in subcortical regions. *Neuropsychologia*, 59, 35–46. <https://doi.org/10.1016/j.neuropsychologia.2014.04.010>.
- Gabay, S., Nestor, A., Dundas, E., & Behrmann, M. (2014). Monocular advantage for face perception implicates subcortical mechanisms in adult humans. *Journal of Cognitive Neuroscience*, 26(5), 927–937. <https://doi.org/10.1162/jocn.a.00528>.
- Golarai, G., Ghahremani, D. G., Eberhardt, J. L., & Gabrieli, J. D. (2015). Distinct representations of configural and part information across multiple face-selective regions of the human brain. *Frontiers in Psychology*, 6, 1710. <https://doi.org/10.3389/fpsyg.2015.01710>.
- Hole, G. J. (1994). Configurational factors in the perception of unfamiliar faces. *Perception*, 23, 65–74.
- 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(7), 1025.
- Johnson, M. H. (2005). Subcortical face processing. *Nature Reviews Neuroscience*, 6(10), 766–774. <https://doi.org/10.1038/nrn1766>.
- Johnson, M. H., Dziurawiec, S., Ellis, H., & Morton, J. (1991). Newborns' preferential tracking of face-like stimuli and subsequent decline. *Cognition*, 40, 1–19.
- Johnson, M. H., Senju, A., & Tomalski, P. (2015). The two-process theory of face processing: Modifications based on two decades of data from infants and adults. *Neuroscience & Biobehavioral Reviews*, 50C, 169–179. <https://doi.org/10.1016/j.neubiorev.2014.10.009>.
- Kveraga, K., Boshyan, J., & Bar, M. (2007). Magnocellular projections as the trigger of top-down facilitation in recognition. *Journal of Neuroscience*, 27(48), 13232–13240. <https://doi.org/10.1523/JNEUROSCI.3481-07.2007>.
- LeDoux, J. E. (1994). Emotion, memory and the brain. *Scientific American*, 270(6), 50–57. <https://doi.org/10.1038/scientificamerican069450>.
- Lee, Michael D., & Wagenmakers, Eric-Jan (2014). *Bayesian cognitive modeling: A practical course*. Cambridge university press.
- Liu, T. T., & Behrmann, M. (2014). Impaired holistic processing of left-right composite faces in congenital prosopagnosia. *Frontiers in Human Neuroscience*, 8, 750. <https://doi.org/10.3389/fnhum.2014.00750>. Retrieved from <https://www.frontiersin.org/article/10.3389/fnhum.2014.00750>.
- Maurer, D., Grand, R. L., & Mondloch, C. J. (2002). The many faces of configural processing. *Trends in Cognitive Sciences*, 6(6), 255–260.
- Maurer, D., O'Craven, K. M., Le Grand, R., Mondloch, C. J., Springer, M. V., Lewis, T. L., et al. (2007). Neural correlates of processing facial identity based on features versus their spacing. *Neuropsychologia*, 45(7), 1438–1451. <https://doi.org/10.1016/j.neuropsychologia.2006.11.016>.
- 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(12), 2086–2106. <https://doi.org/10.1162/jocn.a.00469>.
- Menon, R. S., Ogawa, S., Strupp, J. P., & Ugurbil, K. (1997). Ocular dominance in human V1 demonstrated by functional magnetic resonance imaging. *Journal of Neurophysiology*, 77(5), 2780–2787.
- Miller, S. E., Legan, A. W., Henshaw, M. T., Ostevik, K. L., Samuk, K., Uy, F. M. K., & Sheehan, M. J. (2019). Evolutionary dynamics of recent selection of cognitive abilities. *Proceedings of the National Academy of Sciences of the United States of America*, www.pnas.org/cgi/doi/10.1073/pnas.1918592117.
- Morton, J., & Johnson, M. H. (1991). CONSPEC and CONLERN: A two-process theory of infant face recognition. *Psychological Review*, 98(2), 164–181.
- Pascalis, O., & Kelly, D. J. (2009). The Origins of Face Processing in Humans: Phylogeny and Ontogeny. *Perspectives on Psychological Science*, 4(2), 200–209. <https://doi.org/10.1111/j.1745-6924.2009.01119.x>.
- Rezlescu, C., Susilo, T., Wilmer, J. B., & Caramazza, A. (2017). The Inversion, Part-Whole, and Composite Effects Reflect Distinct Perceptual Mechanisms With Varied Relationships to Face Recognition. *Journal of Experimental Psychology: Human Perception and Performance*. <https://doi.org/10.1037/xhp0000400>.
- Richler, J. J., Cheung, O. S., & Gauthier, I. (2011). Holistic processing predicts face recognition. *Psychological Science*, 22(4), 464–471. <https://doi.org/10.1177/0956797611401753>.
- Richler, J. J., Cheung, O. S., Wong, A. C., & Gauthier, I. (2009). Does response interference contribute to face composite effects? *Psychonomic Bulletin & Review*, 16(2), 258–263. <https://doi.org/10.3758/PBR.16.2.258>.
- Richler, J. J., & Gauthier, I. (2014). A meta-analysis and review of holistic face processing. *Psychological Bulletin*, 140(5), 1281–1302. <https://doi.org/10.1037/a0037004>.
- Richler, J. J., Mack, M. L., Gauthier, I., & Palmeri, T. J. (2009). Holistic processing of faces happens at a glance. *Vision Research*, 49(23), 2856–2861. <https://doi.org/10.1016/j.visres.2009.08.025>.
- Richler, J. J., Mack, M. L., Palmeri, T. J., & Gauthier, I. (2011). Inverted faces are (eventually) processed holistically. *Vision Research*, 51(3), 333–342. <https://doi.org/10.1016/j.visres.2010.11.014>.
- 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>.
- Richler, J. J., Tomarken, A. J., Sunday, M. A., Vickery, T. J., Ryan, K. F., Floyd, R. J., et al. (2019). Individual differences in object recognition. *Psychological Review*, 126(2), 226–251. <https://doi.org/10.1037/rev0000129>.
- Richler, J. J., Wong, Y. K., & Gauthier, I. (2011). Perceptual Expertise as a Shift from Strategic Interference to Automatic Holistic Processing. *Current Directions in Psychological Science*, 20(2), 129–134. <https://doi.org/10.1177/0963721411402472>.
- Rossion, B. (2008). Picture-plane inversion leads to qualitative changes of face perception. *Acta Psychologica (Amst)*, 128(2), 274–289. <https://doi.org/10.1016/j.actpsy.2008.02.003>.
- Rossion, B. (2013). The composite face illusion: A whole window into our understanding of holistic face perception. *Visual Cognition*, 21(2), 139–253.
- Rossion, B., Dricot, L., Devolder, A., Bodart, J. M., Crommelinck, M., Gelder, B. D., & Zoontjes, R. (2000). Hemispheric asymmetries for whole-based and part-based face processing in the human fusiform gyrus. *Journal of Cognitive Neuroscience*, 12(5), 793–802.
- Scherf, K. S., Whyte, E. M., Minshew, N., & Behrmann, M. (2020). Training up holistic processing in autism: Converging longitudinal intervention studies. *Journal of Cognitive Neuroscience*. in press subject to revision.
- Schiltz, C., Dricot, L., Goebel, R., & Rossion, B. (2010). Holistic perception of individual faces in the right middle fusiform gyrus as evidenced by the composite face illusion. *The Journal of Visual Communication and Image Representation*, 10(2). <https://doi.org/10.1167/10.2.25>, 25.1–16.
- Schiltz, C., & Rossion, B. (2006). Faces are represented holistically in the human occipito-temporal cortex. *Neuroimage*, 32(3), 1385–1394. <https://doi.org/10.1016/j.neuroimage.2006.05.037> (Epub 2006 Jul 25. PMID: 16870475).
- Schiltz, C., & Rossion, B. (2006). Faces are represented holistically in the human occipito-temporal cortex. *Neuroimage*, 32(3), 1385–1394. <https://doi.org/10.1016/j.neuroimage.2006.05.037>.
- Schneider, K. A., & Kastner, S. (2005). Visual responses of the human superior colliculus: A high-resolution functional magnetic resonance imaging study. *Journal of Neurophysiology*, 94(4), 2491–2503. <https://doi.org/10.1152/jn.00288.2005>.
- Schwarzkopf, D. S., Schindler, A., & Rees, G. (2010). Knowing with which eye we see: Utricular discrimination and eye-specific signals in human visual cortex. *PLoS ONE*, 5(10), Article e13775.
- Sheehan, M. J., & Tibbetts, E. A. (2011). Specialized face learning is associated with individual recognition in paper wasps. *Science*, 334(6060), 1272–1275. <https://doi.org/10.1126/science.1211334>.
- Simion, F., Valenza, E., Umiltà, C., & Dalla Barba, B. (1998). Preferential orienting to faces in newborns: A temporal-nasal asymmetry. *Journal of Experimental Psychology: Human Perception and Performance*, 24(5), 1399–1405.
- Solcà, M., Guggisberg, A. G., Schnider, A., & Leemann, B. (2015). Facial blindsight. *Frontiers in Human Neuroscience*, 9, 522.
- Soloveichick, M., Kimchi, R., & Gabay, S. (2021). Functional involvement of subcortical structures in global-local processing. *Cognition*, 206, 2021. <https://doi.org/10.1016/j.cognition.2020.104476>.
- Stein, T., Seymour, K., Hebart, M. N., & Sterzer, P. (2014). Rapid fear detection relies on high spatial frequencies. *Psychological Science*, 25(2), 566–574. <https://doi.org/10.1177/0956797613512509>.
- Tamietto, M., Cauda, F., Corazzini, L. L., Savazzi, S., Marzi, C. A., Goebel, R., et al. (2010). Collicular Vision Guides Nonconscious Behavior. *Journal of Cognitive Neuroscience*, 22(5), 888–902. <https://doi.org/10.1162/jocn.2009.21225>.
- Taubert, J., et al. (2018). Amygdala Lesions Eliminate Viewing Preferences for Faces in Rhesus Monkeys. *Proceedings of the National Academy of Sciences*, 115(31), 8043–8048. <https://doi.org/10.1073/pnas.1807245115>.
- Tibbetts, E. A. (2002). Visual signals of individual identity in the wasp *Polistes fuscatus*. *Proceedings of the Royal Society B: Biological Sciences*, 269(1499), 1423–1428. <https://doi.org/10.1098/rspb.2002.2031>.
- Turati, C., Sangrigoli, S., Ruel, J., & de Schonen, S. (2004). Evidence of the face inversion effect in 4-month old infants. *Infancy*, 6(2), 275–297.
- Van den Stock, J., Tamietto, M., Zhan, M., Heinecke, A., Hervais-Adelman, A. G., Legrand, L., et al. (2014). Neural correlates of body and face perception following bilateral destruction of the primary visual cortices. *Frontiers in Behavioral Neuroscience*, 8, 30.

- Vuilleumier, P., Armony, J. L., Driver, J., & Dolan, R. J. (2003). Distinct spatial frequency sensitivities for processing faces and emotional expressions. *Nature Neuroscience*, 6 (6), Article 624631.
- Wheatstone, C. (1838). Contributions to the physiology of vision. Part the first: On some remarkable and unobserved phenomena of binocular vision. *Philosophical Transactions of the Royal Society*, 128, 371–394.
- Yin, R. K. (1969). Looking at upside-down faces. *Journal of Experimental Psychology*, 81 (1), 141–145. <https://doi.org/10.1037/h0027474>.
- Young, A. W., Hellawell, D., & Hay, D. C. (1987). Configurational information in face perception. *Perception*, 16, 747–759.