Neuron, Vol. 44, 889–898, December 2, 2004, Copyright ©2004 by Cell Press

Face Perception: Domain Specific, Not Process Specific

Galit Yovel* and Nancy Kanwisher
McGovern Institute for Brain Research
Department of Brain and Cognitive Sciences
Massachusetts Institute of Technology
Cambridge, Massachusetts 02139

Summary

Evidence that face perception is mediated by special cognitive and neural mechanisms comes from fMRI studies of the fusiform face area (FFA) and behavioral studies of the face inversion effect. Here, we used these two methods to ask whether face perception mechanisms are stimulus specific, process specific, or both. Subjects discriminated pairs of upright or inverted faces or house stimuli that differed in either the spatial distance among parts (configuration) or the shape of the parts. The FFA showed a much higher response to faces than to houses, but no preference for the configuration task over the part task. Similarly, the behavioral inversion effect was as large in the part task as the configuration task for faces, but absent in both part and configuration tasks for houses. These findings indicate that face perception mechanisms are not process specific for parts or configuration but are domain specific for face stimuli per se.

Introduction

Extensive behavioral and neural evidence suggests that special mechanisms are engaged in the processing of faces (Bentin et al., 1996; Kanwisher, 2000; Moscovitch et al., 1997; Yin, 1969). However, the extent to which the FFA is specific for faces and the precise nature of the processing that it carries out remain matters of active debate. The goal of the present study was to investigate whether apparently face-specific mechanisms are *domain specific* (i.e., engaged by faces, regardless of the type of processing), *process specific* (i.e., engaged in a specific process that can be applied to any stimulus class), or both stimulus and process specific (i.e., engaged in a specific process that is applied only to faces). To investigate these questions, we made use of the two most well-established markers of face processing: fMRI response of the fusiform face area or "FFA" (Kanwisher et al., 1997; McCarthy et al., 1997), and the drop in face perception performance when faces are presented upside down, i.e., the "face inversion effect" (Yin, 1969).

FFA

The substantially stronger fMRI response to faces than other kinds of stimuli in the mid-fusiform gyrus (Kanwisher et al., 1997; McCarthy et al., 1997) has been replicated in a large number of studies (Kanwisher, 2004). Evidence that this region is not only activated by but is *necessary for* face perception comes from studies showing that patients apparently lacking this area are severely impaired on face perception tasks (Barton et al., 2002; Wada and Yamamoto, 2001). However, there is considerable disagreement in the literature concerning both the nature of the processing that occurs in the FFA and the question of whether the FFA is exclusively involved in face perception (Gauthier and Nelson, 2001; Gauthier et al., 1999; Haxby et al., 2001). Here, we tackle both issues by attempting to induce face-like processing on nonface stimuli. This strategy provides a critical test of our hypotheses, because a strong engagement of the FFA when nonface stimuli are processed like faces would argue strongly against domain specificity and for process specificity.

A serious effort to induce face-like processing of nonface stimuli requires a good hypothesis about how faces are processed. Several suggestions have been proposed in the literature, for example, that face processing is holistic (for review, see Farah et al., 1998) or requires expertise (Diamond and Carey, 1986; Gauthier and Nelson, 2001). Here we test the widespread view that face perception critically involves configural processing, that is, the precise distances among face parts (Freire et al., 2000; Le Grand et al., 2001; Mondloch et al., 2002; Rhodes et al., 1993). Thus, our study specifically tested whether the FFA extracts configural information, rather than part-based information, from faces and nonfaces.

We designed two closely-matched perceptual tasks for overall difficulty, in which subjects had to discriminate sequentially presented image pairs of faces or houses that could differ in only the spatial relations between parts, in one condition, or in only the shapes of parts, in the other condition (Figure 1) (also see Freire et al., 2000; Le Grand et al., 2001 for a similar task with faces). Measuring the strength of the response of the FFA under the four conditions that arise from crossing task (discrimination of configuration versus parts) with stimulus (faces versus houses) enabled us to test each of our hypotheses (see Figure 2). If the FFA is domain specific for faces, we should see a main effect of a higher FFA response to upright faces than to upright houses, with no difference in response between the configuration and part tasks (Figure 2A). If the FFA is process specific for the extraction of configuration, we should see a main effect of a higher FFA response to the configuration rather than to the parts task, with no difference between response to upright faces and upright houses (Figure 2B). If the FFA is specific for both faces and for configural processing, we should see a main effect of a higher response to upright faces than to upright houses and a main effect of a higher response to the configural rather than the parts task (Figure 2C). Finally, if the FFA is specific for configural processing of only faces, we should see an interaction, with the highest response occurring when subjects conduct the configuration task with face stimuli (Figure 2D).

Behavioral Inversion Effect

Long before the advent of fMRI, the behavioral face inversion effect, i.e., the lowered ability to recognize

---

*Correspondence: galit@mit.edu

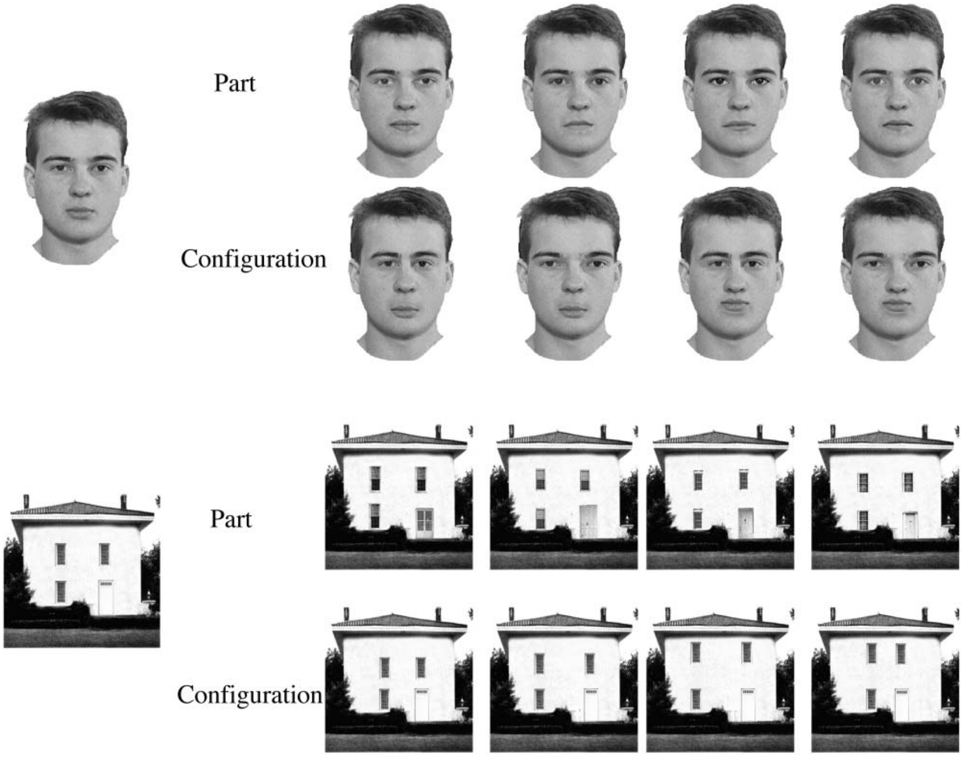

Figure 1. Closely Matched Face and House Tasks Were Used

In each task, an image of a face or a house was manipulated in one of two different ways. For the part set, the shapes of the parts (eyes and mouth in faces, windows and doors in houses) were manipulated to generate four different stimuli that differed in parts, but shared the same configuration. For the configuration set, the spacing between these parts was manipulated to generate four stimuli that shared the same parts, but differed in configuration.

upside-down versus upright faces (Farah et al., 1995; Rhodes and Tremewan, 1994; Valentine, 1988; Yin, 1969), had already been presented as evidence that faces engage special processing mechanisms. Because this inversion effect is larger for faces than for many nonface stimuli (Yin, 1969), it has often been used as a behavioral marker of face-specific processing, in much the same way that the FFA has served as a neural marker of face-specific processing. Similarly, in this study we tested the domain and process specificity of face processing mechanisms by asking whether the inversion effect (i.e., the difference in performance between upright and inverted faces) is larger for faces than for houses, but comparable for the configuration and part tasks (Figure 2A); larger for the configuration versus the part task, but similar on the face and house tasks (Figure 2B); larger for faces than houses and for the configuration versus the part-based task (Figure 2C); or larger for configuration versus the part-based information for faces only (Figure 2D).

It has been suggested that the face inversion effect primarily reflects the loss of configural information when faces are inverted (Freire et al., 2000; Le Grand et al., 2001; Leder and Bruce, 2000). Studies that have tested this hypothesis have reported a larger inversion effect on configuration than part-based matching tasks on faces (Barton et al., 2001; Freire et al., 2000; Le Grand et al., 2001; Leder and Bruce, 2000). However, no study to date, to our knowledge, has tested whether the inversion effect for configural information is specific for faces or whether it might occur when such information is extracted from nonface stimuli. This is the critical condition required to test the domain-specificity and process-specificity hypotheses.

In sum, the current study uses carefully matched stimuli and tasks to investigate the extent to which face processing mechanisms are domain or process specific, using both a neural marker (the response of the FFA) and a behavioral marker (the inversion effect) of face processing (see Figure 2). In a final step, we tested whether these two markers of face processing reflect the same underlying mechanism, a hypothesis that predicts that the two markers should mirror each other, i.e., that the FFA response should be higher for upright than inverted faces (previous studies reported mixed results: Aguirre et al., 1999; Haxby et al., 1999; Kanwisher et al., 1998).

Results

Functional MRI Study

To directly assess whether the FFA is stimulus specific for faces or process specific for the processing of configural information, we compared the responses of the right and left FFA to the configuration and part conditions of upright faces and upright houses. The right and left FFAs were localized in a separate scan during which subjects performed a one-back task on separate blocks of faces and objects. Voxels in the mid-fusiform gyrus that showed a significantly higher response ($p < 10^{-4}$

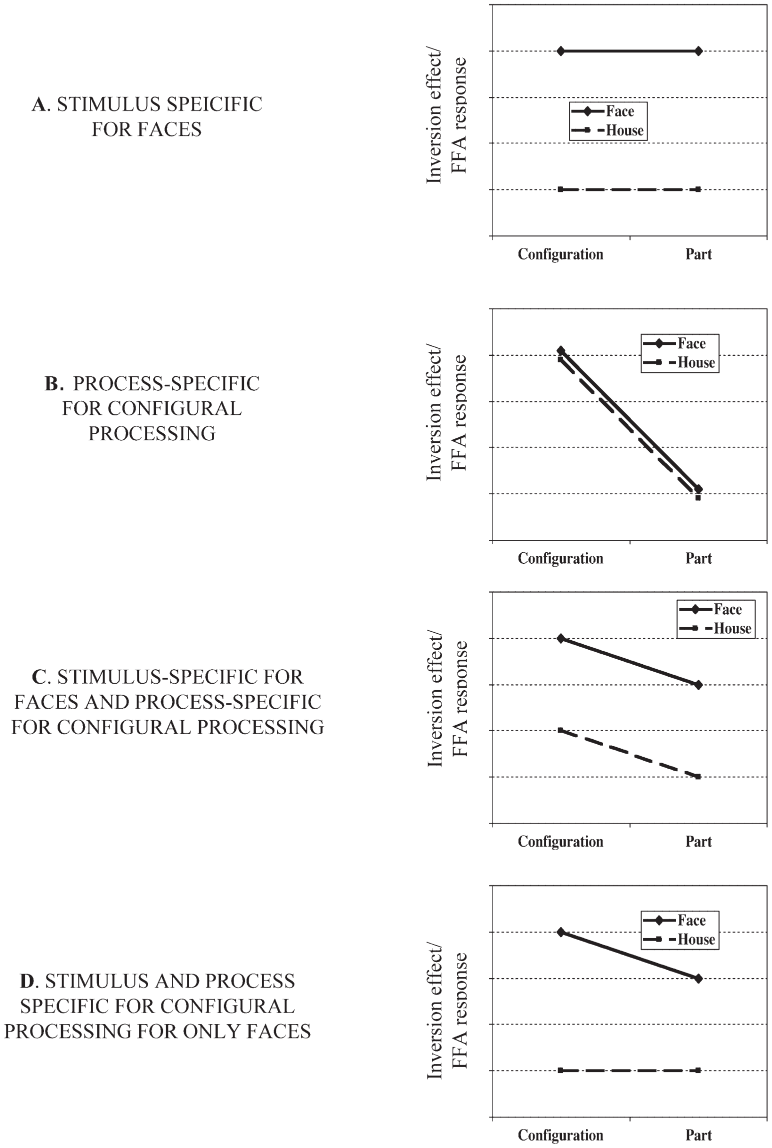

Figure 2. The Four Hypotheses Tested in the Behavioral Inversion Effect and the fMRI Study

The y axis reflects the magnitude of the fMRI response in the FFA or the magnitude of the inversion effect in the behavioral study (performance for upright − inverted stimuli). (A) A higher FFA response and a larger inversion effect for faces than houses on both the configuration and part tasks would support the stimulus (domain)-specificity hypothesis. (B) A higher FFA response and a larger inversion effect for configural than part processing for faces and houses would support the process specificity hypothesis. (C) A higher FFA response and a larger inversion effect to faces than houses and to the configuration task than the part task would support the stimulus *and* process specificity hypothesis. (D) A higher FFA response and a larger inversion effect for faces than houses and for configuration rather than parts for faces only would support a stimulus and process specificity hypothesis for faces only.

uncorrected) to faces than to objects defined the FFA. All 15 subjects showed a right FFA (1710.35 mm^3), but only 14 of them showed also a left FFA (945.62 mm^3). The analysis below was conducted on the 14 subjects who showed both right and left FFAs.

Figure 3 shows the mean percent signal change in the right and left FFA for each condition. A repeated measures ANOVA of percent signal change for upright faces and houses with Hemisphere, Stimulus (Face, House), and Task (Configuration, Part) as within-subject factors revealed a much higher response to faces than houses [main effect of Stimulus, $F(1,13) = 64.04$, $p < 0.0001$] that was larger over the right FFA than the left [interaction of Stimulus and Hemisphere, $F(1,13) = 5.54$, $p < 0.05$]. Further, we found no evidence that the FFA responds more to configural than to part information, either for faces (thus arguing against the hypothesis that the FFA primarily conducts configural processing of faces) or for houses (also arguing against a more general role for the FFA in configural processing). Neither the main effect of Task [$F(1,13) = 1.85$, $p = 0.19$] nor the interaction of Stimulus and Task [$F(1,13) < 1$] was significant. A marginally significant three-way interaction of Hemisphere, Stimulus, and Task [$F(1,13) = 4.38$, $p = 0.06$] reflects a higher response to the house-part than the house-configuration condition ($p < 0.005$) only in the left FFA, in the absence of a difference between the configuration and part tasks for faces in either the right or left FFA or for houses in the right FFA.

Our findings of a similar response to the configuration and part face tasks in the right and left FFAs may at first seem inconsistent with a prior report (Rossion et al., 2000) of a higher response when subjects attended to a whole face relative to a face part in the right FFA and vice versa in the left FFA, in the absence of such a difference for houses. However, in our task, subjects attended to the whole stimulus in both the part and configuration conditions, presumably accounting for the different pattern of FFA activation that we observed.

Second, the higher response to the house part than to the house configuration task in the left FFA contrasts with the similar response for both conditions in the right

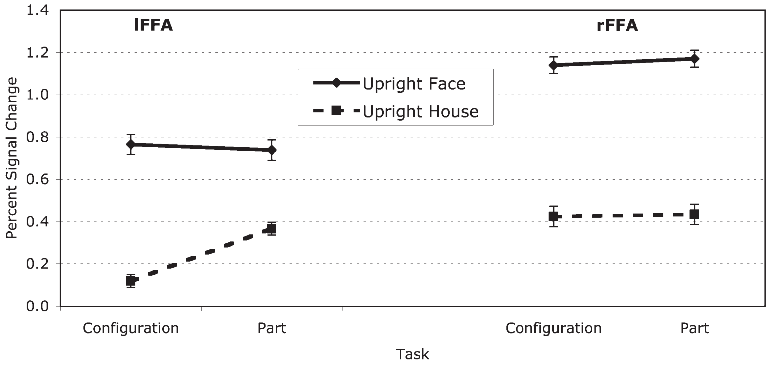

Figure 3. Percent Signal Change in the Right and Left FFA for the Upright Face and Upright House Stimuli Showing a Higher Response to Faces than Houses and a Similar Response to the Configuration and Part Task for Both Stimulus Types

These findings, in particular in the right FFA, support a domain-specific hypothesis (see Figure 2A). Error bars represent the standard error of the difference between responses to the configural and part tasks.

FFA. Interestingly, the left FFA response to houses is similar to the pattern of results we observed in object-selective regions (see below). Although the left FFA shows a higher response to faces than to nonfaces, it may mediate processes that are not as face specific as those of the right FFA and may not be a good marker of face-specific mechanisms. Further evidence for the less specific role that the left FFA might play in face processing is the fact that the response of the left FFA to faces and its face inversion effect are significantly smaller than those of the right FFA (see below). Finally, the size of the left FFA is significantly smaller than that of the right FFA [$t(13) = 3.27$, $p < 0.01$]. We therefore suggest that the right FFA may be engaged in more selective face processing than the left FFA and is therefore a better marker of specific face processing mechanisms.

FFA Inversion Effects

Figure 4 shows that upright faces elicited a larger response than did inverted faces, in particular in the right FFA. A repeated measures ANOVA with Hemisphere, Orientation, and Task revealed a significant face inversion effect [$F(1,13) = 16.5$, $p < 0.001$], larger in the right than the left hemisphere [interaction of Hemisphere and Orientation $F(1,13) = 15.32$, $p < 0.005$]. Prior studies either found no inversion effect in the FFA (Aguirre et al., 1999; Haxby et al., 1999) or a very small one (Kanwisher et al., 1998). The present data reveal, for the first time, a sizable inversion effect in the right FFA (rFFA percent signal increase for faces: 1.15 upright versus 0.85 inverted), strengthening the link between the behavioral inversion effect and the FFA. These findings are consistent with those from intracranial ERP recordings from the surfaces of the fusiform gyrus that show a lower amplitude for inverted than upright faces (McCarthy et al., 1999).

Several behavioral studies have reported greater inversion effects for the configuration task than the part task (Barton et al., 2001; Freire et al., 2000; Le Grand et al., 2001; Leder and Bruce, 2000). Thus, if the FFA is specialized for processing of spatial information, we should see a larger drop in the FFA response to inverted faces in the configuration condition versus the part condition. However, the FFA inversion effect (Figure 4) was large for both the configuration and the part conditions (no interaction of Task and Orientation, $p = 0.65$). This finding further argues against any special role for the FFA in processing spatial information from images of faces (or houses).

Response of Object-Selective Regions

Here we identified in a localizer scan the lateral occipital complex (LOC), which responds more strongly ($p < 10^{-4}$) to images of objects than to scrambled images of objects (Malach et al., 1995; Kanwisher et al., 1996). As Figure 5 shows, the pattern of response of the object-selective region differed in several respects from the response pattern in the FFA, in particular the right FFA. First, the object area showed a higher response to the part rather than the configuration conditions [$F(1,13) = 16.32$, $p < 0.005$] for upright faces and upright houses (no interaction of stimulus and type). Second, there was no difference between responses to houses or faces [$F(1,13) = 3.58$, $p = 0.08$]. Both patterns contrast with the right FFA, which responds more strongly to faces than houses and responds similarly in the configuration and part tasks [interaction of Area (LOC, rFFA) and Task

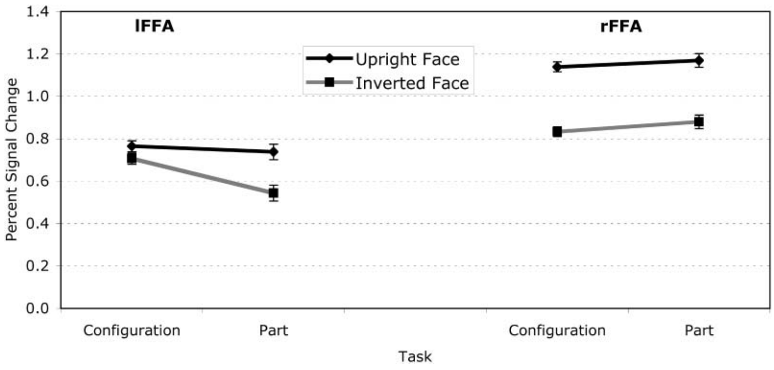

Figure 4. A Similar Right FFA Face Inversion Effect—Drop in fMRI Signal for Inverted Relative to Upright Stimuli—Was Seen on the Part and the Configuration Tasks

These findings mirror our behavioral inversion effect findings and suggest that upright-face processing involves the extraction of both configural and part information. Error bars represent the standard error of the difference between responses to the upright and inverted faces.

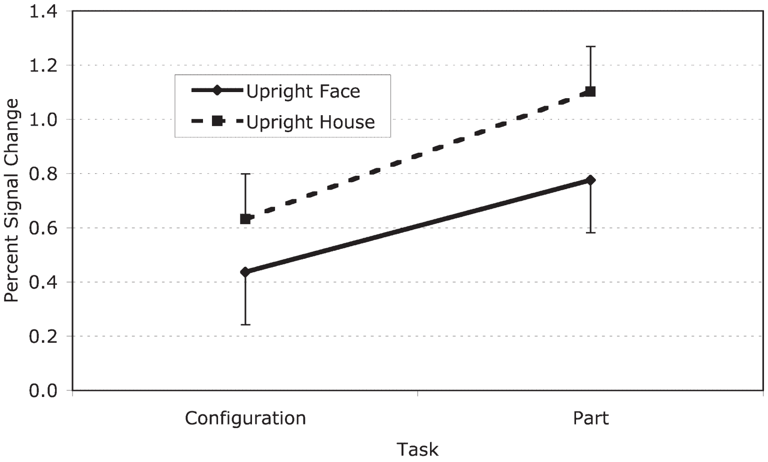

Figure 5. Percent Signal Change in Object-Selective Regions, Showing a Higher Response to the Part Than the Configuration Task

This pattern of response differs significantly from the rFFA response that did not differ for parts and configuration (see Figure 3). Error bars represent the standard error of the difference between responses to the configural and part tasks.

$F(1,13) = 11.30$, $p < 0.005$]. These results underline the functional differences between the FFA and nearby shape-processing regions.

Results of Behavioral Data Collected in the Scanner

To ensure that differences in the fMRI signal during the different task conditions are not due to differences in task difficulty, we asked whether subjects showed similar levels of performance during the configuration and part tasks for upright faces and houses. A repeated measures ANOVA with Stimulus (Face, House) and Task (Configuration, Part) was performed on accuracy in the matching task. Performance for houses (84%) was slightly higher than for faces (78%) [$F(1,15) = 6.93$, $p < 0.02$]. Neither the effect of Task ($p = 0.28$) nor the interaction of Task and Stimulus ($p = 0.80$) was significant. Performance for upright faces did not differ for the part (79%) and configuration tasks (78%) [$F(1,15) = 1.64$, $p = 0.22$]. Similarly, for the house stimuli, performance in the configuration task (84%) did not differ from that in the part task (85%) [$F(1,15) < 1$]. As for the fMRI findings, behavioral data that were collected in the scanner revealed an inversion effect (i.e., better performance for upright than inverted stimuli) in both the face configuration [$t(15) = 2.14$, $p < 0.05$] and the face part tasks [$t(15) = 5.31$, $p < 0.001$]. An interaction of Orientation and Task [$F(1,15) = 7.98$, $p < 0.05$] reflects a larger inversion effect in the part than the configuration task.

Behavioral Study

The finding that the FFA did not show a higher response in the configuration versus the part task for either faces or houses argues against the process-specific hypothesis, that the FFA responds selectively during processing of spatial information related to any stimuli. However, the FFA may respond automatically to faces regardless of the type of information the subject extracts from the stimuli. We therefore sought converging evidence using the behavioral inversion effect, which has been shown to be sensitive to such manipulations. In a behavioral study conducted outside the scanner, we included upright faces, inverted faces, upright houses, and inverted houses. If the inversion effect for the configuration and/or part task for houses is smaller than the inversion effect for the configuration and/or part task for faces, respectively (or absent), we can conclude that face processing mechanisms are stimulus specific for faces, not process specific for configural processing.

Inversion Effect

To ensure that our findings are not confounded with task difficulty, we first asked whether performance in processing upright faces and houses was equivalent across the configuration and part tasks. Performance for upright faces was 79% for the configuration task and 77% for the part task. Performance for upright houses was 79% for the configuration task and 81% for the part task. A repeated measures ANOVA with Stimulus (Face/House) and Task (Configuration/Part) yielded neither a main effect of Task [$F(1,73) < 1$] nor a main effect of Stimulus [$F(1,73) = 2.63$, $p = 0.11$]. The difference between configuration and part tasks was not significant for either faces [$F(1,73) = 3.08$, $p = 0.08$] or houses [$F(1,73) = 1.20$, $p = 0.27$] (Figure 6). Thus, our upright stimuli are well matched for difficulty both across face and house stimuli, and across configuration and part tasks. The critical finding of this experiment, shown in Figure 6C, is the lower performance for inverted than upright stimuli for both the part and configuration tasks for faces and the absence of any inversion effect for houses. This pattern was confirmed by a significant three-way interaction of Stimulus (Face/House), Orientation (Upright/Inversion), and Task (Configuration/Part) [$F(1,73) = 13.80$, $p < 0.0001$]. For faces, the inversion effect was highly significant for both part [$F(1,73) = 103.01$, $p < 0.0001$] and configuration tasks [$F(1,73) = 34.3$, $p < 0.0001$]. An Orientation by Task interaction for faces [$F(1,73) = 12.97$, $p < 0.0001$] reflects a larger inversion effect in the face part than in the face configuration task. House stimuli elicited no inversion effect for either part or configuration tasks. Performance was higher for the inverted versus the upright house-part task [$F(1,73) = 5.78$, $p = 0.02$] and there was no orientation effect for the house-configuration task [$F(1,73) < 1$]. Thus, our behavioral marker of face processing (i.e., the inversion effect) mirrors the findings from our neural marker (FFA response); both markers show (1) a similar effect for the configuration and part tasks for faces (suggesting a lack of specificity for the processing of config-

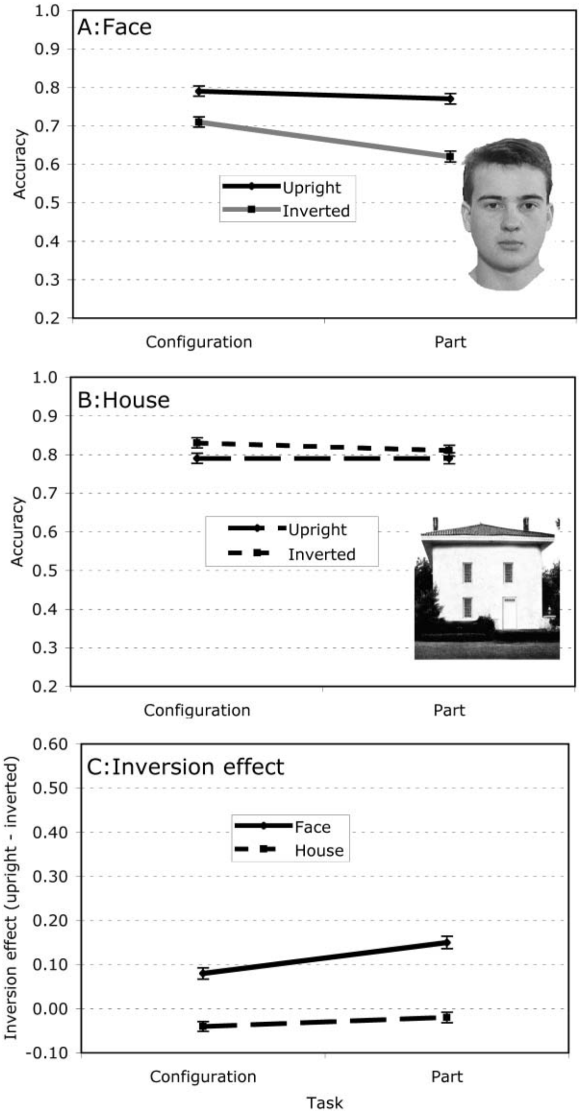

Figure 6. Lower Performance for Inverted Faces Than for Upright Faces on the Configural and Part Tasks

(A) No inversion effect for houses (B) Note that performance levels for the upright conditions were matched across the two tasks for faces and houses. The difference between performance for upright and inverted stimuli (inversion effect) that for the configural and part tasks was present for faces, but not for houses (C), parallels the rFFA findings and suggests that mechanisms of face perception are stimulus specific for faces rather than process specific for configural processing (see Figure 2A). Error bars represent the standard error of the difference between responses to the configural and part tasks.

ural information), and (2) a much stronger effect for faces than houses (supporting stimulus specificity).

What Is the Significance of the Larger Inversion Effect for the Part versus the Configuration Task? According to one alternative hypothesis, the inversion effect that we observed during performance of the part task reflects subjects' use of a configural strategy even in this condition. After all, even though we tried hard to minimize these effects, any change in parts will inevitably produce a subtle change in configuration. However, if subjects use only configural information to discriminate faces in our part task, performance on the part task should be much lower than on the configuration task, as the difference in configural information between the faces in the part set is much more subtle than in the configuration set (see Figure 1). The fact that we were able to match performance on the configuration and part tasks suggests that subjects did not rely on configural information to discriminate among faces in the part set (see Supplemental Data http://www.neuron.org/cgi/content/full/44/5/889/DC1/ for reaction time analyses). Although the exact reason for the larger inversion effect in the part than the configuration task that we observed is not clear, the key point of our study is simply that inversion effects are at least as strong for both configuration and part discrimination tasks.

Resolving Discrepancies with Prior Behavioral Studies. The effect of stimulus exemplar. Because our findings of inversion effects for faces in both the configuration and part tasks are inconsistent with those of some previous reports (Barton et al., 2001; Freire et al., 2000; Le Grand et al., 2001; Leder and Bruce, 2000), we explored this effect further by creating another set of configuration and part stimuli from another face exemplar (presented in Figure 1). We again manipulated the configuration and the part information in a way that yielded similar performance on the two tasks for the upright condition (77% for configuration and 76% for part condition) [$t(15) = 0.487$, $p = 0.63$]. The inversion effect was again highly significant in both the configuration [$t(15) = 3.06$, $p < 0.01$] and the part tasks [$t(15) = 5.21$, $p < 0.001$]. A 3-way ANOVA in which Face exemplar was also included as a factor, revealed no interaction of Face with the Orientation and/or Task factors, which suggests that the pattern of results we found are not specific for the face exemplar that we chose, but instead generalizeable across different face exemplars (Riesenhuber et al., 2004).

The effect of task design: mixed versus blocked. Finally, previous studies that reported a larger inversion effect for configuration than part face stimuli used a blocked design in which subjects were informed about the way that the faces were manipulated. Our behavioral study used a mixed design in which subjects were not informed about the type of manipulation. However, as reported above, in the fMRI experiment we used a blocked design and informed subjects before each block the nature of the task manipulation (configuration or parts). As reported in the previous section, our findings show that even when subjects are instructed about whether to look for configuration or part differences, the inversion effect was larger for the part than for the configuration condition. Thus, consistent with our fMRI results, the behavioral inversion effects seen across different face stimuli and task designs suggest that specialized face perception mechanisms are stimulus specific for faces rather than process specific for configural processing.

Discussion

We used two independent markers of face processing, the response of the FFA measured with fMRI and the face inversion effect measured behaviorally, to ask (1) whether the mechanisms used in face perception are specifically engaged in extracting information about spatial distances among parts, and (2) whether these mechanisms can also be engaged in the processing of nonfaces. In both studies, subjects matched face or house stimuli that differed either in configuration or part information. Our fMRI and behavioral results mirrored each other perfectly. First, contrary to a widespread view in the literature (Barton et al., 2002; Freire et al., 2000; Le Grand et al., 2001; Leder and Bruce, 2000; Mondloch et al., 2002), we found no evidence that face processing mechanisms are specifically engaged in the extraction of spatial distance among parts: both FFA responses (Figure 3) and behavioral inversion effects (Figure 6C) are no greater when subjects discriminate faces based on configural information (i.e., the distance among parts) than when the discriminations are based on the identity of individual face parts. Second, we found no evidence that the FFA responds differently to configural compared to part information in nonfaces, as well as no inversion effect for the configural or the part house task. Finally, evidence that our neural marker (the FFA response) and behavioral marker (the inversion effect) of face processing reflect the same underlying mechanisms comes from the clear inversion effect that we observed with fMRI in the FFA for both configuration and part processing of faces (Figure 4). Thus, our data provide strong evidence that common face processing mechanisms reflected in both FFA activity and in behavioral inversion effects are domain specific for processing faces per se, rather than process specific for extracting configural information (Figure 2A). Next we discuss these results and their implications in more detail.

Characterizing Face Processing

Our results of similar FFA responses and similar inversion effects for the configuration and part tasks for faces argue against the widespread view that faces primarily engage configural processing. Several prior studies have reported larger inversion effects for the configuration than for the part tasks. A few of these studies, however, did not manipulate the shape of a face part but instead its brightness or color (Barton et al., 2002; Leder and Bruce, 2000; Searcy and Bartlett, 1996). Such color discrimination tasks might be primarily mediated by lower-level visual processes, and, to a lesser extent, by specialized face mechanisms, and therefore may not provide valid tests of the extent to which the face processing system is specialized for configural processing but not for face parts. A number of reports that did manipulate the shape of face parts used the “Jane” face (Le Grand et al., 2001; Mondloch et al., 2002) and found a larger inversion effect for the configuration than the part task. However, in these studies, the part task is easier than the configuration task, which may account for the relatively small inversion effect seen during this task (see also Freire et al., 2000, for a similar pattern of results with a different face exemplar). A valid comparison of inversion effects requires that the two tasks be matched for performance under the upright condition. Further, performance on both tasks should be in the middle range to avoid ceiling or floor effects. Indeed, Mondloch et al. (2002) reported that when they increased the distance among face parts of the “Jane” face beyond the distance they used in their study, the inversion effect for the configuration task was also eliminated. This finding further highlights the effect of task difficulty on the magnitude of the inversion effect and the importance of matching task difficulty on both tasks. When task difficulty is matched, we find no evidence for a larger inversion effect on the configuration versus the part task (see also Malcom et al., 2005; Riesenhuber et al., 2004). The finding of a similar FFA response to configuration and part tasks argues against a differential role for processing of configuration and parts in face perception. This conclusion is consistent with current computational models of face processing (Biederman and Kalocsai, 1997; Riesenhuber and Poggio, 1999). Furthermore, our findings are consistent with reports that C.K., a neurological patient who suffers from object agnosia, nonetheless shows intact processing of both configural and part information from faces (Moscovitch et al., 1997).

Evidently, the distinction between configuration and part information does not capture the essence of face processing and the way it differs from the processing of nonfaces. Another related but distinct proposal is that faces are processed more holistically than are nonface stimuli (Farah et al., 1995; Tanaka and Farah, 1993, 2003; Tanaka and Sengco, 1997; Young et al., 1987). While the precise meaning of “holistic processing” has been operationalized in a variety of different ways in prior studies (Tanaka and Farah, 1993; Williams et al., 2004; Young et al., 1987; Yovel et al., 2005), the crux of the idea is that face perception entails the mandatory processing of the face as a whole, rather than the independent and noninteractive processing of individual parts. This idea is not inconsistent with the present finding of strong FFA engagement and behavioral (and fMRI) inversion effects for the part task, because subjects may be processing the parts in an interactive (and hence “holistic”) manner in our part task.

Domain Specificity of Face Processing

The second question addressed in this study is whether face processing engages cognitive and neural mechanisms that are domain specific for processing faces. In an effort to provide a strong test of this hypothesis, we attempted to induce face-like processing of nonface stimuli (houses) by matching discrimination tasks on houses as closely as we could to discrimination tasks on faces. This effort failed dramatically: we found neither an inversion effect for houses nor a strong FFA engagement for house stimuli in either the configuration or the part task. These data support the domain specificity of face processing. At the same time, the challenge remains for others to explore a better characterization of face processing (perhaps along the “holistic” lines discussed above) to induce “face-like” processing of nonfaces. If such efforts ever produce strong engagement of the FFA by nonfaces and face-sized behavioral inversion effects, that will argue against the domain specificity of face processing.

Several researchers have argued that objects of expertise present just such a case: they are processed like faces and they engage the FFA. However, the evidence for this hypothesis is weaker than generally appreciated (for review, see McKone and Kanwisher, 2005), and several new findings argue strongly against it. For example, Duchaine et al. (2004) have demonstrated a powerful dissociation in a developmental prosopagnosic subject between face recognition (severely impaired) and the acquisition of expertise with novel objects (i.e., normal "Greeble expertise"). Also, trial-by-trial correlations have been demonstrated between face identification performance and the magnitude of both the FFA and face-selective M170 responses, yet no such effects were observed for car identification in car experts (Grill-Spector et al., 2004; Xu et al., 2005). Thus, the current literature on expertise does not contain compelling evidence against the domain specificity of face processing.

Finally, the pattern of response of the FFA during part and configuration tasks with face and house stimuli was markedly different than that of object-selective regions (Figure 5), further highlighting the fact that face processing is qualitatively different from object processing.

Conclusion

In summary, contrary to a widespread assumption in the face processing literature, we found no evidence that the face system is specialized for processing the spatial relations among parts to a greater extent than for processing the shape of those parts. Further, we were unable to engage face processing mechanisms during the perception of nonface stimuli, even after matching the perceptual requirements as closely as possible for face and nonface tasks. These findings argue that the mechanisms used in face perception are not process specific for the processing of configural information, but are instead domain specific for faces.

Experimental Procedures

Functional MRI

Subjects

Seventeen healthy subjects participated in the experiment. Prior to the analysis, two subjects were excluded due to excessive head motion in the scanner (>8 mm).

Stimuli

Stimuli were presented using Psychtoolbox implemented in MATLAB (Brainard, 1997). Photoshop was used to generate the configuration and part sets of selected face and house images.

Face Stimuli. Two sets of four face stimuli were generated from a picture of a male face. The Configuration set consisted of four faces in which eyes were either closer or farther apart and the mouth was either closer or farther from the nose. In the Part set, the two eyes and the mouth were replaced in each of four faces by eyes and mouths of similar shapes from different faces. Figure 1 shows a face stimulus generated by the same procedure, which yielded similar behavioral data (see Results section) as the face stimulus used in the fMRI experiment (that face is not presented in the figure because we could not find the owner of the face to obtain permission to publish it).

The generation of the face stimuli followed the method used by Le Grand et al. (2001), except for one key difference. The configuration and part manipulations were based on performance in a behavioral pilot study. In particular, we manipulated the stimuli until they yielded an average performance level of about 80% correct (a dynamic range) in both the configuration and the part tasks for both the face and house stimuli. The 21 subjects who participated in this pilot study were not included in the fMRI experiment or the main behavioral experiment that we report here.

House Stimuli. A method similar to the one we used to create the face stimuli was used to create the house stimuli. An image of a house, which included two upper windows, a lower left-side window and a lower right-side door, was manipulated in one of two ways. In the Configuration set, four houses had the left-side windows and the right-side window and door closer together or farther apart or the right and left upper windows closer to or farther from the roof. In the Part set, all three windows and the door were replaced by windows and doors taken from different house images (See Figure 1).

All possible combinations from the original and the four different manipulated stimuli yielded 20 different pairs for the configuration condition and 20 different pairs for the part condition.

After we selected the configuration and part stimuli that yielded similar performance levels, we measured the physical difference for each of the different pairs of configuration and part stimuli of faces and houses. Specifically, for each stimulus pair that was presented in the experiment, we computed the Euclidian distance (the square root of the sum of the squared difference between each pixel of one image and its corresponding pixel on the second image). There were ten possible pairs of different stimulus exemplars for each condition (Face/House × Configuration/Part). The physical difference for the faces in the configuration set (mean, 1344.7; SD, 258.6) was similar to the physical difference in the part set (mean, 1094.9; SD, 175.3). For houses, the physical difference for the house configuration set (mean, 8600.3; SD, 1340.1) was in the same range (i.e., within 2 standard deviations) as the house part set (mean, 6893.1; SD, 1031.4).

Procedure

fMRI Data Acquisition. Scanning was performed with a Siemens 3.0-T research scanner at the MGH-NMR Center in Charlestown, MA. During the anatomical scan we collected 128 sagittal slices (TR = 11 ms; TE = 4 mm; FOV = 256). For the functional scans, 28 contiguous 4 mm thick axial slices (matrix size: 64 × 64) were oriented parallel to the temporal lobe and covered the entire brain. Gradient echo pulse sequences with TR 2 s; TE 30 ms; flip angle 90°; resolution 3.1 × 3.1 × 4 mm were used.

The experiment consisted of six runs of the experimental task and five runs of the localizer (described below), which were presented interleaved starting with the experimental task.

Experimental Task. Each experimental run included six blocked conditions: Face Upright Configuration; Face Upright Part; Face Inverted Configuration; Face Inverted Part; House Upright Configuration; House Upright Part. To allow maximal power to these relevant six conditions, inverted houses were not included in the experiment because these stimuli were not critical to test our hypotheses (see Figure 2). Each scan consisted of 20 s epochs of each of the six experimental conditions interleaved with 20 s of fixation (baseline) epochs. The baseline epoch included 17 s of fixation, 2 s of a cue, which specified the type of task (e.g., Face Upright Part), and 1 s of fixation, after which the specified condition commenced (total of 260 s). Thus, the subjects were informed before each block by a lexical cue on which dimension the stimuli would vary. This ensured that subjects indeed attended and processed the relevant dimension (configuration or parts) in each of the different conditions.

Each experimental epoch included ten trials. Each trial consisted of a first face or house stimulus for 250 ms, an interstimulus interval for 500 ms, and a second face or house stimulus for 250 ms. The intertrial interval lasted 1000 ms, during which subjects pressed one key for "same" and another for "different" to indicate their response. The second stimulus was presented in a different location, so apparent motion cannot be used to discriminate pairs that differ in configuration. The order of the six conditions was counterbalanced across runs, such that each condition was preceded and followed by a different condition across the runs.

Localizer Scan. The localizer scan consisted of five stimulus categories, including faces, objects, and scrambled images of the objects. The two other categories were not included in the analyses reported here. Each scan consisted of two groups of five consecutive 16 s epochs, one for each category, presented in one order in one group and the opposite order in the other group. The blocks were preceded and followed by 16 s fixation (total time 208 s). The serial position of the categories was counterbalanced across scans. Each image was presented for 200 ms with 800 ms blank intervals between presentations. Subjects were instructed to press a key whenever they noticed two images repeated in a row (one-back).

This happened twice per epoch and ensured that subjects were awake and attentive. Eighty different images were presented for each category across the five runs.

fMRI Data Analysis

Data were analyzed separately for each subject using Fsfast (http://surfer.nmr.mgh.harvard.edu/). Motion correction was performed prior to data analysis and was conducted using the AFNI motion correction algorithm (Cox and Jesmanowicz, 1999) to align all the images to the first time image of the first run. Data from the localizer were spatially smoothed with a Gaussian filter (full-width half-maximum = 5 mm) and were used to define the ROIs separately for each subject. Data from the experimental task were not smoothed. A γ function with delta = 2.25 and tau = 1.25 was used to estimate the hemodynamic response (HDR).

ROI Analysis. For each subject, voxels that responded significantly more to faces than to objects (t test, $p < 10^{-4}$) in the localizer scan were defined as face-selective regions. Voxels that responded more to objects than to scrambled images of objects were defined as object-selective regions. fROI (http://froi.sourceforge.net/) was used to select face-selective voxels in the right and left fusiform gyrus and object-selective voxels in the lateral occipital and the ventral temporal regions for each subject individually. The response to each of the experimental conditions was the amplitude of the HDR averaged across the six runs and across the voxels included in the ROIs.

Behavioral Experiment

Subjects

Seventy-seven subjects participated in the behavioral experiment. One subject did not complete the house-upright and inverted tasks, one did not complete the house-upright task, and one did not complete the face-inverted task. These subjects were excluded from further analyses.

Procedure

Subjects completed the face and the house matching tasks in addition to seven other tasks, each designed to test a different hypothesis (face memory, face/house categorization, grating detection, gender, expression, gaze, and identity categorization). Here we report only results from the face- and house-upright and -inverted tasks. The upright task was presented before the inverted task. Each trial started with a 500 ms fixation dot at the center of the screen. A first stimulus was presented for 250 ms followed by a 1000 ms interstimulus interval during which the fixation dot was on the screen. The second face was on the screen for 250 ms. Subjects used one key to indicate that the presenting image was the same as the former image and another key to indicate its difference, after which a fixation dot appeared again. Subjects had to press the space bar to initiate the next trial. The part and configuration trials were presented in a randomized order. Subjects were not informed that the stimuli would differ in configuration or parts, but were only told that some of the pairs were identical whereas others were slightly different.

Each of the 20 different pairs of the configuration and the part sets was presented once during the task. Same pairs included repeated presentation of each of the four faces in the configuration set, the four faces in the part set, or the original face (see Figure 1). Each of these repeated pairs was presented four times to yield a similar number of same and different trials.

Acknowledgments

We would like to thank Chris Baker, Hans Op de Beeck, Elinor McKone, and Rachel Robbins for their comments on the manuscript. We also thank Nao Gamo, Kevin Der, and Michael Ogrydziak for help with behavioral data collection as well as Stephanie Chow for help with stimulus generation. This research was supported by NIMH grant 66696 (N.K.); NEI grant EY13455 (N.K.); the National Center for Research Resources (P41-RR14075, R01 RR16594-01A1, and the NCRR BIRN Morphometric Project BIRN002); and the Mental Illness and Neuroscience Discovery (MIND) Institute.

Received: May 27, 2004
Revised: August 24, 2004
Accepted: October 18, 2004
Published: December 1, 2004

References

Aguirre, G.K., Singh, R., and D'Esposito, M. (1999). Stimulus inversion and the responses of face and object-sensitive cortical areas. Neuroreport *10*, 189–194.

Barton, J.J., Keenan, J.P., and Bass, T. (2001). Discrimination of spatial relations and features in faces: Effects of inversion and viewing duration. Br. J. Psychol. *92*, 527–549.

Barton, J.J., Press, D.Z., Keenan, J.P., and O'Connor, M. (2002). Lesions of the fusiform face area impair perception of facial configuration in prosopagnosia. Neurology *58*, 71–78.

Bentin, S., Allison, T., Puce, A., Perez, E., and McCarthy, G. (1996). Electrophysiological studies of face perception in humans. J. Cogn. Neurosci. *8*, 551–565.

Biederman, I., and Kalocsai, P. (1997). Neurocomputational bases of object and face recognition. Philos. Trans. R. Soc. Lond. B Biol. Sci. *352*, 1203–1219.

Brainard, D.H. (1997). The psychophysics toolbox. Spat Vis. *10*, 433–436.

Cox, R.W., and Jesmanowicz, A. (1999). Real-time 3D image registration for functional MRI. Magn. Reson. Med. *42*, 1014–1018.

Diamond, R., and Carey, S. (1986). Why faces are and are not special: an effect of expertise. J. Exp. Psychol. Gen. *115*, 107–117.

Duchaine, B.C., Dingle, K., Butterworth, E., and Nakayama, K. (2004). Normal greeble learning in a severe case of developmental prosopagnosia. Neuron *43*, 469–473

Farah, M.J., Tanaka, J.W., and Drain, H.M. (1995). What causes the face inversion effect? J. Exp. Psychol. Hum. Percept. Perform. *21*, 628–634.

Farah, M.J., Wilson, K.D., Drain, M., and Tanaka, J.N. (1998). What is "special" about face perception? Psychol. Rev. *105*, 482–498.

Freire, A., Lee, K., and Symons, L.A. (2000). The face-inversion effect as a deficit in the encoding of configural information: direct evidence. Perception *29*, 159–170.

Gauthier, I., and Nelson, C.A. (2001). The development of face expertise. Curr. Opin. Neurobiol. *11*, 219–224.

Gauthier, I., Tarr, M.J., Anderson, A.W., Skudlarski, P., and Gore, J.C. (1999). Activation of the middle fusiform 'face area' increases with expertise in recognizing novel objects. Nat. Neurosci. *2*, 568–573.

Grill-Spector, K., Knouf, N., and Kanwisher, N. (2004). The fusiform face area subserves face perception, not generic within-category identification. Nat. Neurosci. *7*, 555–562.

Haxby, J.V., Ungerleider, L.G., Clark, V.P., Schouten, J.L., Hoffman, E.A., and Martin, A. (1999). The effect of face inversion on activity in human neural systems for face and object perception. Neuron *22*, 189–199.

Haxby, J.V., Gobbini, M.I., Furey, M.L., Ishai, A., Schouten, J.L., and Pietrini, P. (2001). Distributed and overlapping representations of faces and objects in ventral temporal cortex. Science *293*, 2425–2430.

Kanwisher, N. (2000). Domain specificity in face perception. Nat. Neurosci. *3*, 759–763.

Kanwisher, N.G. (2003). The ventral visual object pathway in humans: Evidence from fMRI. In The Visual Neurosciences, L. Chalupa and J. Werner, eds. (Cambridge, MA: MIT Press), pp. 1179–1189.

Kanwisher, N., Chun, M.M., McDermott, J., and Ledden, P.J. (1996). Functional imaging of human visual recognition. Brain Res. Cogn. Brain Res. *5*, 55–67.

Kanwisher, N., McDermott, J., and Chun, M.M. (1997). The fusiform face area: A module in human extrastriate cortex specialized for face perception. J. Neurosci. *17*, 4302–4311.

Kanwisher, N., Tong, F., and Nakayama, K. (1998). The effect of face inversion on the human fusiform face area. Cognition *68*, B1–B11.

Le Grand, R., Mondloch, C.J., Maurer, D., and Brent, H.P. (2001). Neuroperception. Early visual experience and face processing. Nature *410*, 890.

Leder, H., and Bruce, V. (2000). When inverted faces are recognized:

the role of configural information in face recognition. Q. J. Exp. Psychol. A *53*, 513–536.

Malach, R., Reppas, J.B., Benson, R.R., Kwong, K.K., Jiang, H., Kennedy, W.A., Ledden, P.J., Brady, T.J., Rosen, B.R., and Tootell, R.B. (1995). Object-related activity revealed by functional magnetic resonance imaging in human occipital cortex. Proc. Natl. Acad. Sci. USA *92*, 8135–8139.

Malcom, G., Leung, C., and Barton, J.J.S. (2005). Regional salience and the inversion effect for faces: differential effects for feature shape, spatial relations, and external contour. Perception, in press.

McCarthy, G., Puce, A., Gore, J.C., and Allison, T. (1997). Face-specific processing in the human fusiform gyrus. J. Cogn. Neurosci. *9*, 605–610.

McCarthy, G., Puce, A., Belger, A., and Allison, T. (1999). Electrophysiological studies of human face perception. II: Response properties of face-specific potentials generated in occipitotemporal cortex. Cereb. Cortex *9*, 431–444.

McKone, E., and Kanwisher, N. (2005). Does the human brain process objects of expertise like faces? A review of the evidence. In From Monkey Brain to Human Brain, S. Dehaene, J.-R. Duhamel, M.D. Hauser, and G. Rizzolatti, eds. (Cambridge MA: MIT Press), in press.

Mondloch, C.J., Le Grand, R., and Maurer, D. (2002). Configural face processing develops more slowly than featural face processing. Perception *31*, 553–566.

Moscovitch, M., Winocur, G., and Behrmann, M. (1997). What is special about face recognition? Nineteen experiments on a person with visual object agnosia and dyslexia but normal face recognition. J. Cogn. Neurosci. *9*, 555–604.

Rhodes, G., and Tremewan, T. (1994). Understanding face recognition: caricature effects, inversion, and the homogeneity problem. Visual Cognition *1*, 275–311.

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

Riesenhuber, M., and Poggio, T. (1999). Hierarchical models of object recognition in cortex. Nat. Neurosci. *2*, 1019–1025.

Riesenhuber, M., Jarudi, I., Gilad, S., and Sinha, P. (2004). Face processing in humans is compatible with a simple shape-based model of vision. Biol. Lett., in press.

Rossion, B., Dricot, L., Devolder, A., Bodart, J.M., Crommelinck, M., De Gelder, B., and Zoontjes, R. (2000). Hemispheric asymmetries for whole-based and part-based face processing in the human fusiform gyrus. J. Cogn. Neurosci. *12*, 793–802.

Searcy, J.H., and Bartlett, J.C. (1996). Inversion and processing of component and spatial-relational information in faces. J. Exp. Psychol. Hum. Percept. Perform. *22*, 904–915.

Tanaka, J., and Farah, M. (2003). Holistic face recognition. In Perception of Faces, Objects and Scenes. Analytic and Holistic Processes, M.A. Peterson, and G. Rhodes, eds. (New York: Oxford University Press) pp. 53–74.

Tanaka, J.W., and Farah, M.J. (1993). Parts and wholes in face recognition. Q. J. Exp. Psychol. A. *46*, 225–245.

Tanaka, J.W., and Sengco, J.A. (1997). Features and their configuration in face recognition. Mem. Cognit. *25*, 583–592.

Valentine, T. (1988). Upside-down faces: a review of the effect of inversion upon face recognition. Br. J. Psychol. *79*, 471–491.

Wada, Y., and Yamamoto, T. (2001). Selective impairment of facial recognition due to a haematoma restricted to the right fusiform and lateral occipital region. J. Neurol. Neurosurg. Psychiatry *71*, 254–257.

Williams, M.A., Moss, S.A., and Bradshaw, J.L. (2004). A unique look at face processing: the impact of masked faces on the processing of facial features. Cognition *91*, 155–172.

Xu, Y., Liu, J., and Kanwisher, N. (2005). The M170 is selective for faces, not for expertise. Neuropsychologia, in press.

Yin, R. (1969). Looking at upside down faces. J. Exper. Psychol. *81*, 141–145.

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

Yovel, G., Paller, K.A., and Levy, J. (2005). A whole face equals the sum of its halves. Interactive processing in face perception. Vis. Cogn., in press.