Cognition 169 (2017) 25–35

Contents lists available at ScienceDirect

Cognition

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

Original Articles

Revealing the mechanisms of human face perception using dynamic apertures

Jennifer Murphy [a,*], Richard Cook [b]

[a] MRC Social, Genetic and Developmental Psychiatry Centre, Institute of Psychiatry, Psychology, and Neuroscience, King's College London, London, UK
[b] Department of Psychological Sciences, Birkbeck, University of London, London, UK

ARTICLE INFO

Article history:
Received 19 January 2017
Revised 2 August 2017
Accepted 2 August 2017

Keywords:
Face perception
Holistic processing
Face inversion effect
Aperture viewing
Facial features

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

* Corresponding author at: MRC Social, Genetic and Developmental Psychiatry Centre, Institute of Psychiatry, Psychology and Neuroscience, De Crespigny Park, Denmark Hill, London SE5 8AF, UK.
E-mail address: Jennifer.Murphy@kcl.ac.uk (J. Murphy).

http://dx.doi.org/10.1016/j.cognition.2017.08.001
0010-0277/© 2017 Published by Elsevier B.V.

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, 2017). 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 et al., 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., 2017; 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 & Yovel, 2009; McKone et al., 2007; Piepers & Robbins, 2013; Rossion, 2008).

In recent years, however, the functional significance of the composite face effect has proved controversial (Murphy et al., 2017; 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., 2017; 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., 2017; 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).

[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, Candrian, Huber, & Bruce, 2001; Rossion, 2008, 2009).