Neuropsychologia 50 (2012) 2454–2459

Contents lists available at SciVerse ScienceDirect

Neuropsychologia

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

The response of face-selective cortex with single face parts and part combinations

Lindsay R. Arcurio, Jason M. Gold, Thomas W. James *

Department of Psychological and Brain Sciences, Indiana University, 1101 E Tenth St., Bloomington, IN 47405, United States

ARTICLE INFO

Article history:
Received 27 February 2012
Received in revised form
21 May 2012
Accepted 18 June 2012
Available online 26 June 2012

Keywords:
Object recognition
Face recognition
fMRI
Occipital face area
Feature-based processing

ABSTRACT

A critical issue in object recognition research is how the parts of an object are analyzed by the visual system and combined into a perceptual whole. However, most of the previous research has examined how changes to object parts influence recognition of the whole, rather than recognition of the parts themselves. This is particularly true of the research on face recognition, and especially with questions related to the neural substrates. Here, we investigated patterns of BOLD fMRI brain activation with internal face parts (features) presented singly and in different combinations. A preference for single features over combinations was found in the occipital face area (OFA) as well as a preference for the two-eyes combination stimulus over other combination stimulus types. The fusiform face area (FFA) and lateral occipital cortex (LO) showed no preferences among the single feature and combination stimulus types. The results are consistent with a growing view that the OFA represents processes involved in early, feature-based analysis.

1. Introduction

The processes involved in object recognition, and especially in face recognition, are often dichotomized into part/feature-based and holistic/configural (Farah, Wilson, Drain, & Tanaka, 1998; Maurer, Le Grand, & Mondloch, 2002; McKone & Yovel, 2009; Rossion, 2008). Although there has been a considerable amount of research investigating the behavioral and neural markers of holistic/configural processing and also of feature changes on holistic/configural processing, there has been relatively little research investigating markers of single part-based processing. Studies that restrict viewing to isolated features converge with eye-movement studies and suggest that face recognition relies largely on the eye/eyebrow regions, followed by the mouth regions, followed by the nose regions (Blais, Jack, Scheepers, Fiset, & Caldara, 2008; Caldara, Zhou, & Miellet, 2010; Haig, 1986; James, Huh, & Kim, 2010; Yarbus, 1967). Studies using response classification or reverse correlation techniques converge with the other methods to suggest that face recognition relies mostly on eye/eyebrow regions, followed by mouth regions (Schyns, Bonnar, & Gosselin, 2002; Sekuler, Gaspar, Gold, & Bennett, 2004). Finally, ideal observer techniques converge with the other methods to show that eye/eyebrow regions carry the most information for face recognition, followed by the mouth regions, followed by the nose regions (Gold, Bennett, & Sekuler, 1999; Gold, Mundy, & Tjan, 2012). Thus, the results of these behavioral studies suggest that, despite the fact that faces may tend to be analyzed using highly configural/holistic strategies, there are parts of the face that are more informative than others and that are analyzed preferentially.

There are only a few studies that have investigated the neural substrates involved in processing these informative parts of a face, but they suggest several important points about the patterns of brain activation found in regions of face- and object-selective cortex. First, there is some evidence that the activation in the FFA, which to some is taken as the hallmark of whole face processing (Kanwisher & Yovel, 2006), is just as sensitive to partial images of faces as it is to whole face images (James et al., 2010; Tong, Nakayama, Moscovitch, Weinrib, & Kanwisher, 2000). Second, fragments of faces that are high in "diagnosticity" produce greater levels of activation in the FFA, occipital face area (OFA), and lateral occipital cortex (LO), than fragments that are low in diagnosticity (Lerner, Epshtein, Ullman, & Malach, 2008; Nestor, Vettel, & Tarr, 2008). Third, despite the fact that the FFA has been shown to be equally sensitive to whole and partial faces, the FFA has been shown to play a greater role in processing whole faces than the OFA and the OFA plays a greater role in the processing face parts than the FFA (Betts & Wilson, 2010; Nichols, Betts, & Wilson, 2010), suggesting that the processing of wholes and parts may not be all-or-none. Most recently, a series of studies using TMS to disrupt processing in the OFA has found evidence that it is highly involved in the processing of face parts (Pitcher, Charles, Devlin, Walsh, & Duchaine, 2009; Pitcher, Duchaine, Walsh, Yovel, & Kanwisher, 2011; Pitcher, Walsh, & Duchaine, 2011; Pitcher, Walsh, Yovel, & Duchaine, 2007). These results have led to the

* Corresponding author. Tel.: +1 812 856 0841.
E-mail address: thwjames@indiana.edu (T.W. James).

0028-3932/$ - see front matter © 2012 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.neuropsychologia.2012.06.016

hypothesis that the OFA may represent a site of early-stage face-part processing that feeds into the FFA, which represents a site of late-stage whole face processing (Pitcher et al., 2011).

The goal of the present study was to extend research on the neural substrates of feature-based processing of faces. fMRI was used to measure BOLD activation with single internal face features and different combinations of those features. We hypothesized that the OFA would respond strongly with face features and would show a gradient of sensitivity across different face features, with highest sensitivity for eye features and lowest sensitivity for nose features. Based on the idea that the FFA represents a later stage of processing that may involve the integration of features, we hypothesized that the FFA would show more activation with combination stimuli than with single features. Alternatively, the FFA may respond weakly and equivalently across the stimulus types, because they all lack a whole-face context.

2. Materials and methods

2.1. Subjects

Fourteen healthy adults (seven males, ages 21–32) participated for payment. Two subjects were excluded from the fMRI analysis due to motion artifacts in the functional imaging data. All subjects signed informed consent forms, and the Indiana University Institutional Review Board approved the study.

2.2. Stimuli

Twelve face images (six males and six female) were created with FaceGen 3.2 (http://www.facegen.com) and are shown in the top of Fig. 1. Parameters in FaceGen were selected such that all faces were between the ages of 20 and 30, symmetric, and equally attractive. Faces were rendered as 256×256 pixel grayscale images. Different sized apertures were used for the eye/eyebrow, nose, and mouth features, but across face images, the size and position of the apertures was kept constant. The final set of single feature stimuli included left eye/eyebrow, right eye/eyebrow, nose, and mouth. Multi-feature combination stimuli were created by combining two, three, or four single features, always taken from the same face, and always positioned in the correct spatial location. The combinations used were the two eyes (2-feature), eyes and mouth (3-feature), and eyes, nose, and mouth (4-feature). It is worthwhile noting that in these seven single feature and combination feature stimulus types, no face outline was used. Examples of the stimuli in Fig. 1 are shown embedded in the same level of noise used during scanning (see procedures below). The top row is shown at the mean contrast level used during scanning. The bottom row is shown at a much higher contrast level for illustration purposes.

2.3. Scanning session procedures

Subjects underwent a pre-scan practice procedure in a MRI simulator located in the Indiana University Imaging Research Facility to familiarize the subjects with the MRI environment, familiarize the subjects with the task, and to help limit any perceptual learning during the subsequent scanning session.

Subjects lay supine in the scanner bore with their head secured in the head coil by foam padding. Subjects viewed stimuli through a mirror that was mounted above the head coil. This allowed the subjects to see the stimuli on a rear-projection screen (40.2×30.3 cm) placed behind the subject in the bore. Stimuli were projected onto the screen with a Mitsubishi LCD projector (model XL30U). The viewing distance from the mirror to the eyelid was 11.5 cm, and the distance from the screen to the mirror was 79 cm, giving a total viewing distance of 90.5 cm. When projected in this manner, the size of the entire 256×256 pixel stimulus image subtended approximately 6° of visual angle.

Each scanning session consisted of one localizer run and seven experimental runs. The localizer run was included to independently, functionally localize object- and face-selective brain regions, specifically the OFA, FFA, and LO for region of interest (ROI) analyses. During the functional localizer run, full contrast, noise-free, grayscale images of familiar objects (e.g., chair, toaster), faces (different from those used in the experimental runs), and phase-scrambled images (derived from the object and face stimuli) were presented in a blocked design while participants fixated the center of the screen. Six stimuli per block were presented for 1.5 s each with an inter-stimulus interval (ISI) of 500 ms, producing a block time of 12 s. Blocks were presented in 48 s cycles of noise–objects–noise–faces. There were eight cycles in the single run and the run began and ended with 12 s of rest, making the total run length 6 min and 48 s.

During experimental runs, each of the seven stimulus types was presented at each of three separate contrast levels in a full-factorial 3×7 design. The contrast levels were determined individually for each subject. This was done to bring behavioral performance below ceiling, to reduce variability across subjects, and to assess the influence of stimulus quality on brain activation. For each subject, the exact contrast levels used were determined from the data collected during the pre-scan practice session. The low contrast for each subject was the level that produced 75% accuracy with the eyes–nose–mouth stimulus during practice trials and the high contrast was the level that produced 75% with the mouth stimulus during practice trials. The middle contrast was a level midway between the low and high levels on a log contrast scale. Contrast is reported as the square root of contrast energy (RMS contrast) and signal-to-noise ratio (SNR) is reported as the ratio of signal contrast energy and noise contrast energy.

The stimuli were presented in a blocked design while participants performed a one-back matching task. Six stimuli per block were all selected from the same stimulus type. Stimuli in a block of trials were selected from one stimulus type and presented at one contrast level (i.e., trials in a block were taken from the same cell in the 3×7 full-factorial design). Stimuli were embedded in Gaussian noise of constant variance (RMS contrast = 0.1) that was re-sampled each trial. Stimuli were presented for 1 s each with an ISI of 2 s, producing a total block length of 18 s. Stimulus blocks were separated by fixation blocks 12 s in length. Matlab R2008a (http://www.mathworks.com) combined with the Psychophysics Toolbox (http://www.psychtoolbox.org; Brainard, 1997; Pelli, 1997) was used to create the stimuli, adjust the signal-to-noise levels, present the stimuli during the scanning session, and collect the behavioral responses. Each experimental run contained 15 stimulus blocks and 16 fixation blocks, for a total run length of 7 min and 42 s. Across the seven runs, there were a total of 105 stimulus blocks, equally divided among the seven stimulus types, resulting in 15 blocks per stimulus type.

2.4. Imaging parameters and analysis

Imaging data were acquired with a Siemens Magnetom TRIO 3-T whole-body MRI. During data collection, an upgrade was performed to TIM TRIO such that

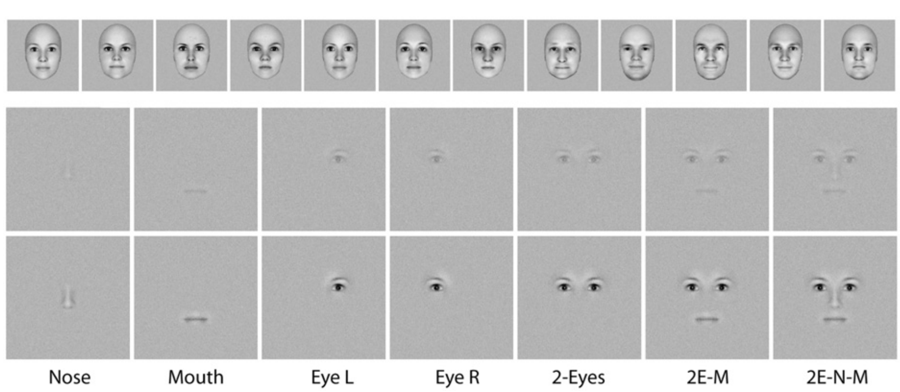

Fig. 1. Single feature and combination stimuli. The top row shows the 12 different faces from which the stimuli were drawn. In the bottom two rows, from left to right, the stimuli are single features of nose, mouth, left eye, and right eye; and combinations of two-eyes, eyes–mouth, and eyes–nose–mouth. Stimuli in the second from bottom row are shown at a contrast level equal to the mean threshold contrast across subjects for the eyes–nose–mouth stimulus, which was the highest contrast level used in the scanner. Stimuli in the bottom row are shown at 10 times that contrast level too make the stimuli easier for the reader to view.