Acta Psychologica 188 (2018) 25–38

Contents lists available at ScienceDirect

Acta Psychologica

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

Tolerance to spatial-relational transformations in unfamiliar faces: A further challenge to a configural processing account of identity recognition

Martina Lorenzino, Martina Caminati, Corrado Caudek*

Department of Neuroscience, Psychology, Drug Research, and Child Health, Section of Psychology, Università degli Studi di Firenze, Via di San Salvi 12, Complesso di San Salvi, Padiglione 26, Firenze (FI) 50135, Italy

ARTICLE INFO

Keywords:
Face perception
Facial identity
Facial expressions
Configural processing
Face-inversion effect
Other-race effect

ABSTRACT

One of the most important questions in face perception research is to understand what information is extracted from a face in order to recognize its identity. Recognition of facial identity has been attributed to a special sensitivity to "configural" information. However, recent studies have challenged the configural account by showing that participants are poor in discriminating variations of metric distances among facial features, especially for familiar as opposed to unfamiliar faces, whereas a configural account predicts the opposite. We aimed to extend these previous results by examining classes of unfamiliar faces with which we have different levels of expertise. We hypothesized an inverse relation between sensitivity to configural information and expertise with a given class of faces, but only for neutral expressions. By first matching perceptual discriminability, we measured tolerance to subtle configural transformations with same-race (SR) versus other-race (OR) faces, and with upright versus upside-down faces. Consistently with our predictions, we found a lower sensitivity to at-threshold configural changes for SR compared to OR faces. We also found that, for our stimuli, the face inversion effect disappeared for neutral but not for emotional faces – a result that can also be attributed to a lower sensitivity to configural transformations for faces presented in a more familiar orientation. The present findings question a purely configural account of face processing and suggest that the role of spatial-relational information in face processing varies according to the functional demands of the task and to the characteristics of the stimuli.

1. Introduction

To recognize the face of a familiar individual is an extremely challenging task, because the image of a person's face changes dramatically under different poses, expressions, and illuminations conditions, whereas the images of the faces of different persons are relatively similar to each other, in that they all share the same general configuration (*i.e.*, two eyes above a nose and mouth). Despite this challenge, most adults are extremely good at recognizing familiar faces.

The "configural account," which maintains that face recognition can be explained by an analysis of the image-plane distances among face features, is currently a popular account of how familiar faces are recognized (*e.g.*, Piepers & Robbins, 2012; Tanaka & Gordon, 2011). However, the configural account has come under considerable criticism recently, because it seems better suited to explain perceptual performance in artificial image-matching tasks involving unfamiliar faces, rather than the most important phenomenon of familiar face recognition under highly-variable naturalistic settings (Burton, 2013; Burton, Schweinberger, Jenkins, & Kaufmann, 2015). The present study intends to provide a further test of the configural account by examining a perceptual task (*i.e.*, image matching of at-threshold unfamiliar face variations) that, in principle, affords the most favorable conditions for this theoretical proposal. In the following, we first proceed with a clarification of the terminology and, before presenting our hypothesis and experimental design, we briefly describe the empirical results supporting the configural account and their main limitations.

1.1. Relational, configural, and holistic processing

Face perception differs from the perception of other objects because faces are endowed of emergent properties that are not "deducible" from the properties of their constituent parts. These emergent properties have been described by using terms such as configural, relational, and holistic. However, there is no consensus about terminology and a diversity of meanings is still present in current usage of these terms in the literature (for a discussion, see Piepers & Robbins, 2012).

An influential distinction has been proposed by Diamond and Carey (1986), according to whom identity discrimination can be based on

* Corresponding author.
E-mail addresses: martina.lorenzino@unifi.it (M. Lorenzino), caminati@ordinepsicologiliguria.it (M. Caminati), corrado.caudek@unifi.it (C. Caudek).

https://doi.org/10.1016/j.actpsy.2018.05.004
Received 12 October 2017; Received in revised form 4 May 2018; Accepted 9 May 2018

relatively isolated facial information (*i.e.*, features) or on the simultaneous processing of two or more parts of a face (*i.e.*, relational information). Another influential distinction comes from Rhodes (1988), according to whom *first-order features* describe image properties that "can be characterized independently of other features in the face and as a result can be labelled by single words in natural language," *second-order features* describe the spatial relations between first-order features, and *higher-level features* "are a function of several first- and/or second-order features" (p. 43).

The term *configural processing* has been used to describe "the integration of all, or just some, of this second-order configural information within the face" (Piepers & Robbins, 2012, p. 2). Such term has been used, sometimes interchangeably, sometimes with a different meaning, with the term *holistic*. In the strictest interpretation of holistic processing, faces are perceived as wholes that cannot be broken down into parts, whereas objects are processed on the basis of their constituent parts (Tanaka & Farah, 1993). According to Piepers and Robbins (2012), such extreme view of the holistic model would require the storage in memory of an enormous number of face templates and, thus, does not seem well supported. A compromise that has been proposed, the *holistic/part-based model*, posits that, by acting in parallel, both part-based and holistic processing contribute to face perception (Moscovitch, Winocur, & Behrmann, 1997).

What is important for the present purposes is to be clear about what we mean by configural processing. According to Piepers and Robbins (2012), in the holistic/part-based model, the terms holistic and configural can be taken to mean the same thing if they refer to the integration of spacing information involving *all, as opposed to just some*, facial features. What is meant by "all" in the present context may be clarified by saying that holistic processing cannot be reduced to the measurement of the distances between the major face blobs (McKone & Yovel, 2009). Instead, "[h]olistic processing may derive from calculations of distances between multiple very local landmark points on faces, or it may derive directly from an image unprocessed for key points or parts at any level of scale (e.g., *pixel-by-pixel intensity coding*)" (p. 780, emphasis added). In other words, "holistic processing could derive from a coding in which there is no decomposition of the image at shape-related boundaries at all. For example, the input to holistic processing might be a list of pixel intensities covering the whole face, rather than a set of parts or a set of landmark points." In this sense, "the representation of metric information within the face […] would be purely implicit, not explicit" (p. 792). In our reading, this is this most encompassing definition of "configural information," because it provides the seemingly least committal description of the kind of relational information that is used by the visual system for facial identity recognition. One example of this kind of approach is provided by the computational models of face recognition based on principal component analysis, which take lists of pixel intensities as their input (Hancock, Burton, & Bruce, 1996).

1.2. Sensitivity to configural properties in face perception

Having provided a very broad definition of what is meant by configural information, let us examine some of the face-perception phenomena that have been explained by configural processing. In the Face-Inversion Effect (FIE; Yin, 1969), the decrease in recognition for upside-down faces has been attributed to an impairment of configural processing for inverted faces (Valentine, 1988). According to some authors, the Other-Race Effect (*i.e.*, the better recognition performance for same-race (SR) faces compared to other-race (OR) faces; ORE) is determined by a better configural processing for SR faces compared to OR faces (Michel, Rossion, Han, Chung, & Caldara, 2006; Rossion & Michel, 2011). The part-whole effect, in which the recognition of a face part is easier if it is presented within the context of the whole face rather than in isolation, has been interpreted as showing that configural processing favors the recognition of single facial features (Tanaka & Sengco, 1997). Despite the wealth of evidence supporting the configural account, this approach is not without criticisms and limitations.

1.3. Limitations of the configural processing account

Several studies have questioned the hypothesis that face recognition can be reduced to a sensitivity to configural information. In fact, faces can be still recognized when configural information is changed. For example, Hole, George, Eaves, and Rasek (2002) found that familiar face recognition is not impaired even for dramatic changes in configural information, such as the vertical stretching of twice the original height of the face images. Moreover, Burton et al. (2015) showed that face identity recognition is impaired by nonconfigural changes (*e.g.*, photographic negation).

Recent studies have also re-examined the relation between configural processing and familiarity. People can easily recognize the identity of familiar faces across a wide range of viewing conditions, but these high levels of accuracy do not hold for unfamiliar faces (*e.g.*, Davis & Valentine, 2009). According to the configural account, these results are due to the fact that recognition of familiar face relies more strongly on configural processing than does recognition of unfamiliar faces (Sandford & Burton, 2014). The configural account thus predicts that a perturbation of configural information will have a stronger negative impact on recognition performance for familiar than for unfamiliar faces. However, this hypothesis has been contradicted by several studies. For example, Sandford and Burton (2014) presented faces in a wrong aspect ratio and asked participants to adjust the stimulus images in order to eliminate these distortions. They found that participants were very poor at this task, which suggests a weak sensitivity to configural information. Interestingly, participants' performance in this task was worse for familiar than for unfamiliar faces. According to Sandford and Burton, this result indicates that the representation of a familiar face (but not of unfamiliar faces) requires the tolerance for subtle changes of the face spatial layout in order to maintain a stable representation of face identity and to facilitate recognition (see also Baker, Laurence, & Mondloch, 2017; Burton, Kramer, Ritchie, & Jenkins, 2016). In a similar vein, Caudek (2013) showed that categorical processes in face perception determine a loss in memorial discriminability for the transient image transformations that are irrelevant for identity recognition. He found that configural changes, which do not alter identity recognition, are represented in working memory with lower fidelity for faces than for non-face objects. These results directly contradict the predictions of the configural account.

1.4. Aims of the present study

A first aim of the present study was, as in the study of Sandford and Burton (2014), to investigate sensitivity to configural transformations. However, we did not compare familiar to unfamiliar faces, but rather different classes of unfamiliar faces. Upright faces are more "face-like" than upside-down faces, and the same can be said for SR as compared to OR faces. As suggested by Caudek (2013), we hypothesized a stronger tolerance to configural changes for those classes of unfamiliar faces with which we have a greater expertise. Therefore, participants were expected to notice fewer at-threshold configural transformations for SR than for OR faces, and for upright than for upside-down faces.

A second aim of the present study was to compare neutral and emotional faces. Cognitive processing may "filter out" the perceptual and memorial representations of subtle within-identity configural changes in order to facilitate recognition of familiar neutral faces (Burton et al., 2015). However, configural transformations also transmit affectively valent information. We thus hypothesized that sensitivity to configural changes would be modulated by emotional state: Subtle configural changes should be "filtered-out" in the case of neutral but not expressive