Revealing the mechanisms of human face perception using dynamic apertures

Jennifer Murphy[1,*] & Richard Cook[2]

[1]MRC Social, Genetic and Developmental Psychiatry Centre,
Institute of Psychiatry, Psychology, and Neuroscience,
King's College London, London, U.K.

[2]Department of Psychological Sciences,
Birkbeck, University of London, London, U.K.

*Corresponding author:
Jennifer.Murphy@kcl.ac.uk
MRC Social, Genetic and Developmental Psychiatry Centre,
Institute of Psychiatry, Psychology and Neuroscience
De Crespigny Park, Denmark Hill,
London, U.K, SE5 8AF

Highlights

- We sought to compare the visual processing of upright and inverted faces
- In one condition observers categorized faces presented in their entirety
- In the other, a dynamic aperture moved incrementally across the facial image
- Whole-face processing improved the perception of both upright and inverted faces

Abstract

Faces are notoriously hard to perceive when turned upside-down. It is often claimed that perceptual decrements reflect a qualitative switch from parallel whole-face processing, to a serial analysis of individual features. To test this view, we examined observers' ability to categorize faces presented in their entirety, or viewed through a dynamic aperture that moved incrementally across the facial image. Exposing faces region-by-region disrupts holistic processing, but permits serial analysis of local features. In line with the holistic accounts, we predicted that aperture viewing would greatly impair judgements of upright, but not inverted faces. As expected, identity, gender, age, and expression were categorized more precisely when faces were presented upright and in their entirety. Contrary to holistic accounts, however, the detrimental effects of inversion seen in the whole-face condition were no greater than in the aperture condition. Moreover, we found comparable aperture effects for upright and inverted faces; observers exhibited less decision noise when faces were viewed in their entirety, than when viewed through the aperture, irrespective of orientation. We replicate these findings in control experiments and show that the same pattern is seen irrespective of the direction of aperture transition or the nature of the fill used to replace the occluded regions of the to-be-judged image. These results challenge holistic accounts of the face inversion effect and support an alternative interpretation. First, in line with previous findings, they indicate that perceptual decrements when viewing upside-down faces result from impoverished descriptions of local regions, not the loss of whole-face processing. Second, when interpreting inverted faces, access to the wider face context appears to be far more important than currently believed.

1. Introduction

1.1 Holistic face processing

Theories of holistic face processing argue that upright and inverted faces recruit qualitatively different visual processing; local features are thought to be integrated into a unified whole when observers view upright faces, whereas inverted faces are thought to recruit a 'piecemeal' or 'parts-based' analysis of local features (Farah, Wilson, Drain, & Tanaka, 1998; Maurer, Le Grand, & Mondloch, 2002; McKone, Kanwisher, & Duchaine, 2007; McKone & Yovel, 2009; Piepers & Robbins, 2013; Rossion, 2008). Holistic processing is thought to be causally related to face recognition ability (Farah et al., 1998; Maurer et al., 2002; Piepers & Robbins, 2013) as well as perceptual expertise more broadly (Richler, Wong, & Gauthier, 2011), and diminished holistic processing is frequently cited as the cause of face recognition difficulties seen in autism spectrum disorder (Behrmann, Thomas, & Humphreys, 2006) and prosopagnosia (Avidan, Tanzer, & Behrmann, 2011; DeGutis, Cohan, & Nakayama, 2014). Holistic processing has often been linked with the superior representation of inter-feature spatial relations (Rossion, 2008; Searcy & Bartlett, 1996). More recently, however, it has been suggested that holistic processing improves the perception of local features as well as their configuration (Farah et al., 1998; Hayward, Crookes, Chu, Favelle, & Rhodes, 2016; Rhodes, Hayward, & Winkler, 2006; Yovel & Kanwisher, 2008).

Proponents of holistic face processing have been criticized for stating their assumptions informally (Fifić & Townsend, 2010; Fitousi, 2015, 2016; Wenger & Townsend, 2001). Nevertheless, some important features of the hypothesized information processing characteristics may be delineated. For example, its characterization as 'the simultaneous integration of the multiple parts of a face into a single perceptual representation' (e.g., Rossion, 2008, 2013) suggests that upright faces are processed by multiple parallel channels, each describing a particular local region. The activations of these parallel channels may be subsequently combined in a single output channel that conveys an integrated representation of the whole face, upon which perceptual decisions are based (see coactivation architecture; e.g., Fifić & Townsend, 2010). Assertions that integrated representations are 'more than the sum of their parts' (e.g., Shen & Palmeri, 2015) imply that lateral interactions between different channels may improve processing accuracy and efficiency for upright faces. In contrast, the characterization of inverted face processing as

"parts-based" or "piecemeal" suggests that decisions about inverted faces depend on evidence accumulated through a serial analysis of local features. In the absence of lateral interactions between parallel channels, the description of one region remains unaffected by the processing of other regions; i.e., processing independence is hypothesized.

In its strongest form, holistic processing theory argues that parallel whole-face processing is engaged only in the presence of an intact whole face[1] (Farah et al., 1998; McKone & Yovel, 2009; Tanaka & Farah, 1993; Tsao & Livingstone, 2008). Such gating may increase neurocognitive efficiency by ensuring that resource-intensive processing is engaged only where appropriate (Tsao & Livingstone, 2008). The possibility that holistic face processing is dependent on the detection of canonical first-order facial information has been used to explain diminished integration of information from upper and lower face halves when composite face arrangements are spatially misaligned or inverted (Murphy, Gray, & Cook, 2016). This view is also consistent with reports that judgments about cropped facial features presented in isolation (i.e., in the absence of a facial context) are relatively insensitive to orientation inversion (Palermo et al., 2011; Rhodes, Brake, & Atkinson, 1993), however this remains controversial (see Leder, Candrian, Huber, & Bruce, 2001; Rakover & Teucher, 1997). Finally, it accords with findings that scrambled faces – where local features do not appear in their typical locations – do not elicit putative markers of holistic processing, including inversion effects (Martini, McKone, & Nakayama, 2006; Tanaka & Farah, 1993; Towler, Parketny, & Eimer, 2015).

1.2 Composite face effects
To date, the principal line of evidence for the holistic processing account comes from the composite face effect (Hole, 1994; Young, Hellawell, & Hay, 1987). When the top half of one face is aligned with the bottom half of another, and presented upright, the halves appear to 'fuse' resulting in a compelling percept of a novel facial configuration (Murphy et al., 2016; Rossion, 2013). When composite arrangements are presented upside-down, however, fusion is greatly diminished (McKone et al., 2013; Susilo, Rezlescu, & Duchaine, 2013). This effect suggests that the visual system integrates information from disparate regions when observers view upright, but not inverted, faces, consistent with theories of holistic face processing (Farah et al., 1998; Maurer et al., 2002; McKone et al., 2007; McKone & Yovel, 2009; Piepers & Robbins, 2013; Rossion, 2008).

In recent years, however, the functional significance of the composite face effect has proved controversial (Murphy et al., 2016; Richler, Wong et al., 2011; Rossion, 2013). In particular, it remains unclear whether susceptibility to this illusion is related to face recognition ability. While some studies have found that differences in susceptibility correlate with face recognition ability (DeGutis, Wilmer, Mercado, & Cohan, 2013; Richler, Cheung, & Gauthier, 2011b), others have found little or no evidence of association (Konar, Bennett, & Sekuler, 2010; Rezlescu, Susilo, Wilmer, & Caramazza, 2017; Wang, Li, Fang, Tian, & Liu, 2012). Notably, many individuals with acquired (Finzi, Susilo, Barton, & Duchaine, 2016) and developmental (Biotti et al., in press; Le Grand et al., 2006; Susilo et al., 2010) prosopagnosia exhibit typical effects, suggesting a complex relationship between illusion susceptibility and face recognition ability.

The literature on the composite face effect has also been dogged by methodological issues. For example, the size and pattern of composite effects may be affected by the presence or absence of a small gap between the target and distractor halves (Rossion & Retter, 2015), the presence of subtle signs of facial emotion in distractor regions (Gray, Murphy, Marsh, & Cook, 2017), and observers' attentional strategy (Fitousi, 2016). Moreover, there has been considerable debate about the respective merits of the original matching design and a more recent congruency variant (DeGutis et al., 2013; Murphy et al., 2016; Richler, Wong et al., 2011; Rossion, 2013). Whilst the size of composite effects measured using the original design may be affected by response bias (Richler, Cheung, & Gauthier, 2011a), estimates of susceptibility obtained using the congruency paradigm may be contaminated by general factors including response priming and interference (Rossion, 2013).

1.3 Aperture paradigms
Given the controversy surrounding the composite face effect, it is important that complementary tests of holistic face processing are developed. One promising line of innovation uses aperture viewing to compare the visual processing recruited by upright and inverted faces. Aperture paradigms restrict observers' field-of-view, such that only a small part of an image is visible. During aperture viewing, participants must therefore base perceptual decisions on information extracted from exposed local regions, either viewed serially (e.g., in dynamic aperture paradigms) or in isolation (e.g., in static aperture

paradigms). Similar aperture techniques have been used elsewhere in the vision sciences to investigate a range of topics including shape perception, object recognition, and reading (Anstis & Atkinson, 1967; Craddock, Martinovic, & Lawson, 2012; McConkie & Rayner, 1975; Morgan, Findlay, & Watt, 1982; Rieger, Grüschow, Heinze, & Fendrich, 2007; Rock, 1981).

In the context of face perception research, authors have employed aperture techniques to address two questions. First, aperture methods, including 'Bubbles' (e.g., Gosselin & Schyns, 2001) and reverse correlation methods (e.g., Sekuler, Gaspar, Gold, & Bennett, 2004), have been used to reveal which facial regions are informative when making particular judgments (Gosselin & Schyns, 2001; Haig, 1985; Sekuler et al., 2004). Studies in this tradition have repeatedly highlighted the value of the information contained within the eye-region (for discussion see Rossion, 2008). Second, aperture techniques have been used to investigate holistic face processing (Evers, Van Belle, Steyaert, Noens, & Wagemans, 2017; Gold, Mundy, & Tjan, 2012; Tanaka & Farah, 1993; Van Belle, De Graef, Verfaillie, Busigny, & Rossion, 2011; Van Belle, De Graef, Verfaillie, Rossion, & Lefèvre, 2010). Crucially, because aperture viewing prevents access to the whole-face, these manipulations block or reduce observers' ability to process faces holistically. By comparing performance in aperture and whole-face viewing conditions, researchers can directly assess the contribution of holistic processing when observers judge upright and inverted faces (Tanaka & Farah, 1993; Van Belle et al., 2011).

Employing this logic, a recent study compared the perception of upright and inverted faces using a gaze-contingent aperture paradigm, in which the to-be-revealed portion of the target face is determined by the fixation behavior of the observer (Van Belle et al., 2010). Consistent with holistic accounts, observers show marked inversion effects only when judging the whole face; when basing decisions on local information sampled serially through the aperture, observers' upright and inverted performance was broadly comparable. Gaze-contingent aperture paradigms potentially allow observers to fixate on the stimulus naturally. It is not clear, however, how the aperture manipulation alters observers' fixation strategies; many observers appear to fixate the regions between features more often in the aperture conditions, relative to free-viewing conditions (Van Belle et al., 2010). Moreover, gaze-dependent paradigms relinquish control of stimulus

presentation to the participant; different participants receive different visual input depending on their fixation strategies. Consequently, manipulations of facial orientation (Van Belle et al., 2010) and observer group (Evers et al., 2017; Van Belle et al., 2011), may be confounded with the basic visual information presented to observers.

1.4 Present study

In the present study we sought to test the predictions of holistic face processing theory using a novel dynamic aperture paradigm in which participants were unable to influence the speed or trajectory of the viewing window. Observers made simple binary categorization judgments about faces presented in their entirety, or viewed through a dynamic aperture that moved incrementally across the image. Observers' perceptual ability in each viewing condition (their 'decision noise') was inferred from the slope of the resulting psychometric function. If our expert processing of upright faces is a product of holistic processing, aperture viewing should result in marked performance decrements. Conversely, judgments about inverted faces – thought to depend on a serial part-based analysis should be relatively unaffected by aperture viewing.

2. Materials and methods

2.1 Participants

In all experiments, observers were healthy adults with normal or corrected-to-normal vision. Participants were screened for developmental prosopagnosia using the PI20 self-report instrument (Shah, Gaule, Sowden, Bird, & Cook, 2015). Sample size was determined *a priori*, informed by previous psychophysical investigations of face processing (e.g., Cook, Aichelburg, & Johnston, 2015; Cook, Brewer, Shah, & Bird, 2014; Shah, Bird, & Cook, 2016). Participants whose response function could not be modelled using a cumulative Gaussian were replaced and their data not subject to further analysis. Ethical clearance was granted by the local ethics committee and the study was conducted in line with the ethical guidelines laid down in the 6th (2008) Declaration of Helsinki. All participants gave informed consent and were fully debriefed upon task completion.

2.2 Stimuli

Video stimuli presented a single facial image drawn from morph continua (Figure 1), either upright or inverted. Continua consisted of seven levels that varied attribute strength

from 20% to 80% in increments of 10%. Images were morphed using Morpheus Photo Morpher Version 3.11 (Morpheus Software, Indianapolis, IN). Bitmap image sequences were compiled in Matlab (The MathWorks, Natick, MA) and saved as .avi files. Video stimuli were presented at 30 frames-per-second on a LCD display (60 Hz refresh rate).
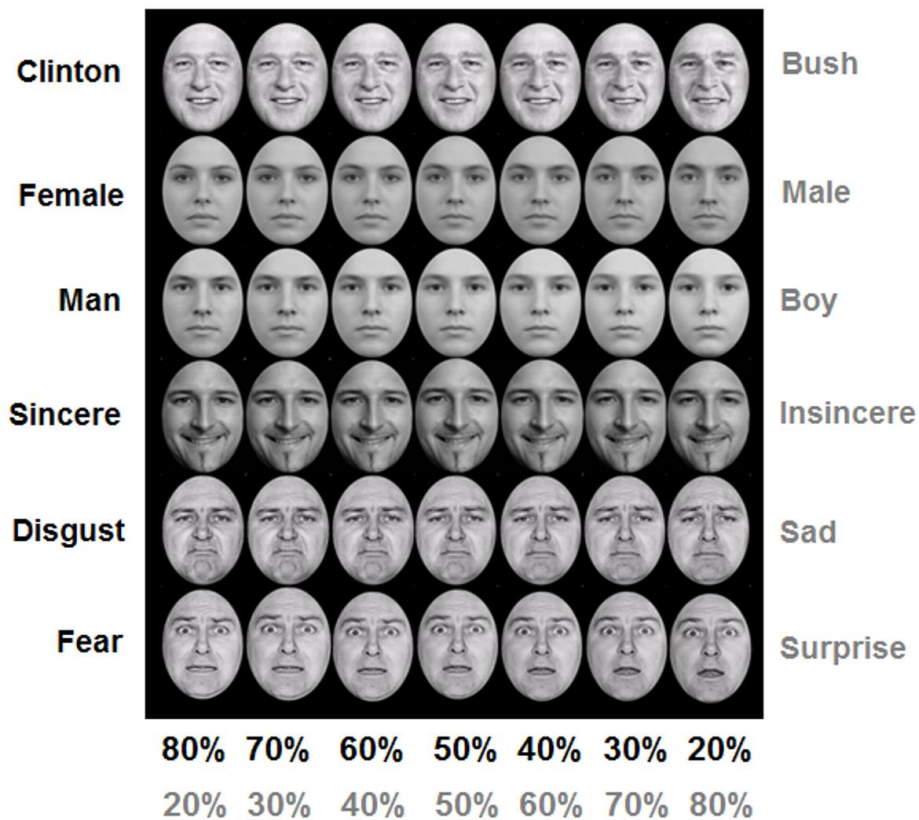

Figure 1. Video stimuli presented a single facial image drawn from morph continua, either upright or inverted. The continua consisted of seven levels that varied attribute strength from 20% to 80% in increments of 10%.

The identity continuum was created by morphing images of George W. Bush and Bill Clinton (sourced online). The age continuum was created by morphing a composite of eight male child faces with a similar composite of eight adult faces. The gender continuum was created by morphing a composite of eight male faces with a similar composite of eight female faces. The original images for the age and gender morphs were sourced from the Radboud Face Database (Langner et al., 2010). The smile-sincerity (Ipser & Cook, 2016), as well as the Surprise-Fear and Disgust-Sadness continua (Biotti & Cook, 2016), have been described elsewhere.

2.3 Procedure

In the aperture condition, a viewing window moved over the image (Figure 2a). The aperture was 12.5% the height of the face (~6.5° of visual arc when viewed at 58 cm) and

took 6 secs to move across the face. In the whole-face condition, faces were presented in their entirety for 1000 ms (Figure 2b). On each trial participants were presented with a single video stimulus and were required to make a binary categorization judgment about the face presented (e.g., "Bush or Clinton?" when judging identity). Following stimulus offset, a response screen was visible until a response was registered. The procedure comprised 448 experimental trials in total (4 conditions × 16 presentations × 7 stimulus levels). Trials for the four conditions (upright whole-face, inverted whole-face, upright aperture, inverted aperture) were randomly interleaved within four mini-blocks of 112 trials. All experiments were programmed in MATLAB using the Psychophysics Toolbox (Brainard, 1997; Pelli, 1997).
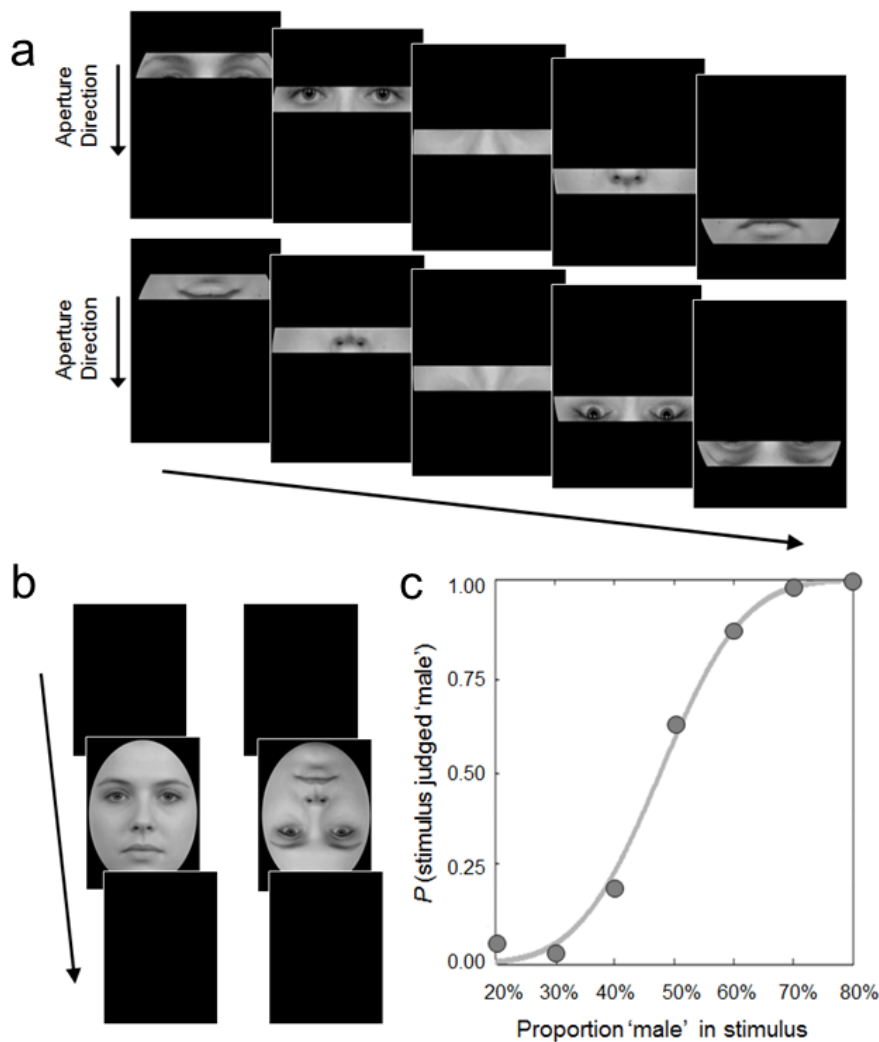

Figure 2. (a) In the aperture condition, a viewing window moved steadily over the image, taking 6 secs to move from top to bottom. (b) In the whole-face condition, faces were presented in their entirety for 1 sec. (c) Participants were required to make binary categorization judgments about facial stimuli (e.g., 'male' or 'female'?). Responses were analyzed by fitting psychometric functions. The slope of the function is an estimate of decision noise; steep and shallow slopes are associated with low and high noise estimates, respectively.

2.4 Analysis

Participants' responses were analyzed by modelling psychometric functions (Figure 2c), using the Palamedes toolbox (Prins & Kingdom, 2009). For each participant, separate cumulative Gaussian functions were fitted for each of the four conditions. For the purposes of the current study the key parameter of interest is decision noise; a measure of the precision with which stimuli are categorized. Noise estimates were inferred directly from the slope of the psychometric function, defined as the standard deviation of the symmetric Gaussian distribution underlying each cumulative Gaussian function. Lower noise (steep slopes) estimates indicate that observers can perceive subtle differences in stimulus strength and vary their responses accordingly. Greater noise estimates (shallow slopes) reveal that participants' responses are relatively invariant to changes in stimulus strength. Estimates of decision noise were analyzed using ANOVA with Stimulus Orientation (upright, inverted) and Viewing Condition (aperture, whole-face) as within-subjects factors.

3. Results

In our first set of experiments we examined observers' judgments of stimuli drawn from morph continua blending facial identity, age, and gender incrementally. These attributions are thought to depend on the encoding of face structure; a semi-permanent, durable source of facial variation that changes slowly over time (Bruce & Young, 1986; Haxby, Hoffman, & Gobbini, 2000). In our second set of experiments, we examined observers' judgments of transient facial expressions. The visual analysis of facial expression may partially dissociate from the analysis of semi-permanent features such as identity (Biotti & Cook, 2016; Bruce & Young, 1986; Calder & Young, 2005; Duchaine, Parker, & Nakayama, 2003; Haxby et al., 2000) and is thought to depend on information sampled from different local regions (Gosselin & Schyns, 2001). Observers categorized stimuli drawn from morph continua blending expressions of fear and surprise, disgust and sadness, sincere and insincere happiness, respectively. Contrary to the predictions of the holistic account, we found no evidence of disproportionate aperture effects when observers judged upright faces. In contrast, however, two striking findings replicated consistently in all experiments: i) substantial inversion effects when faces were viewed through the aperture, and ii) substantial aperture effects when viewing inverted faces. Mean decision noise estimates obtained in the four viewing conditions are shown in Figure 3.
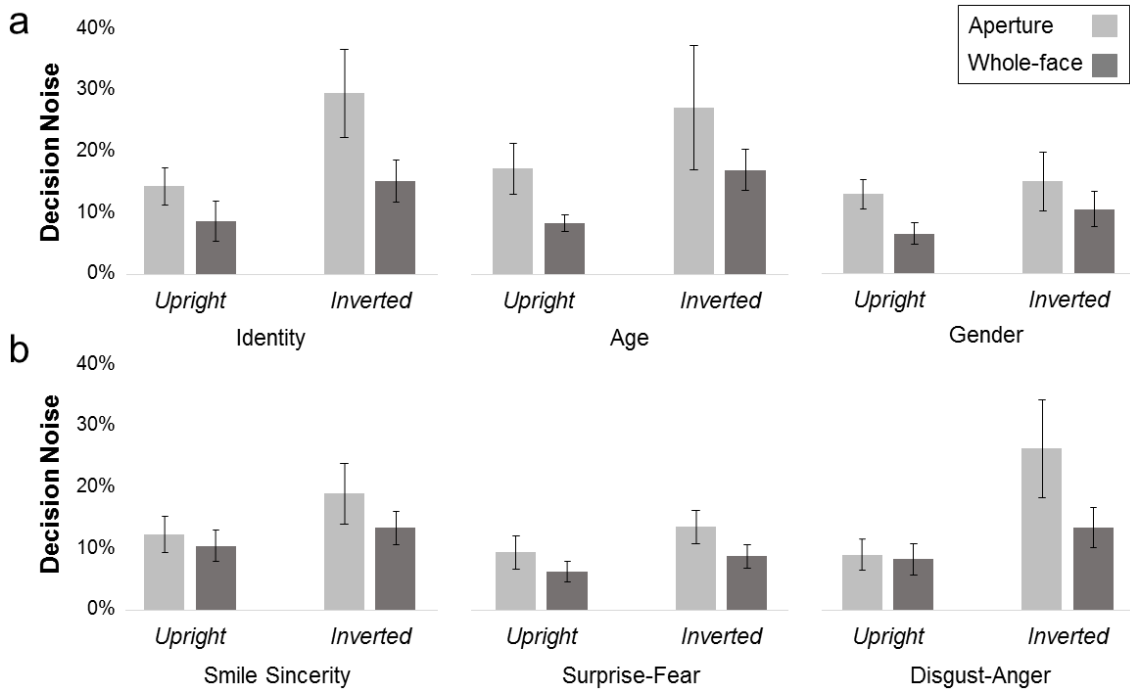
Figure 3. (a) Mean estimates of decision noise for structural attributes. (b) Mean estimates of decision noise for transient expression attributes. Error bars denote 95% confidence intervals.

3.1 Structural facial attributes

Identity classification. Sixteen observers (7 males, $M_{age} = 28.31$, $SD_{age} = 6.45$) made identity judgments, responding 'Bush' or 'Clinton'. A main effect of Viewing Condition was observed [$F(1,15) = 31.26$, $p < .001$, $\eta^2 = .68$] whereby decision noise was greater when faces were viewed through the aperture. Significant effects of Viewing Condition were seen for upright [$t(15) = 4.85$, $p < .001$] and inverted faces [$t(15) = 4.91$, $p < .001$]. A main effect of Orientation was also observed [$F(1,15) = 43.23$, $p < .001$, $\eta^2 = .74$], whereby decision noise was greater when faces were viewed upside-down. Simple contrasts revealed significant effects of Orientation for aperture [$t(15) = 5.84$, $p < .001$] and whole-face presentation [$t(15) = 4.41$, $p = .001$]. A Viewing Condition × Orientation interaction was observed [$F(1,15) = 10.70$, $p = .005$, $\eta^2 = .42$], whereby aperture viewing had a greater detrimental effect when faces were inverted. Decision noise was broadly similar when observers viewed upright faces through apertures and when inverted faces were viewed in their entirety [$t(15) = .55$, $p > .250$].

Age classification. Sixteen observers (4 males, $M_{age} = 27.50$, $SD_{age} = 6.27$) made age judgments, responding 'Man' or 'Boy'. A main effect of Viewing Condition was observed [$F(1, 15) = 13.26$, $p = .002$, $\eta^2 = .47$], whereby decision noise was greater in the aperture

12

viewing condition. Significant effects of Viewing Condition were seen for upright [$t(15)$ = 5.95, p < .001] and inverted faces [$t(15)$ = 2.59, p = .020]. A main effect of Orientation was also observed [$F(1,15)$ = 23.79, p < .001, η^2 = .61], whereby noise estimates were greater when faces were inverted. Simple contrasts revealed significant effects of Orientation for aperture [$t(15)$ = 3.13, p = .007] and whole-face presentation [$t(15)$ = 7.95, p < .001]. No Viewing Condition × Orientation interaction was observed [$F(1,15)$ = .24, p > .250, η^2 = .02]. Decision noise was broadly similar when observers viewed upright faces through apertures and when inverted faces were viewed in their entirety [$t(15)$ = .20, p >.250].

Gender classification. Sixteen observers (4 males, M_{age} = 27.86, SD_{age} = 6.33) made gender judgments, responding 'Male' or 'Female'. A main effect of Viewing Condition was observed [$F(1,15)$ = 36.39, p < .001, η^2 = .71], whereby decision noise was greater when faces were viewed through the aperture. Significant effects of Viewing Condition were seen for upright [$t(15)$ = 5.73, p < .001] and inverted faces [$t(15)$ = 3.57, p = .003]. A main effect of Orientation was also observed [$F(1,15)$ = 5.46, p = .034, η^2 = .27], whereby decision noise was greater when faces were inverted. Simple contrasts revealed significant effects of Orientation for whole-face [$t(15)$ = 3.25, p = .005], but not aperture presentation [$t(15)$ = 1.18, p > .250]. No Viewing Condition × Orientation interaction was observed [$F(1,15)$ = 1.53, p > .235, η^2 = .09]. Decision noise was greater when observers viewed upright faces through apertures than when inverted faces were viewed in their entirety [$t(15)$ = 2.16, p = .048].

3.2 Transient facial attributes
Smile-sincerity classification. Sixteen observers (5 males, M_{age} = 27.0, SD_{age} = 4.77) made smile-sincerity judgments, responding 'Sincere' or 'Insincere'. A main effect of Viewing Condition was observed [$F(1,15)$ = 7.43, p = .016, η^2 = .33], whereby decision noise was greater when faces were viewed through the aperture. Significant effects of Viewing Condition were seen for inverted [$t(15)$ = 2.83, p = .013], but not upright presentation [$t(15)$ = 1.23, p > .20]. A main effect of Orientation was also observed [$F(1,15)$ = 17.12, p = .001, η^2 = .53], whereby decision noise was greater when faces were inverted. Simple contrasts revealed significant effects of Orientation for aperture [$t(15)$ = 3.64, p = .002] and whole-face presentation [$t(15)$ = 2.23, p = .041]. No significant Viewing Condition × Orientation interaction was observed [$F(1,15)$ = 3.05, p = .101, η^2 = .17]. Decision noise

was similar when observers viewed upright faces through apertures and when inverted faces were viewed in their entirety [$t(15) = 1.01$, $p > .250$].

Surprise-fear classification. Sixteen observers (6 males, $M_{age} = 27.88$, $SD_{age} = 6.39$) made surprise-fear judgments, responding 'Surprise' or 'Fear'. A main effect of Viewing Condition was observed [$F(1,15) = 15.08$, $p = .001$, $\eta^2 = .50$], whereby decision noise was greater when faces were viewed through the aperture. The effect of Viewing Condition was only marginally significant for upright faces [$t(15) = 2.12$, $p = .051$], but was clearly significant for inverted faces [$t(15) = 3.80$, $p = .002$]. A main effect of Orientation was also observed [$F(1,15) = 12.23$, $p = .003$, $\eta^2 = .45$], whereby decision noise was greater when faces were inverted. Simple contrasts revealed significant effects of Orientation for aperture [$t(15) = 2.50$, $p = .025$] and whole-face presentation [$t(15) = 2.92$, $p = .011$]. No significant Viewing Condition $\times$ Orientation interaction was observed [$F(1,15) = .76$, $p > .250$, $\eta^2 = .05$]. Decision noise was similar when observers viewed upright faces through apertures and when inverted faces were viewed in their entirety [$t(15) = .38$, $p > .250$].

Disgust-sadness classification. Sixteen observers (9 males, $M_{age} = 28.19$, $SD_{age} = 7.35$) made disgust-sadness judgments, responding 'Disgust' or 'Sadness'. A main effect of Viewing Condition was observed [$F(1,15) = 19.02$, $p = .001$, $\eta^2 = .56$], whereby decision noise was greater when faces were viewed through the aperture. A significant effect of Viewing Condition was seen for inverted [$t(15) = 4.62$, $p < .001$], but not upright presentation [$t(15) = 0.79$, $p > .250$]. A main effect of Orientation was also observed [$F(1,15) = 33.48$, $p < .001$, $\eta^2 = .69$], whereby decision noise was greater when faces were inverted. Simple contrasts revealed significant effects of Orientation for aperture [$t(15) = 5.67$, $p < .001$] and whole-face presentation [$t(15) = 3.59$, $p = .003$]. A significant Viewing Condition $\times$ Orientation interaction was also observed [$F(1,15) = 18.93$, $p = .001$, $\eta^2 = .56$] whereby aperture viewing had a greater detrimental effect when faces were inverted. Decision noise was greater when inverted faces were viewed in their entirety, than when upright faces were viewed through the aperture [$t(15) = 2.92$, $p = .011$].

3.3 High-contrast occlusion

In the foregoing experiments we found no evidence of disproportionate aperture effects for upright faces. One possibility is that compensatory processes obscured the detrimental effects of aperture viewing in the upright condition. In the aperture conditions, the

occluded facial regions appeared black (pixel values were set to zero). When faces were presented in the upright aperture condition, it is possible that participants were able to retain previously exposed face parts in visual working memory, facilitating better performance. We examined this possibility in a further set of experiments. Observers judged structural attributes (identity, age, gender) using an identical procedure to that employed previously. Crucially, however, occluded regions were replaced with a mask constructed from high-contrast greyscale ellipses (Figure 4a). If visual working memory augmented perceptual decisions in previous upright aperture conditions, high-contrast occlusion ought to increase the detrimental effects of aperture viewing. Contrary to this prediction, however, the addition of the high-contrast occluder had no material effect on the pattern of results (Figure 4b).
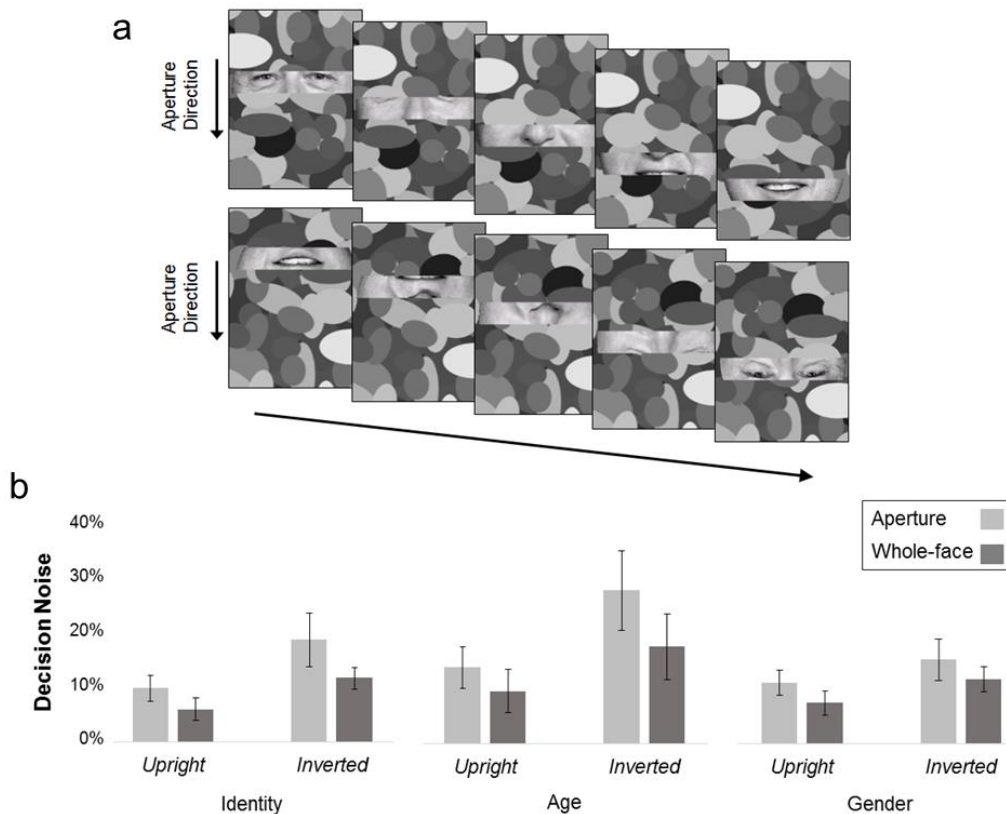
Figure 4. (a) Occluded regions were replaced with a mask constructed from high-contrast greyscale ellipses to prevent observers holding previously revealed regions in visual working memory. (b) Mean estimates of decision noise for structural attributes obtained using the high-contrast occluder. Error bars denote 95% confidence intervals.

Identity classification. Sixteen observers (5 males, $M_{age} = 29.56$, $SD_{age} = 7.21$) made identity judgments, responding 'Bush' or 'Clinton'. A main effect of Viewing Condition was observed [$F(1,15) = 26.52$, $p < .001$, $\eta^2 = .64$] whereby noise estimates were greater when faces were viewed through the aperture. Significant effects of Viewing Condition

were seen for upright [$t(15) = 3.79$, $p = .002$] and inverted faces [$t(15) = 3.87$, $p = .001$]. A main effect of Orientation was also observed [$F(1,15) = 32.30$, $p < .001$, $\eta^2 = .68$], whereby decision noise was greater when faces were inverted. Simple contrasts revealed significant effects of Orientation for aperture [$t(15) = 4.29$, $p = .001$] and whole-face presentation [$t(15) = 5.38$, $p < .001$]. No significant Viewing Condition × Orientation interaction was observed [$F(1,15) = 2.18$, $p = .160$, $\eta^2 = .13$]. Decision noise was greater when inverted faces were viewed in their entirety than when upright faces were viewed through the aperture [$t(15) = 2.31$, $p = .036$].

Age classification. Sixteen observers (4 males, $M_{age} = 24.69$, $SD = 5.53$) made age judgments, responding 'Man' or 'Boy'. A main effect of Viewing Condition was observed [$F(1,15) = 23.93$, $p < .001$, $\eta^2 = .62$] whereby noise estimates were greater when viewed through the aperture. Significant effects of Viewing Condition were seen for upright [$t(15) = 2.61$, $p = .020$] and inverted faces [$t(15) = 5.07$, $p < .001$]. A main effect of Orientation was also observed [$F(1,15) = 19.63$, $p < .001$, $\eta^2 = .57$], whereby decision noise was greater when faces were inverted. Simple contrasts revealed significant effects of orientation inversion following aperture [$t(15) = 5.31$, $p < .001$] and whole-face presentation [$t(15) = 2.89$, $p = .011$]. A Viewing Condition × Orientation interaction was observed [$F(1,15) = 6.61$, $p = .021$, $\eta^2 = .31$], whereby aperture viewing had a greater detrimental effect when faces were inverted. Decision noise was broadly similar when observers viewed upright faces through apertures and when inverted faces were viewed in their entirety [$t(15) = 2.01$, $p = .063$].
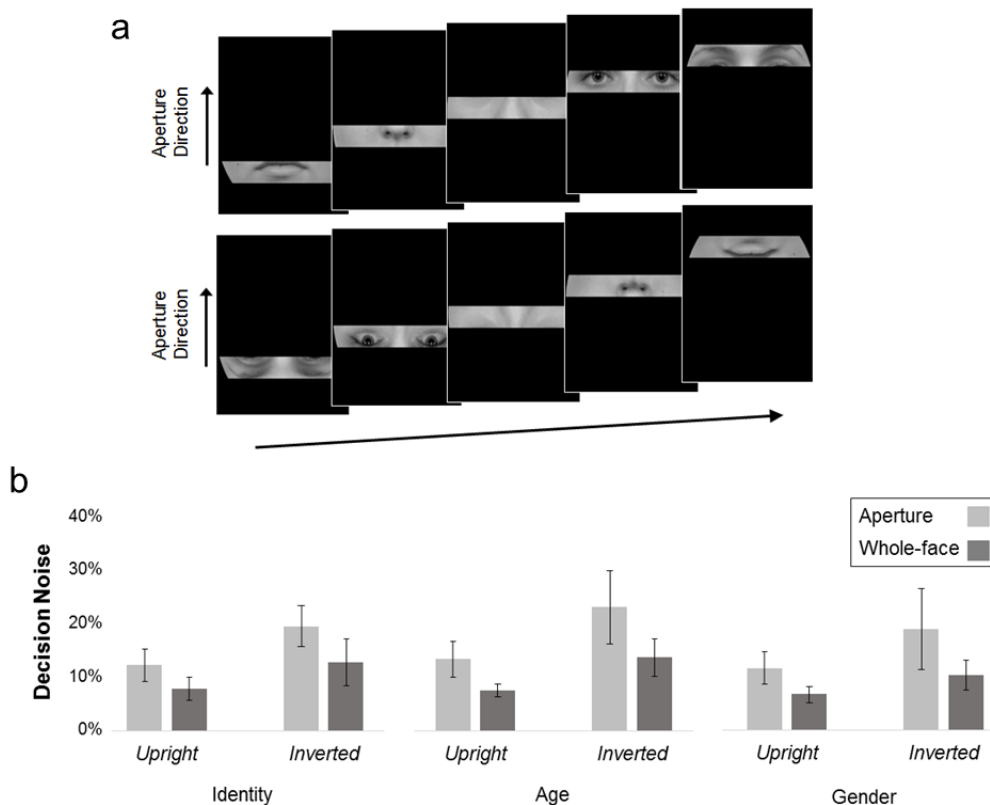
Gender classification. Sixteen observers (5 males, $M_{age} = 29.25$, $SD = 7.85$) made gender judgments, responding 'Male' or 'Female'. A main effect of Viewing Condition was observed [$F(1,15) = 17.66$, $p = .001$, $\eta^2 = .54$] whereby noise estimates were greater when viewed through the aperture. Significant effects of Viewing Condition were seen for upright [$t(15) = 2.99$, $p = .009$] and inverted faces [$t(15) = 2.52$, $p = .023$]. A main effect of Orientation was also observed, [$F(1,15) = 33.03$, $p < .001$, $\eta^2 = .69$], whereby decision noise was greater when faces were inverted. Simple contrasts revealed significant effects of Orientation following aperture [$t(15) = 2.70$, $p = .017$] and whole-face presentation [$t(15) = 5.16$, $p < .001$]. No significant Viewing Condition × Orientation interaction was observed [$F(1,15) = .003$, $p > .250$, $\eta^2 = .001$]. Decision noise was broadly similar when

observers viewed upright faces through apertures and when inverted faces were viewed in their entirety [$t(15) = 0.54$, $p > .250$].

3.4 Reversing the aperture direction

In the aperture conditions described above, the viewing window always moved downwards – from the top of the display to the bottom – irrespective of the orientation of the face. In the upright aperture condition, observers therefore sampled the eye-region before the nose, and then the mouth. Conversely, in the inverted aperture condition, observers first sampled the mouth, then the nose, and finally the eyes. Certain features of our results, notably the high levels of decision noise seen in the inverted aperture conditions, may therefore reflect the order in which perceptual evidence was accumulated. To determine whether the results obtained above are sensitive to the direction of aperture motion, we compared observers' judgments of structural attributes in the whole-face condition with a novel aperture condition, in which the viewing window moved upwards – from the bottom of the display to the top (Figure 5a). Once again, however, this modification failed to change the pattern of results observed (Figure 5b).

Figure 5. (a) To determine whether the results obtained above are sensitive to the direction of aperture motion, we compared observers' judgments of structural attributes in the whole-face condition with a novel aperture condition, in which the viewing window moved upwards. (b) Mean estimates of decision noise for structural attributes obtained using the alternate aperture direction. Error bars denote 95% confidence intervals.

Identity classification. Sixteen observers (6 males, M_{age} = 28.44, SD_{age} = 7.70) completed the identity judgment task, responding 'Bush' or 'Clinton'. A main effect of Viewing Condition was observed [$F(1,15)$ = 38.06, $p < .001$, η^2 = .72] whereby decision noise was greater when faces were viewed through the aperture. Significant effects of Viewing Condition were seen both for upright [$t(15)$ = 4.06, $p = .001$] and inverted faces [$t(15)$ = 3.49, $p = .003$]. A main effect of Orientation was also observed [$F(1,15)$ = 50.89, $p < .001$, η^2 = .77], whereby decision noise was greater when faces were viewed upside-down. Simple contrasts revealed significant effects of orientation inversion for both aperture [$t(15)$ = 4.37, $p = .001$] and whole-face presentation [$t(15)$ = 3.49, $p = .003$]. No significant Viewing Condition × Orientation interaction was observed [$F(1,15)$ = .79, $p > .250$, η^2 = .05]. Decision noise was broadly similar when observers viewed upright faces through apertures and when inverted faces were viewed in their entirety [$t(15)$ = .49, $p > .250$].

Age classification. Sixteen observers (4 males, M_{age} = 27.75, SD_{age} = 7.38) completed the age judgment task, responding 'Man' or 'Boy'. A main effect of Viewing Condition was observed [$F(1, 15)$ = 19.00, $p = .001$, η^2 = .56], whereby decision noise was greater in the aperture viewing condition. Significant effects of Viewing Condition were seen for both upright [$t(15)$ = 4.49, $p < .001$] and inverted faces [$t(15)$ = 3.14, $p = .007$]. A main effect of Orientation was also observed [$F(1,15)$ = 33.84, $p < .001$, η^2 = .69], whereby noise estimates were greater when faces were inverted. Simple contrasts revealed significant effects of orientation for both aperture [$t(15)$ = 3.98, $p <=.001$] and whole-face presentation [$t(15)$ = 4.06, $p = .001$]. No significant Viewing Condition × Orientation interaction was observed [$F(1,15)$ = 1.41, $p > .250$, η^2 = .09]. Decision noise was broadly similar when observers viewed upright faces through apertures and when inverted faces were viewed in their entirety [$t(15)$ = .30, $p > .250$].

Gender classification. Sixteen observers (4 males, M_{age} = 26.75, SD_{age} = 8.92) completed the gender judgment task, responding 'Male' or 'Female'. A main effect of Viewing Condition was observed [$F(1,15)$ = 15.19, $p = .001$, η^2 = .50], whereby decision noise was greater when faces were viewed through the aperture. Significant effects of Viewing Condition were seen for both upright [$t(15)$ = 4.23, $p = .001$] and inverted faces [$t(15)$ = 3.35, $p = .004$]. A main effect of Orientation was also observed [$F(1, 15)$ = 9.75, $p = .007$,

$\eta^2 = .39$], whereby decision noise was greater when faces were shown inverted. Simple contrasts revealed significant effects of inversion following whole-face presentation [$t(15) = 3.20, p = .006$] and aperture presentation [$t(15) = 2.80, p = .013$]. No Viewing Condition $\times$ Orientation interaction was observed [$F(1,15) = 3.55, p = .079, \eta^2 = .19$]. Decision noise was broadly similar when observers viewed upright faces through apertures than when inverted faces were viewed in their entirety [$t(15) = .10, p > .250$].

4. Discussion

The present study used a novel dynamic aperture paradigm to compare the visual processing of upright and inverted faces. We examined observers' ability to categorize upright and inverted face stimuli viewed in their entirety, or viewed through an aperture that revealed the image incrementally.

4.1 Aperture judgments are impaired by face inversion

As expected, we found that judgments of faces presented in their entirety were impaired by orientation inversion; decision noise was greater when judging upside-down faces, than upright faces. Holistic processing accounts argue that this decrement is caused by a switch from whole-face processing, to piecemeal, parts-based processing (Farah et al., 1998; Maurer et al., 2002; McKone & Yovel, 2009; Rossion, 2008). When viewing faces through apertures, however, observers were also markedly better when the occluded face was upright, than when it was inverted. Strikingly, the magnitude of the inversion effect seen in the aperture conditions was comparable to, and in some case larger than, those seen in the whole face condition. Insofar as aperture viewing forces observers to process visual information in a serial, parts-based manner, this finding is hard to reconcile with holistic processing accounts; one would expect smaller inversion effects where observers' ability to process upright faces holistically is restricted.

One possibility is that the mechanisms for encoding local regions are orientation sensitive, i.e., they are tuned to upright faces. If orientation inversion impairs the description of local regions, detrimental effects would be expected in both the whole-face and aperture viewing conditions. It has been argued that judgments about isolated features (i.e., shown in the absence of a facial context) are insensitive to orientation inversion (McKone & Yovel, 2009). However, some authors have reported that substantial inversion effects can

be obtained with cropped features (Rakover & Teucher, 1997). Interestingly, when judging inter-ocular distance, observers not only exhibit substantial inversion effects when the eye-region is presented in isolation, but the magnitude of the effect does not increase when the occluded areas of the face are made visible (Leder et al., 2001). Consistent with the present findings, these results suggest that orientation inversion may impair the perception of local regions, as well as whole-face processing. When faces are shown upside-down, noisy local descriptions may not only be less informative, but may also be difficult to combine in an optimal manner, contributing to greater decision noise (Gold et al., 2012).

A second possibility is that our aperture window permitted some holistic processing in the upright condition. In its strongest form, holistic processing theory argues that parallel whole-face processing is gated, engaged only in the presence of an intact whole face (e.g., Farah et al., 1998; McKone & Yovel, 2009; Tanaka & Farah, 1993; Tsao & Livingstone, 2008). Our aperture condition would be expected to block this type of processing completely. However, some authors have suggested that local regions viewed in isolation may still recruit some form of orientation-specific holistic (or 'configural') processing that improves observers' sensitivity to inter-feature spatial relations (e.g., Leder & Bruce, 2000; Leder et al., 2001; Rossion, 2008). It is possible that the aperture window used here enabled observers to see enough of the face – for example, both eyes could be viewed simultaneously – to utilize some residual configural processing when faces are presented upright. Viewing windows of smaller sizes or different shapes may block this residual configural ability (e.g., Van Belle et al., 2010). Understanding why different types of aperture window differentially affect the orientation sensitivity of observers' judgments may prove significant for theories of holistic face processing.

Recently it has been suggested that the mechanisms for feature encoding may exhibit retinotopic tuning; the perception of eyes and mouths may be better when presented in the upper and lower visual field, respectively (de Haas et al., 2016). According to this account, perceptual decrements associated with orientation inversion may reflect the presentation of mouths and eyes in the upper and lower visual field, respectively. While this interesting possibility warrants further investigation, it seems unlikely to account for the inversion effect seen in our aperture conditions. Because observers were required to fixate a portion

of the image subtending less than 1° as it shifted across the image, exposed regions were likely foveated during each aperture trial.

4.2 Comparable whole-face advantage for upright and inverted faces

Consistent with the view that upright faces benefit from a whole-face advantage (Farah et al., 1998; Maurer et al., 2002; McKone & Yovel, 2009; Rossion, 2008), observers exhibited less decision noise when they viewed upright faces in their entirety, than when viewed through the aperture. While the aperture manipulation blocks whole-face processing, local feature information is preserved. These results therefore provide clear, direct evidence for a whole-face advantage; the ability to process facial information in parallel improves perceptual decisions, relative to serial analysis. This was the case when observers judged a variety of structural (identity, gender, age) and transient (facial affect, smile sincerity) attributes, suggesting that whole-face presentation aids the encoding of face structure at a relatively early stage of face processing (Bruce & Young, 1986; Calder & Young, 2005; Duchaine & Yovel, 2015; Haxby et al., 2000).

Strikingly, aperture viewing also had large detrimental effects on the perception of inverted faces equal to, or greater than, those seen for upright faces. The fact that occluding contextual information in the inverted aperture condition resulted in further perceptual decrements contradicts the view that the perception of inverted faces depends on a serial piecemeal analysis of local features. Instead, this finding indicates that both upright and inverted face perception exhibit a whole-face advantage. This suggestion accords with current thinking about the role of context in human vision. It has been argued that stimulus ambiguity and contextual influence are closely linked; where the physical attributes of a target stimulus do not clearly distinguish between different perceptual hypotheses, contextual information may be used to resolve the ambiguity (Bar, 2004; Friston, 2005; Gregory, 1997; Lawson, Rees, & Friston, 2014). Where orientation inversion disrupts the encoding of local features, residual ability to use information from surrounding regions may be crucial. For example, the wider context contains valuable cues to lighting conditions and face shape – perceptual evidence that can help confirm or reject perceptual hypotheses.

The observation of a whole-face advantage for inverted face processing appears inconsistent with the orientation sensitivity of the composite face illusion (Murphy et al., 2016; Rossion, 2013; Young et al., 1987). In the composite illusion, the top half of one face appears to fuse perceptually with the bottom half of another, when the two halves are aligned and presented upright (Hole, 1994; Young et al., 1987). However, this compelling perceptual bias is greatly reduced when arrangements are shown upside-down, suggestive of weaker integration (McKone et al., 2013; Susilo et al., 2013). How might this feature of the composite face effect be reconciled with our finding of a whole-face advantage for inverted face processing? Crucially, context may still be informative even where it fails to induce demonstrable illusory distortion. One possibility is that the scale of the integration window may be reduced by inversion; residual integration processes may be sufficient to combine information from proximal features of inverted faces, but not to bind distal regions (see also Perceptual Field Hypothesis; Rossion, 2009). Alternatively, perceptual predictions may be weaker, or integration processes slower and less automatic (see Richler, Wong et al., 2011).

4.3 Do upright and inverted faces recruit different visual processing?

Sizeable face inversion effects (Yin, 1969) are often cited as evidence that upright and inverted faces recruit different types of visual processing (Farah et al., 1998; Maurer et al., 2002; McKone et al., 2007; McKone & Yovel, 2009; Richler, Wong et al., 2011; Rossion, 2008). Consistent with this view, upright faces are more effective at capturing attention (Shah, Gaule, Bird, & Cook, 2013) and are harder to mask from conscious awareness (Stein, Reeder, & Peelen, 2016; Stein, Sterzer, & Peelen, 2012), than inverted faces. At the neural level, the fMRI signal change seen in the fusiform face area is larger for upright faces than for inverted faces (Yovel & Kanwisher, 2005), and orientation inversion amplifies and delays the N170, an event-related potential elicited by faces (Bentin, Allison, Puce, Perez, & McCarthy, 1996; Eimer, 2000; Rossion et al., 1999). Relative to typical observers, individuals with acquired or developmental prosopagnosia – conditions characterized by face recognition difficulties – often exhibit striking deficits when viewing upright faces, but broadly typical discrimination of inverted faces (Farah, Wilson, Drain, & Tanaka, 1995; Shah, Gaule, Gaigg, Bird, & Cook, 2015).

Our finding of comparable aperture effects in both orientations accords with previous evidence that the visual processing of upright and inverted faces differs quantitatively, not qualitatively. Should upright and inverted faces engage qualitatively different modes of processing, one might expect a sudden 'jump' in perceptual ability as faces approach their canonical orientation. However, recognition ability varies linearly as a function of orientation rotation (Valentine & Bruce, 1988). Observers also base their judgments of upright and inverted faces on information sampled from similar regions, notably from around the eyes (Sekuler et al., 2004). Composite face effects can be observed for inverted faces when larger sample sizes are utilized (Susilo et al., 2013), or longer presentation times allowed (Richler, Mack, Palmeri, & Gauthier, 2011). At the neural level, upright and inverted faces both elicit strong signal changes in occipital and fusiform face areas (Yovel & Kanwisher, 2005). Similarly, the application of disruptive transcranial magnetic stimulation to the occipital face area impairs judgments of both upright and inverted faces (Pitcher, Duchaine, Walsh, Yovel, & Kanwisher, 2011).

4.4. Conclusion

In summary, when viewing faces through a dynamic aperture, observers continue to show striking inversion effects, suggesting that perceptual decrements induced by inversion may reflect problems encoding local regions, not the loss of whole-face processing. Furthermore, we find evidence of a comparable whole-face advantage for inverted faces, contrary to the view that holistic processing is recruited by upright faces only. It remains to be seen how access to the wider face context modulates the processing of upright and inverted faces. Nevertheless, it is clear that the contribution of whole-face processing to inverted face perception has been underestimated.

Footnotes

[1]Some authors have advanced an alternative view, arguing that isolated local regions may recruit some form of orientation-specific holistic or 'configural' processing in the absence of the wider face context (see Leder & Bruce, 2000; Leder et al., 2001; Rossion, 2008, 2009).

References

Anstis, S. M., & Atkinson, J. (1967). Distortions in moving figures viewed through a stationary slit. *The American Journal of Psychology, 80*(4), 572-585.

Avidan, G., Tanzer, M., & Behrmann, M. (2011). Impaired holistic processing in congenital prosopagnosia. *Neuropsychologia, 49*(9), 2541-2552.

Bar, M. (2004). Visual objects in context. *Nature Reviews Neuroscience, 5*(8), 617-629.

Behrmann, M., Thomas, C., & Humphreys, K. (2006). Seeing it differently: visual processing in autism. *Trends in Cognitive Sciences, 10*(6), 258-264.

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.

Biotti, F., & Cook, R. (2016). Impaired perception of facial emotion in developmental prosopagnosia. *Cortex, 81*, 126-136.

Biotti, F., Wu, E., Yang, H., Jiahui, G., Duchaine, B., & Cook, R. (in press). Normal composite face effects in developmental prosopagnosia. *Cortex*.

Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision, 10*(4), 433-436.

Bruce, V., & Young, A. W. (1986). Understanding face recognition. *British Journal of Psychology, 77*, 305-327.

Calder, A. J., & Young, A. W. (2005). Understanding the recognition of facial identity and facial expression. *Nature Reviews Neuroscience, 6*(8), 641-651.

Cook, R., Aichelburg, C., & Johnston, A. (2015). Illusory feature slowing: evidence for perceptual models of global facial change. *Psychological Science, 26*(4), 512-517.

Cook, R., Brewer, R., Shah, P., & Bird, G. (2014). Intact facial adaptation in autistic adults. *Autism Research, 7*(4), 481-490.

Craddock, M., Martinovic, J., & Lawson, R. (2012). An advantage for active versus passive aperture-viewing in visual object recognition. *Perception, 40*(10), 1154-1163.

de Haas, B., Schwarzkopf, D. S., Alvarez, I., Lawson, R. P., Henriksson, L., Kriegeskorte, N., et al. (2016). Perception and processing of faces in the human brain is tuned to typical feature locations. *Journal of Neuroscience, 36*(36), 9289-9302.

DeGutis, J., Cohan, S., & Nakayama, K. (2014). Holistic face training enhances face processing in developmental prosopagnosia. *Brain, 137*(Pt 6), 1781-1798.

DeGutis, J., Wilmer, J., Mercado, R. J., & Cohan, S. (2013). Using regression to measure holistic face processing reveals a strong link with face recognition ability. *Cognition, 126*(1), 87-100.

Duchaine, B., Parker, H., & Nakayama, K. (2003). Normal recognition of emotion in a prosopagnosic. *Perception, 32*(7), 827-838.

Duchaine, B., & Yovel, G. (2015). A revised neural framework for face processing. *Annual Review of Vision Science, 1*, 393-416.

Eimer, M. (2000). Effects of face inversion on the structural encoding and recognition of faces - evidence from event-related brain potentials. *Cognitive Brain Research, 10*, 145-158.

Evers, K., Van Belle, G., Steyaert, J., Noens, I., & Wagemans, J. (2017). Gaze-contingent display changes as new window on analytical and holistic face perception in children with autism spectrum disorder. *Child Development*.

Farah, M. J., Wilson, K. D., Drain, H. M., & Tanaka, J. R. (1995). The inverted face inversion effect in prosopagnosia: Evidence for mandatory, face-specific perceptual mechanisms. *Vision Research, 35*(14), 2089-2093.

Farah, M. J., Wilson, K. D., Drain, M., & Tanaka, J. N. (1998). What is" special" about face perception? *Psychological Review, 105*(3), 482-498.

Fifić, M., & Townsend, J. T. (2010). Information-processing alternatives to holistic perception: Identifying the mechanisms of secondary-level holism within a categorization paradigm. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 36*(5), 1290-1313.

Finzi, R. D., Susilo, T., Barton, J. J. S., & Duchaine, B. (2016). The role of holistic face processing in acquired prosopagnosia: Evidence from the composite face effect. *Visual Cognition*.

Fitousi, D. (2015). Composite faces are not processed holistically: evidence from the Garner and redundant target paradigms. *Attention, Perception, & Psychophysics, 77*(6), 2037-2060.

Fitousi, D. (2016). Comparing the role of selective and divided attention in the composite face effect: Insights from Attention Operating Characteristic (AOC) plots and cross-contingency correlations. *Cognition, 148*, 34-46.

Friston, K. (2005). A theory of cortical responses. *Philosophical Transactions of the Royal Society B- Biological Sciences, 360*(1456), 815-836.

Gold, J. M., Mundy, P. J., & Tjan, B. S. (2012). The perception of a face is no more than the sum of its parts. *Psychological Science, 23*(4), 427-434.

Gosselin, F., & Schyns, P. G. (2001). Bubbles: a technique to reveal the use of information in recognition tasks. *Vision Research, 41*(17), 2261-2271.

Gray, K. L. H., Murphy, J., Marsh, J. E., & Cook, R. (2017). Modulation of the composite face effect by unintended emotion cues. *Royal Society Open Science, 4*(4), 160867.

Gregory, R. L. (1997). Knowledge in perception and illusion. *Philosophical Transactions of the Royal Society B- Biological Sciences, 352*(1358), 1121-1127.

Haig, N. D. (1985). How faces differ - A new comparative technique. *Perception, 14*(5), 601-615.

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.

Hayward, W. G., Crookes, K., Chu, M. H., Favelle, S. K., & Rhodes, G. (2016). Holistic processing of face configurations and components. *Journal of Experimental Psychology: Human Perception and Performance*.

Hole, G. (1994). Configurational factors in the perception of unfamiliar faces. *Perception, 23*(1), 65-74.

Ipser, A., & Cook, R. (2016). Inducing a concurrent motor load reduces categorization precision for facial expressions. *Journal of Experimental Psychology: Human Perception and Performance, 42*(5), 706-718.

Konar, Y., Bennett, P. J., & Sekuler, A. B. (2010). Holistic processing is not correlated with face-identification accuracy. *Psychological Science, 21*(1), 38-43.

Langner, O., Dotsch, R., Bijlstra, G., Wigboldus, D. H. J., Hawk, S. T., & van Knippenberg, A. (2010). Presentation and validation of the Radboud Faces Database. *Cognition & Emotion, 24*(8), 1377-1388.

Lawson, R. P., Rees, G., & Friston, K. J. (2014). An aberrant precision account of autism. *Frontiers in Human Neuroscience, 8*, 302.

Le Grand, R., Cooper, P. A., Mondloch, C. J., Lewis, T. L., Sagiv, N., de Gelder, B., et al. (2006). What aspects of face processing are impaired in developmental prosopagnosia? *Brain and Cognition, 61*(2), 139-158.

Leder, H., & Bruce, V. (2000). When inverted faces are recognized: The role of configural information in face recognition. *The Quarterly Journal of Experimental Psychology: Section A, 53*(2), 513-536.

Leder, H., Candrian, G., Huber, O., & Bruce, V. (2001). Configural features in the context of upright and inverted faces. *Perception, 30*(1), 73-83.

Martini, P., McKone, E., & Nakayama, K. (2006). Orientation tuning of human face processing estimated by contrast matching in transparency displays. *Vision Research, 46*(13), 2102-2109.

Maurer, D., Le Grand, R., & Mondloch, C. J. (2002). The many faces of configural processing. *Trends in Cognitive Sciences, 6*(6), 255-260.

McConkie, G. W., & Rayner, K. (1975). The span of the effective stimulus during a fixation in reading. *Attention, Perception, & Psychophysics, 17*(6), 578-586.

McKone, E., Davies, A. A., Darke, H., Crookes, K., Wickramariyaratne, T., Zappia, S., et al. (2013). Importance of the inverted control in measuring holistic face processing with the composite effect and part-whole effect. *Frontiers in Psychology, 4*(33), 1-21.

McKone, E., Kanwisher, N., & Duchaine, B. (2007). Can generic expertise explain special processing for faces? *Trends in Cognitive Sciences, 11*(1), 8-15.

McKone, E., & Yovel, G. (2009). Why does picture-plane inversion sometimes dissociate perception of features and spacing in faces, and sometimes not? Toward a new theory of holistic processing. *Psychonomic Bulletin & Review, 16*(5), 778-797.

Morgan, M. J., Findlay, J. M., & Watt, R. J. (1982). Aperture viewing: A review and a synthesis. *The Quarterly Journal of Experimental Psychology, 34*(2), 211-233.

Murphy, J., Gray, K. L. H., & Cook, R. (2016). The composite face illusion. *Psychonomic Bulletin & Review*.

Palermo, R., Willis, M. L., Rivolta, D., McKone, E., Wilson, C. E., & Calder, A. J. (2011). Impaired holistic coding of facial expression and facial identity in congenital prosopagnosia. *Neuropsychologia, 49*(5), 1226-1235.

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: transforming numbers into movies. *Spatial Vision, 10*(4), 437-442.

Piepers, D. W., & Robbins, R. A. (2013). A review and clarification of the terms "holistic," "configural," and "relational" in the face perception literature. *Frontiers in Psychology, 3*(559), 1-11.

Pitcher, D., Duchaine, B., Walsh, V., Yovel, G., & Kanwisher, N. (2011). The role of lateral occipital face and object areas in the face inversion effect. *Neuropsychologia, 49*(12), 3448-3453.

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

Rakover, S. S., & Teucher, B. (1997). Facial inversion effects: parts and whole relationship. *Perception and Psychophysics, 59*(5), 752-761.

Rezlescu, C., Susilo, T., Wilmer, J. B., & Caramazza, A. (2017). The inversion, part-whole, and composite effects reflect distinct perceptual mechanisms with varied relationship to face recognition. *Journal of Experimental Psychology: Human Perception and Performance*.

Rhodes, G., Brake, S., & Atkinson, A. P. (1993). What's lost in inverted faces? *Cognition, 47*(1), 25-57.

Rhodes, G., Hayward, W. G., & Winkler, C. (2006). Expert face coding: Configural and component coding of own-race and other-race faces. *Psychonomic Bulletin & Review, 13*(3), 499-505.

Richler, J. J., Cheung, O. S., & Gauthier, I. (2011a). Beliefs alter holistic face processing ... if response bias is not taken into account. *Journal of Vision, 11*(13), 1-13.

Richler, J. J., Cheung, O. S., & Gauthier, I. (2011b). Holistic processing predicts face recognition. *Psychological Science, 22*(4), 464-471.

Richler, J. J., Mack, M. L., Palmeri, T. J., & Gauthier, I. (2011). Inverted faces are (eventually) processed holistically. *Vision Research, 51*(3), 333-342.

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.

Rieger, J. W., Grüschow, M., Heinze, H. J., & Fendrich, R. (2007). The appearance of figures seen through a narrow aperture under free viewing conditions: Effects of spontaneous eye motions. *Journal of Vision, 7*(6), 1-10.

Rock, I. (1981). Anorthoscopic perception. *Scientific American, 244*, 145-153.

Rossion, B. (2008). Picture-plane inversion leads to qualitative changes of face perception. *Acta Psychologica, 128*(2), 274-289.

Rossion, B. (2009). Distinguishing the cause and consequence of face inversion: The perceptual field hypothesis. *Acta Psychologica, 132*(3), 300-312.

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., Delvenne, J. F., Debatisse, D., Goffaux, V., Bruyer, R., Crommelinck, M., et al. (1999). Spatio-temporal localization of the face inversion effect: An event-related potentials study. *Biological Psychology, 50*, 173-189.

Rossion, B., & Retter, T. L. (2015). Holistic face perception: Mind the gap! *Visual Cognition, 23*(3), 379-398.

Searcy, J. H., & Bartlett, J. C. (1996). Inversion and processing of component and spatial-relational information in faces. *Journal of Experimental Psychology: Human Perception and Performance, 22*(4), 904-915.

Sekuler, A. B., Gaspar, C. M., Gold, J. M., & Bennett, P. J. (2004). Inversion leads to quantitative, not qualitative, changes in face processing. *Current Biology, 14*(5), 391-396.

Shah, P., Bird, G., & Cook, R. (2016). Face processing in autism: Reduced integration of cross-feature dynamics. *Cortex*.

Shah, P., Gaule, A., Bird, G., & Cook, R. (2013). Robust orienting to protofacial stimuli in autism. *Current Biology, 23*(24), R1087-R1088.

Shah, P., Gaule, A., Gaigg, S. B., Bird, G., & Cook, R. (2015). Probing short-term face memory in developmental prosopagnosia. *Cortex, 64*, 115-122.

Shah, P., Gaule, A., Sowden, S., Bird, G., & Cook, R. (2015). The 20-item prosopagnosia index (PI20): a self-report instrument for identifying developmental prosopagnosia. *Royal Society Open Science, 2*(6), 140343.

Shen, J., & Palmeri, T. J. (2015). The perception of a face can be greater than the sum of its parts. *Psychonomic Bulletin & Review, 22*(3), 710-716.

Stein, T., Reeder, R. R., & Peelen, M. V. (2016). Privileged access to awareness for faces and objects of expertise. *Journal of Experimental Psychology: Human Perception and Performance, 42*(6), 788-798.

Stein, T., Sterzer, P., & Peelen, M. V. (2012). Privileged detection of conspecifics: Evidence from inversion effects during continuous flash suppression. *Cognition, 125*(1), 64-79.

Susilo, T., McKone, E., Dennett, H., Darke, H., Palermo, R., Hall, A., et al. (2010). Face recognition impairments despite normal holistic processing and face space coding: evidence from a case of developmental prosopagnosia. *Cognitive Neuropsychology, 27*(8), 636-664.

Susilo, T., Rezlescu, C., & Duchaine, B. (2013). The composite effect for inverted faces is reliable at large sample sizes and requires the basic face configuration. *Journal of Vision, 13*(13), 1-9.

Tanaka, J. W., & Farah, M. J. (1993). Parts and wholes in face recognition. *Quarterly Journal of Experimental Psychology, 46*(2), 225-245.

Towler, J., Parketny, J., & Eimer, M. (2015). Perceptual face processing in developmental prosopagnosia is not sensitive to the canonical location of face parts. *Cortex, 74*, 53-66.

Tsao, D. Y., & Livingstone, M. S. (2008). Mechanisms of face perception. *Annual Review of Neuroscience, 31*, 411-437.

Valentine, T., & Bruce, V. (1988). Mental rotation of faces. *Memory & Cognition, 16*(6), 556-566.

Van Belle, G., De Graef, P., Verfaillie, K., Busigny, T., & Rossion, B. (2011). Whole not hole: Expert face recognition requires holistic perception. *Neuropsychologia, 48*(9), 2620-2629.

Van Belle, G., De Graef, P., Verfaillie, K., Rossion, B., & Lefèvre, P. (2010). Face inversion impairs holistic perception: Evidence from gaze-contingent stimulation. *Journal of Vision, 10*(5), 1-10.

Wang, R., Li, J., Fang, H., Tian, M., & Liu, J. (2012). Individual differences in holistic processing predict face recognition ability. *Psychological Science, 23*(2), 169-177.

Wenger, M. J., & Townsend, J. T. (2001). Faces as gestalt stimuli: Process characteristics. In M. J. Wenger & J. T. Townsend (Eds.), *Computational, geometric, and process*

perspectives on facial cognition (pp. 229-284). Mahwah, New Jersey: Lawrence Erlbaum Associates.

Yin, R. K. (1969). Looking at upside-down faces. *Journal of Experimental Psychology, 81*(1), 141-145.

Young, A. W., Hellawell, D., & Hay, D. C. (1987). Configurational information in face perception. *Perception, 16*(6), 747-759.

Yovel, G., & Kanwisher, N. (2005). The neural basis of the behavioral face-inversion effect. *Current Biology, 15*(24), 2256-2262.

Yovel, G., & Kanwisher, N. (2008). The representations of spacing and part-based information are associated for upright faces but dissociated for objects: evidence from individual differences. *Psychonomic Bulletin & Review, 15*(5), 933-939.

Author contributions

JM and RC contributed equally to the design of all experiments and drafted the manuscript for publication. RC constructed the stimuli and wrote the experimental program. JM collected and analyzed the data.