

## Cortical correlates of face and scene inversion: A comparison

Russell A. Epstein<sup>a,\*</sup>, J. Stephen Higgins<sup>a</sup>, Whitney Parker<sup>a</sup>,  
Geoffrey K. Aguirre<sup>b,c</sup>, Samantha Cooperman<sup>a</sup>

<sup>a</sup> Department of Psychology, University of Pennsylvania, 3720 Walnut St., Philadelphia, PA 19104-6241, USA

<sup>b</sup> Center for Cognitive Neuroscience, University of Pennsylvania, PA, USA

<sup>c</sup> Department of Neurology, University of Pennsylvania, PA, USA

Received 29 August 2004; received in revised form 10 October 2005; accepted 12 October 2005

Available online 21 November 2005

### Abstract

Face recognition is more strongly impaired by stimulus inversion than nonface object recognition. This phenomenon, known as the face inversion effect (FIE), suggests that the visual system contains specialized processing mechanisms that are more engaged by upright faces than by inverted faces or nonface objects. Neuroimaging and neuropsychological studies indicate that environmental scenes may also recruit specialized-purpose processing machinery but a comparable inversion effect for scenes has not been established. Here we demonstrate that both face and scene inversion lead to behavioral penalties during performance of a continuous visual matching task; however, the scene inversion effect was less robust and declined in magnitude over the course of the experiment. Scene inversion led to greater neural response in the functionally defined lateral occipital (LO) object area for inverted versus upright scenes and reduced response in the parahippocampal place area (PPA), while face inversion led to greater response in LO and the right middle fusiform (MF) object area for inverted versus upright faces but no change in the fusiform face area (FFA). A whole-brain analysis revealed several regions that responded more strongly to either upright versus inverted faces or upright versus inverted scenes, some of which may be involved in post-recognition processing. These results demonstrate that both face and scene inversion cause a shift from specialized processing streams towards generic object-processing mechanisms; however, this shift only leads to a reliable behavioral penalty in the case of face inversion.

© 2005 Elsevier Ltd. All rights reserved.

**Keywords:** Functional magnetic resonance imaging; Visual cortex; Object recognition; Fusiform face area; Parahippocampal place area; Lateral occipital complex

### 1. Introduction

Many kinds of objects are more difficult to recognize when they are upside-down than when they are in their canonical orientation. Behavioral studies have demonstrated that this inversion penalty is often stronger for faces than for other stimulus categories (Valentine, 1988; Yin, 1969). This phenomenon – known as the face inversion effect (FIE) – has been taken as support for the claim that faces and nonface objects recruit different processing streams (Farah, Wilson, Drain, & Tanaka, 1998). In particular, face recognition is believed to rely more than non-face recognition on configural or holistic information (Carey & Diamond, 1977; Farah, Tanaka, & Drain, 1995; Maurer, Grand, & Mondloch, 2002; Rhodes, Brake, & Atkinson, 1993). Recent

studies have demonstrated that inversion impairs the ability to detect the configural aspects of faces (such as the distance between the two eyes) but not individual faces features (such as the color or shape of the eyes) (Bartlett & Searcy, 1993; Freire, Lee, & Symons, 2000; Searcy & Bartlett, 1996). There is evidence that similar configural or holistic processing mechanisms might be recruited by other object categories for which we have extensive expertise and are also more difficult to recognize after inversion (Diamond & Carey, 1986; Gauthier & Tarr, 2002; Rossion & Gauthier, 2002).

Claims for specialized processing streams have also been made for scenes (Aguirre, Zarahn, & D'Esposito, 1998; Epstein, 2005; Epstein & Kanwisher, 1998) and bodies (Downing, Jiang, Shuman, & Kanwisher, 2001). Faces, scenes, and bodies differentially activate specific regions of cortex that tend to be found in the same location across subjects (Kanwisher, 2004, but see Haxby et al., 2001) and both scene and face perception can be specifically impaired by brain damage (Aguirre &

\* Corresponding author. Tel.: +1 215 573 3532; fax: +1 215 898 1982.  
E-mail address: epstein@psych.upenn.edu (R.A. Epstein).

D'Esposito, 1999; Epstein, DeYoe, Press, Rosen, & Kanwisher, 2001; Mendez & Cherrier, 2003). Given these claims of specialization, one might expect to find scene and body inversion effects comparable to the FIE. Indeed, a "body-inversion effect" has recently been reported (Reed, Stone, Bozova, & Tanaka, 2003) in which subjects are impaired at discriminating between images of bodies exhibiting different postures when the images are inverted. The magnitude of the body inversion effect was similar to the magnitude of the face inversion effect. In contrast, the few studies that have examined scene inversion reported to have either no effect on recognition performance (Diamond & Carey, 1986; Wright & Roberts, 1996) or a small effect that is significantly less than the FIE (Scapinello & Yarmey, 1970; Yin, 1969).

In the current study, we measure the behavioral effects of face and scene inversion and also use functional magnetic resonance imaging (fMRI) to identify cortical regions that respond differentially to upright versus inverted faces and upright versus inverted scenes. This experiment follows on from a number of earlier studies that examined the effect of inversion on the neural response to faces (Aguirre, Singh, & D'Esposito, 1999; Gauthier, Tarr, Anderson, Skudlarski, & Gore, 1999; Haxby et al., 1999; Kanwisher, Tong, & Nakayama, 1998). Several of these studies focused on the fusiform face area (FFA), a region of cortex that responds much more strongly to faces than to nonface objects and is believed to be critical for face recognition (Kanwisher, McDermott, & Chun, 1997; McCarthy, Puce, Gore, & Allison, 1997; Puce, Allison, Asgari, Gore, & McCarthy, 1996; Rhodes, Byatt, Michie, & Puce, 2004). One might expect that inversion would strongly reduce FFA response to faces. Somewhat surprisingly, these earlier studies indicated that inversion reduces face response in the FFA only mildly (Gauthier et al., 1999; Haxby et al., 1999; Kanwisher et al., 1998; Yovel & Kanwisher, 2004) or not at all (Aguirre et al., 1999). In fact, the greatest effect of face inversion was found in putative "object" areas, which responded *more strongly* to inverted faces than to upright faces (Aguirre et al., 1999; Haxby et al., 1999). This neural effect was interpreted as reflecting greater engagement of generic object-processing mechanisms to inverted faces. However, a limitation of these earlier studies was the fact that occipitotemporal visual regions were classified as either "face" or "object" areas based on differential response to either faces versus nonface objects (Aguirre et al., 1999) or faces versus houses (Haxby et al., 1999). Thus, these studies did not distinguish between voxels involved in generic processing of objects from voxels that might be preferentially involved in house/scene processing. Furthermore, most of these earlier studies focused on particular regions of interest in occipitotemporal cortex and thus might have overlooked regions outside of these ROIs that responded differentially to upright versus inverted faces. An exception is a study by Leube et al. (2003), which examined effects of face inversion across the whole brain and found greater response to upright than to inverted faces in the right superior temporal sulcus and right insular cortex.

The neural consequences of scene inversion have been comparatively less studied, probably because the behavioral effects of scene inversion have not been reliably established. Haxby et al. (1999) used upright and inverted houses as the nonface

comparison stimuli in an fMRI study of the face inversion effect. Although house inversion per se was not a focus of this study, the authors reported reduced response to inverted compared to upright houses in "house-selective" voxels in the medial fusiform part of ventral temporal cortex, but a non-significant trend towards greater response for inverted houses in more posterior occipital regions. Insofar as the medial temporal "house-selective" voxels about the scene-selective parahippocampal place area (PPA; Epstein & Kanwisher, 1998), these results suggest that scene inversion may result in decreased engagement of scene-selective regions but increased engagement of generic object processing regions. In other words, scene inversion and face inversion may have qualitatively similar effects on their respective processing streams. However, in the absence of a nonhouse object comparison condition, it is unclear whether the effects of house inversion in this study are specific to houses, or whether they would apply to a wide variety of stimuli.

In the current study, subjects were scanned while they performed a continuous visual matching task on upright and inverted versions of faces, "scenes" consisting of images of houses set in their natural surroundings, and nonface objects. Our primary goal was to compare the effects of face inversion to the effects of scene inversion in order to gain new insight into the neural origins of the face inversion effect. We were particularly interested in four issues. First, would face inversion lead to greater engagement of regions involved in generic object processing, as suggested by earlier studies? Second, to what extent would scene inversion lead to a behavioral penalty, and would that penalty be reflected in a similar shift towards greater engagement of generic object-processing regions? Third, are changes in the fMRI response engendered by face and scene inversion closely tied to the corresponding behavioral effects? Fourth, would inversion of scenes and faces have effects on regions outside of the occipitotemporal visual areas that were the focus of earlier studies? To anticipate, we found that face inversion does lead to greater engagement of generic object processing regions, scene inversion has a qualitatively similar though less dramatic effect, the magnitude of these neural effects are not tightly coupled to the magnitude of the behavioral inversion effects, and both face and scene inversion have large effects on many cortical regions that may be involved in post-recognition processing.

## 2. Methods

### 2.1. Subjects

Twelve healthy right-handed volunteers (six males, median age 20.5 years) were recruited from the local community and gave informed consent according to procedures approved by the University of Pennsylvania institutional review board. All subjects had normal or corrected-to-normal vision.

### 2.2. fMRI parameters

Subjects were scanned at the Hospital of the University of Pennsylvania on a 3 T Siemens Trio scanner equipped with a Siemens body coil and a four-channel head coil. T2\* weighted images sensitive to blood oxygenation level-dependent contrasts were acquired using a gradient-echo echoplanar pulse sequence (TR = 3000 ms, TE = 30 ms, matrix size = 64 × 64, field

of view = 19.2 cm). Each functional volume consisted of 44 3-mm thick axial slices. Prospective motion correction was done online with a PACE sequence. Structural T1\* weighted images for anatomical localization were acquired using a 3D MPAGE pulse sequence (TR = 1620 ms, TE = 3 ms, TI = 950 ms, voxel size = 0.9766 mm × 0.9766 mm × 1 mm, matrix size = 192 × 256 × 160).

### 2.3. Procedure

Scan sessions consisted of four experimental scans followed by two functional localizer scans. In the experimental scans subjects viewed upright and inverted digitized greyscale photographs of faces, buildings, and common everyday objects, all of which were unfamiliar to the subjects prior to the scan. Face images were obtained from the Psychological Image Collection at the Stirling University ([pics.psych.stir.ac.uk](http://pics.psych.stir.ac.uk)). Building photographs included background elements and thus depicted complete environmental “scenes” (see Fig. 1 for examples). In total subjects viewed 72 faces, 72 scenes, and 72 objects within the main experiment, each of which was presented once upright and once inverted. Note that the objects were chosen from a variety of different categories, so within-set similarity was lower for objects than for faces or scenes.

Each experimental scan was 432 s long and was divided into eighteen 18 s picture epochs (three epochs for each of the stimulus categories) interleaved with seven epochs in which the screen was blank except for a fixation point. During each picture epoch subjects viewed six different photographs of the same type for 1 s each with a 1 s interstimulus interval; three of these photographs were presented twice in sequence for a total of nine picture presentations in all. Subjects responded to each photograph by using a button box to indicate whether it was the same as or different from the immediately preceding photograph. Stimuli for the upright and inverted versions of each category were repeated across scans; for example, faces that were presented upright in scans 1 and 3 were presented inverted in scans 2 and 4 (and vice versa).

In the two localizer scans, subjects viewed digitized color photographs of faces, common objects, scenes, and other stimulus categories all presented in the upright orientation. Scans were divided into 15 s picture epochs interleaved with fixation epochs in a blocked design. During each picture epoch, 20 photographs of the same type were presented for 400 ms each with a 350 ms interstimulus interval. Subjects performed a one-back task in which they were required to press a button whenever two identical stimuli appeared in a row. There were two such repetitions in each epoch.

Stimuli were rear projected onto a Mylar screen at the head of the scanner with an Epson 8100 3-LCD projector equipped with a Buhl long-throw lens. A

mirror was mounted to the coil to allow subjects to view the screen. The images subtended a visual angle of  $8.75^\circ \times 8.75^\circ$ . EPrime (Psychology Software Tools Inc.) was used to run the experiments.

### 2.4. Data analysis

Functional images were corrected for differences in slice timing by resampling slices in time to match the first slice of each volume, and then realigned with respect to the first image of the scan using sinc interpolation. The mean realigned image was normalized to the Montreal Neurological Institute (MNI) template using a set of affine and smoothly nonlinear transformations as implemented in SPM2. This transformation was then separately applied to all functional images for the subject. Images were resampled into 3 mm isotropic voxels and spatially smoothed with an 8 mm FWHM gaussian filter.

Data were analyzed using the general linear model as implemented in VoxBo ([www.voxbo.org](http://www.voxbo.org)), including an empirically derived  $1/f$  noise model, filters that removed high and low temporal frequencies, regressors to account for global signal variations, and nuisance regressors to account for between-scan differences (Aguirre, Zarahn, & D’Esposito, 1997; Zarahn, Aguirre, & D’Esposito, 1997). Each stimulus condition was modeled as a boxcar function convolved with a canonical hemodynamic response function. Both regions of interest and whole-brain analyses were performed.

In the region of interest analysis, individual subject ROIs were defined by using linear contrasts to identify sets of contiguous voxels that responded preferentially to faces, scenes, and objects in the localizer scans as predicted by previous experiments. Specifically: (1) the fusiform face area was identified by a face > object contrast in the left hemisphere in 10 subjects and the right hemisphere in all 12 subjects, (2) the superior temporal sulcus face area was identified by a face > object contrast in the right hemisphere in 9 subjects, (3) the parahippocampal place area was identified by a scene > object contrast in both the left and right hemispheres in all 12 subjects, (4) the middle fusiform “object” area was identified by an object > (faces and scenes) contrast in the left hemisphere in 9 subjects and in the right hemisphere in 7 subjects, (5) the posterior portion of the lateral occipital complex was identified by an object > scenes contrast in the left hemisphere in 12 subjects and in the right hemisphere in 11 subjects. The timecourse of MR response during the experimental scans was extracted from each of these regions of interest (averaging over all voxels) and entered into the general linear model in order to calculate parameter estimates (beta values) for all six conditions, which were used as the dependent variables in a random effects analysis. To ensure that results from the middle fusiform

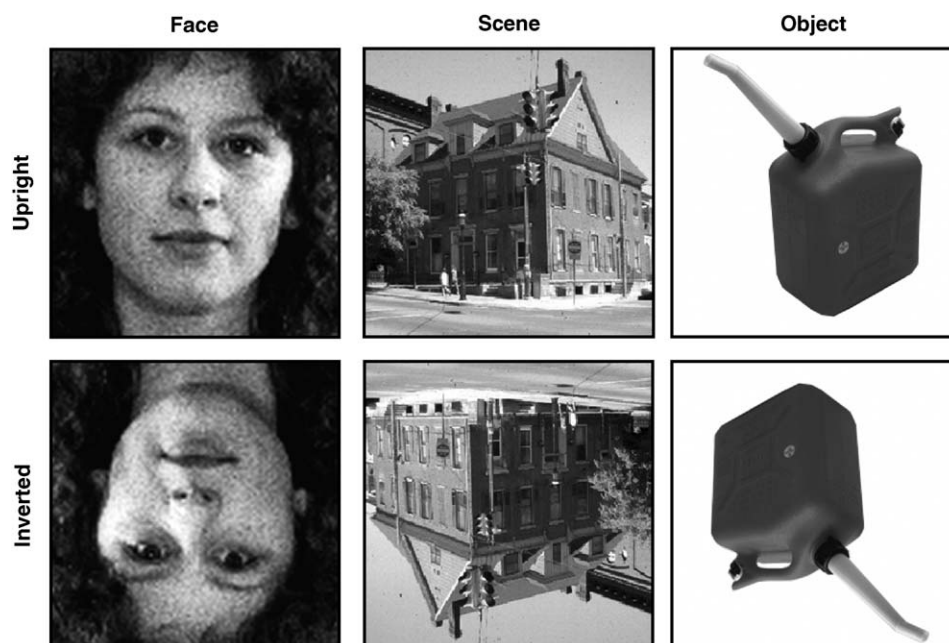


Fig. 1. Examples of stimuli used in the experiment.

object area were not biased by the fact that this region was not defined in several subjects, we performed an additional analysis in which MF was defined for all subjects using data from a group analysis of the localizer data. The results were not qualitatively different from those reported below for the individually defined ROIs.

In the whole-brain analyses, subject-specific *t*-maps were calculated for contrasts of interest and then smoothed to 12 mm FWHM to facilitate between-subject averaging before entry into a random effects analysis. Worsley's *stat\_threshold* program was used to identify clusters with a corrected significance level of  $p < 0.05$  (Cao, 1999; Friston, Worsley, Frackowiak, Mazziotta, & Evans, 1994). At a voxelwise uncorrected significance level of  $p < 0.001$ , this corresponded to a cluster size of  $>53$  voxels. A similar analysis using subject-specific beta maps gave almost identical results to those reported here.

### 3. Results

#### 3.1. Behavioral results

Reaction times to correct trials were examined (Table 1). An omnibus analysis of variance revealed significant main effects of stimulus type [ $F_{(2,22)} = 12.4, p < 0.001$ ] stimulus orientation [ $F_{(1,11)} = 40.6, p < 0.001$ ] and run [ $F_{(3,33)} = 18.8, p < 0.001$ ]. Critically, we also observed a significant stimulus type  $\times$  stimulus orientation interaction [ $F_{(2,22)} = 4.8, p < 0.02$ ], reflecting the fact that increases in reaction time after both face and scene inversion were significantly greater than the increases in RT after object inversion [(face versus object)  $\times$  (upright versus inverted):  $F_{(1,11)} = 6.3, p < 0.05$ ; (scene versus object)  $\times$  (upright versus inverted):  $F_{(1,11)} = 10.1, p < 0.01$ ]. Interestingly, there was a significant stimulus type  $\times$  inversion  $\times$  run interaction [ $F_{(6,66)} = 3.0, p < 0.02$ ] attributable to the fact that the scene inversion effect was stronger for the first scan than for subsequent scans [(scene versus object)  $\times$  (upright versus inverted)  $\times$  run:  $F_{(3,33)} = 3.8, p < 0.05$ ] while the face inversion effect was more constant over the course of the experiment [(face versus object)  $\times$  (upright versus inverted)  $\times$  run:  $F_{(3,33)} = 1.3, n.s.$ ]. Thus, significant inversion effects were observed for both faces and scenes, but the scene inversion effect was less robust insofar as subjects were able to partially compensate for scene inversion over the course of the experiment. Subjects made few errors (4.4%) and the number of errors did not differ significantly across the six experimental conditions.

Table 1  
Reaction time for each condition in ms

Stimulus	Run in experiment				Mean
	Run 1	Run 2	Run 3	Run 4	
Faces					
Upright	586	552	527	519	546 $\pm$ 24
Inverted	629	586	567	534	579 $\pm$ 23
I-U	43	34	40	15	33 $\pm$ 9
Scenes					
Upright	574	542	545	512	543 $\pm$ 26
Inverted	660	557	553	521	573 $\pm$ 24
I-U	86	15	8	9	30 $\pm$ 5
Objects					
Upright	582	539	519	499	535 $\pm$ 27
Inverted	582	538	527	518	541 $\pm$ 26
I-U	0	-1	8	19	6 $\pm$ 5

#### 3.2. fMRI inversion effects: functional regions of interest

We initially focused our analysis on several functionally defined regions of interest in occipitotemporal cortex that have been established by previous studies. In particular, we used data from the localizer scans to identify three cortical regions that respond preferentially to specific stimulus categories: the fusiform face area (Kanwisher et al., 1997), the superior temporal sulcus face area (STS: Hoffman & Haxby, 2000; Puce, Allison, Bentin, Gore, & McCarthy, 1998), and the parahippocampal place area (Aguirre et al., 1998; Epstein & Kanwisher, 1998). In addition, we identified two regions that respond to objects in a nonspecific manner: the posterior portion of the lateral occipital complex (LO: Grill-Spector, Kourtzi, & Kanwisher, 2001; Malach et al., 1995) and the middle fusiform object region (MF: Grill-Spector, 2003; Hasson, Harel, Levy, & Malach, 2003). The latter region is a swath of territory between the PPA and the FFA on the fusiform gyrus that responds at least as strongly to nonface and nonscene objects as it does to faces and scenes. The locations of these regions are illustrated in Fig. 2.

Consistent with earlier results, we observed face inversion effects in the form of *increased* response to inverted versus

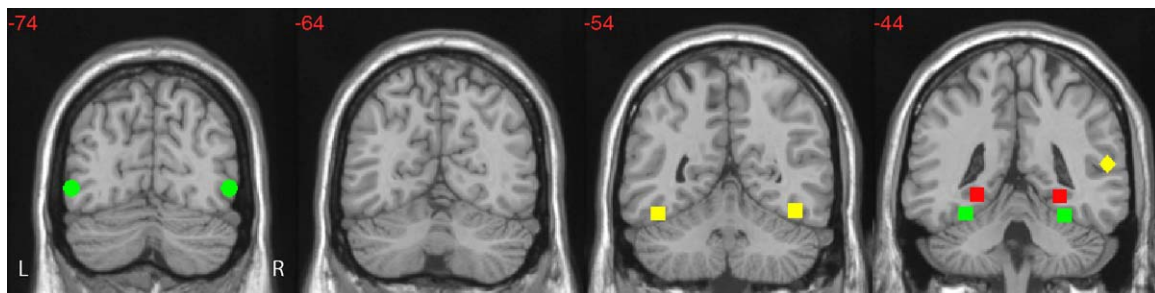


Fig. 2. Functionally defined regions of interest. The cross-subject average locations of the parahippocampal place area (red squares), middle fusiform object area (green squares), fusiform face area (yellow squares), superior temporal sulcus face area (yellow diamond) and posterior lateral occipital object area (green circles) are plotted on coronal slices of a reference brain in standard space. Right hemisphere is on the right. Note the medial to lateral organization of the PPA, MF, and FFA within ventral occipitotemporal cortex. Mean stereotactic coordinates in MNI space for each region were left FFA = (-44, -55, -23), right FFA = (42, -53, -21), right STS = (57, -48, 8), left PPA = (-25, -48, -11), right PPA = (27, -44, -12), left MF = (-32, -46, -23), right MF = (30, -44, -24), left LOC = (-50, -72, -7), right LOC = (49, -72, -7). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

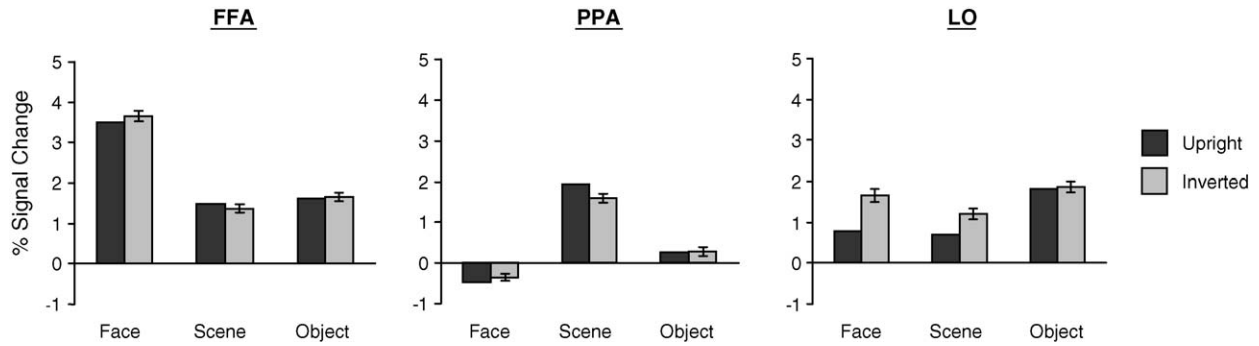


Fig. 3. MR response within the fusiform face area (FFA), parahippocampal place area (PPA), and lateral occipital (LO) object area. Data are averaged over left and right hemisphere. Error bars represent  $\pm 1$ S.E.M. for the upright-inverted difference for each stimulus category. The FFA does not distinguish between upright vs. inverted faces, but the PPA shows reduced response to inverted vs. upright scenes. LO responds more strongly to inverted than to upright stimuli for both faces and scenes. A more complete description of these data can be found in Table 2.

upright faces in generic object processing regions [left LO  $t(11) = 4.9, p < 0.001$ ; right LO  $t(10) = 6.6, p < 0.0001$ ; right MF  $t(6) = 4.1, p < 0.01$ ] (see Fig. 3 and Table 2). We tested for face-specific inversion effects in each region (i.e. face inversion effects that were significantly larger than object inversion effects) by performing analysis of variance with two factors: stimulus type (face versus object) and stimulus orientation (upright versus inverted). The interaction between these factors was significant in left and right LO [left  $F_{(1,11)} = 31.3, p < 0.001$ ; right  $F_{(1,10)} = 37.5, p < 0.001$ ] and in right MF [ $F_{(1,6)} = 18.9, p < 0.01$ ]. Face-specific inversion effects were not observed in the left or right FFA, PPA, left MF, or right STS [all  $p$ 's  $> 0.15$ ]. Indeed, it is interesting to note that FFA responses to upright and inverted faces did not differ [paired  $t$ -test:  $p > 0.15$  in both hemispheres], a pattern that has been observed in some but not all earlier studies.

Increases in fMRI activity in generic object processing regions were also observed after scene inversion [left  $t(11) = 3.6, p < 0.01$ ; right  $t(10) = 9.2, p < 0.0001$ ]. In this case, this increase in LO response was accompanied by a decrease in response in the appropriate category-specific region, the PPA [left  $t(12) = 2.7,$

$p < 0.05$ ; right  $t(12) = 3.0, p < 0.05$ ]. Examination of scene-specific orientation effects revealed significant or nearly significant stimulus type (scene versus object)  $\times$  stimulus orientation (upright versus inverted) interactions in LO [left  $F_{(1,11)} = 4.7, p = 0.51$ ; right:  $F_{(1,10)} = 7.4, p < 0.05$ ] and the PPA [left  $F_{(1,11)} = 5.7, p < 0.05$ ; right  $F_{(1,11)} = 5.2, p < 0.05$ ]. An interesting aspect of the data from the PPA is that response reduction after stimulus inversion is relatively mild compared to the place  $>$  object category difference that was used to define this region. In particular, the response to inverted scenes in the PPA, albeit smaller than the response to upright scenes, is still much greater than the response to objects [left PPA  $t(11) = 8.5$ , right PPA  $t(11) = 8.3$ , both  $p$ 's  $< 0.00001$ ]. Scene-specific inversion effects were not observed in the FFA, STS, or MF (all  $p$ 's  $> 0.25$ ).

### 3.3. Correlations between neural and behavioral inversion effects

The results above demonstrate that both inverted faces and inverted scenes more strongly engage the lateral occipital than their upright counterparts. In addition, increased response was

Table 2  
MR response within functionally defined regions of interest

	Upright faces	Inverted faces	Upright scenes	Inverted scenes	Upright objects	Inverted objects	Face-specific inversion effect	Scene-specific inversion effect
Left FFA	2.96	2.97	1.04	0.99	1.39	1.32	0.09	0.02
Right FFA	3.90	4.11	1.78	1.66	1.87	1.93	0.16	-0.18
Right STS	0.99	0.72	-0.02	-0.23	0.19	-0.01	-0.08	-0.01
Left PPA	-0.55	-0.54	1.63	1.34	0.15	0.14	0.01	-0.29*
Right PPA	-0.41	-0.16	2.21	1.85	0.34	0.43	0.15	-0.45*
Left LOC	0.44	1.34	0.70	1.30	1.99	1.96	0.93***	0.64†
Right LOC	1.39	2.15	0.77	1.26	1.59	1.67	0.68***	0.41*
Left MF	0.40	0.67	1.48	1.56	1.89	2.12	0.04	-0.15
Right MF	1.12	1.83	2.78	2.72	2.85	2.95	0.61**	-0.16

Units are percent signal change relative to fixation baseline. The face-specific inversion effect is defined as (inverted faces – upright faces) – (inverted objects – upright objects) and the scene-specific inversion effect is defined as (inverted scenes – upright scenes) – (inverted objects – upright objects).

\*  $p < 0.05$ .  
 \*\*  $p < 0.01$ .  
 \*\*\*  $p < 0.001$ .  
 †  $p = 0.051$ .

observed in the right middle fusiform object area after face inversion. Can this shift in response emphasis towards generic object processing regions explain the corresponding behavior inversion effects? We tested this idea by examining the correlation across subjects between the strength of the regional fMRI face and scene inversion effects and the corresponding behavioral inversion effects. For this analysis, inversion effects were defined simply as the difference in beta values or reaction time between upright and inverted stimuli. We did not find any significant correlations in any of our regions of interest (all  $p$ 's > 0.15). Nor did we find a correlation between behavioral inversion effects and the *difference* in the strength of the fMRI inversion effects between the PPA and LO or between the FFA and LO, as one might expect to find if the strength of the behavioral inversion effects were directly linked to the size of a shift in processing from the FFA/PPA towards LO after inversion ( $p$ 's > 0.15). These results were substantially the same when behavioral inversion effects were measured in terms of proportional rather than absolute differences to compensate for the typically right-skewed nature of the RT distribution.

Although it is difficult to make definitive conclusions from negative results, the absence of correlation between the fMRI and behavioral inversion effects in the FFA, PPA, MF, and LO suggests that activity in other cortical regions might critically mediate individual differences in the strength of the behavioral effects. To test this, we measured the correlation between the behavioral and fMRI inversion effects for faces and houses across all voxels of the brain. For face inversion, no regions survived the cluster-size threshold for significance. For the scene inversion, we observed a significant correlation between fMRI and behavioral effects in the left medial frontal gyrus, near the boundary of Brodmann's areas 10 and 32, with an extension into more anterior BA10 (see Fig. 4). Specifically, subjects for whom the upright scene-inverted scene fMRI response difference in this region was large and positive showed large behavioral inversion penalties, while subjects for whom the upright scene-inverted scene fMRI response difference was small or negative showed

smaller behavioral inversion penalties. In particular, there were four subjects who showed greater response to inverted than to upright scenes; these were the subjects with the smallest behavioral inversion effect. It is worth noting that the strong correlation between the behavioral and fMRI inversion effect in this medial frontal region ( $r = 0.89$ ,  $p < 0.0001$ ) cannot be completely attributed the presence of these four outlier subjects, as there was still a trend towards a significant fMRI-RT correlation when these subjects were excluded ( $r = 0.67$ ,  $p = 0.07$ ).

The complete pattern of results from this medial frontal region suggests that it may not play a role in scene processing per se; rather, it may mediate some general cognitive process whose engagement or disengagement reflects (perhaps indirectly) the use of different strategies that affect performance on the matching task. In particular, despite the strong correlation between the behavioral and fMRI scene inversion effects, the mean size of the upright scene-inverted scene difference was not significantly different from zero ( $t < 1$ , n.s.). In contrast, there was a significant mean effect of face inversion ( $t(11) = 6.2$ ,  $p < 0.0001$ ), but no significant correlation between the strength of the face inversion effect and the difference in neural response between upright and inverted faces ( $r = -0.47$ ,  $p = 0.13$ , n.s.). Furthermore, the response to all six stimulus conditions was significantly *lower* than the response during the fixation baseline (all  $t$ 's > 3.9, all  $p$ 's < 0.002), suggesting that this region supports processes that are disengaged by the matching task, rather than engaged.

A number of other researchers have observed similar "deactivations" in medial frontal cortex. It has been argued that this region may support a self-referential default processing mode that tends to turn off during attention-demanding cognitive tasks. Thus, we might interpret response differences in this region as reflecting the amount of attention or effort that subjects direct towards different stimulus categories in the matching task, with lower response (i.e. greater deactivation) corresponding to more effort. In this view, individual differences in the medial frontal response to a particular stimulus category might reflect the adoption of various strategies or processing

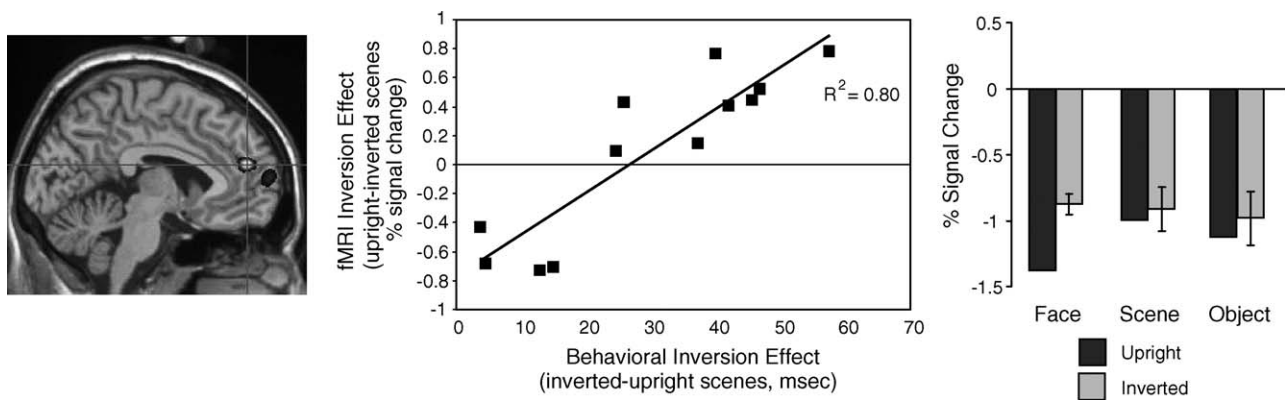


Fig. 4. *Left panel*: Results of a whole-brain group analysis showing a region in medial frontal cortex (BA10) where behavioral and fMRI scene inversion effects were significantly correlated. *Middle panel*: Plot of fMRI scene inversion effect (percent MR signal change to upright scenes-percent MR signal change to inverted scenes) vs. behavioral scene inversion effect (reaction time to inverted scenes-reaction time to upright scenes) in this region. Note the strong correlation, driven in part by the four subjects with the smallest behavioral inversion effect who also showed greater response to inverted scenes than to upright scenes. *Right panel*: Mean fMRI response to all six stimulus conditions in this region. Error bars represent  $\pm 1$  S.E.M. for the upright-inverted difference for each stimulus category. Units on y-axis are percent signal change.

modes that differ in the degree to which they tap attentional resources. In particular, the correlation between the behavioral and fMRI scene inversion effect (which appears to be largely driven by variation in the response to inverted scenes; data not shown) might be attributable to the adoption by some subjects of more efficient recognition strategies that lead to both faster reaction times and less deactivation in the medial frontal lobe for inverted scenes. In contrast, the reliability and stability of the behavioral face inversion effect suggest that the strategies involved in processing upright and inverted faces are more fixed across subjects. The less efficient strategies for inverted faces are reflected in both longer RTs and lower medial frontal response when compared to upright faces; however, we see no evidence that these quantities vary across subjects in a systematic manner.

#### 3.4. fMRI inversion effects: whole-brain analyses

A whole-brain analysis was performed to identify regions outside of the occipitotemporal face/scene/object areas that responded differentially to upright versus inverted faces and upright versus inverted scenes. Several brain regions sensitive to faces and scene inversion were identified (see Table 3, and Figs. 5 and 6). We then tested the corresponding stimulus type (face versus object or scene versus object) by stimulus orientation (upright versus inverted) interaction within each region to verify that the inversion effects were specific to faces or scenes. Regions in which these interactions were signifi-

cant ( $p < 0.05$ ) are described below. We first consider regions affected by face inversion and then the regions affected by scene inversion.

Greater response to inverted compared to upright faces was observed in both the lateral occipital area and also in a large bilateral swath of parietal cortex, roughly following the intraparietal sulcus. Thus, the whole-brain analysis confirms the data from the ROI analysis indicating that LO responds more strongly to inverted than to upright faces. The finding of greater response to inverted faces in parietal cortex was unexpected. When the responses to all six stimulus conditions within this parietal region were examined, it was found that the right parietal region also responded more strongly to inverted scenes than to upright scenes ( $t(11) = 2.7$ ,  $p < 0.02$ ), although this difference was not large enough to observe at corrected significance levels in the whole-brain analysis (see Fig. 7). Furthermore, the response to upright faces was significantly lower than the response to all other stimulus categories in both the left and right parietal region (all  $p$ 's  $< 0.02$ ) and was not significantly different from the fixation baseline ( $t < 1$ , n.s. in both hemispheres). Insofar as parietal cortex has been strongly linked to spatial attention (Kastner, Pinsk, De Weerd, Desimone, & Ungerleider, 1999; Wojciulik & Kanwisher, 1998), the fact that it engages to non-face objects, scenes, and inverted faces suggests that recognition of these stimuli might involve the use of attention to focus on individual object parts or scene elements. In contrast, upright faces are matched on the basis of holistic representations (Farah et al., 1998) and thus may not require the recruitment of spatial

Table 3  
Regions responding differentially to upright vs. inverted faces and upright vs. inverted scenes in a whole-brain group analysis

Contrast and region	MNI coordinates			Max $t$ -value	Size in voxels
	$x$	$y$	$z$		
<b>Upright &gt; inverted faces</b>					
R. Amygdala	23	1	-23	5.55	91
Orbital frontal	0	52	-7	7.73	273
Medial frontal	2	60	20	10.98	327
<i>R anterior middle temporal gyrus</i>	<i>60</i>	<i>-3</i>	<i>-24</i>	<i>6.12</i>	<i>53</i>
<i>L anterior middle temporal gyrus</i>	<i>-60</i>	<i>-18</i>	<i>-21</i>	<i>5.54</i>	<i>82</i>
<i>R anterior superior temporal gyrus</i>	<i>57</i>	<i>-24</i>	<i>3</i>	<i>6.39</i>	<i>136</i>
<i>R superior temporal sulcus</i>	<i>66</i>	<i>48</i>	<i>9</i>	<i>5.52</i>	<i>176</i>
<i>Posterior cingulate</i>	<i>-3</i>	<i>54</i>	<i>27</i>	<i>7.11</i>	<i>193</i>
<b>Inverted &gt; upright faces</b>					
L lateral occipital	-40	-70	-2	-10.32	358
R lateral occipital	40	-78	8	-8.42	545
L posterior parietal	-36	-42	52	-7.4	224
R posterior parietal	33	-62	53	-8.36	569
<b>Upright &gt; inverted scenes</b>					
R hippocampus	34	-16	-17	6.08	63
L parahippocampal cortex	-24	-33	-12	5.58	70
Retrosplenial cortex	-13	-51	13	5.8	163
L superior temporal sulcus	-40	-60	28	6.84	87
R superior occipital	19	-94	27	10.96	222
<b>Inverted &gt; upright scenes</b>					
L lateral occipital	-42	-76	-3	-5.68	185
<i>Cerebellum</i>	<i>-3</i>	<i>-78</i>	<i>-15</i>	<i>-9.81</i>	<i>412</i>

Regions exhibiting either a face-specific inversion effect (larger upright vs. inverted difference for faces than for objects) or a scene-specific inversion effect (larger upright vs. inverted difference for scenes than for objects) in roman type; additional regions in italics.

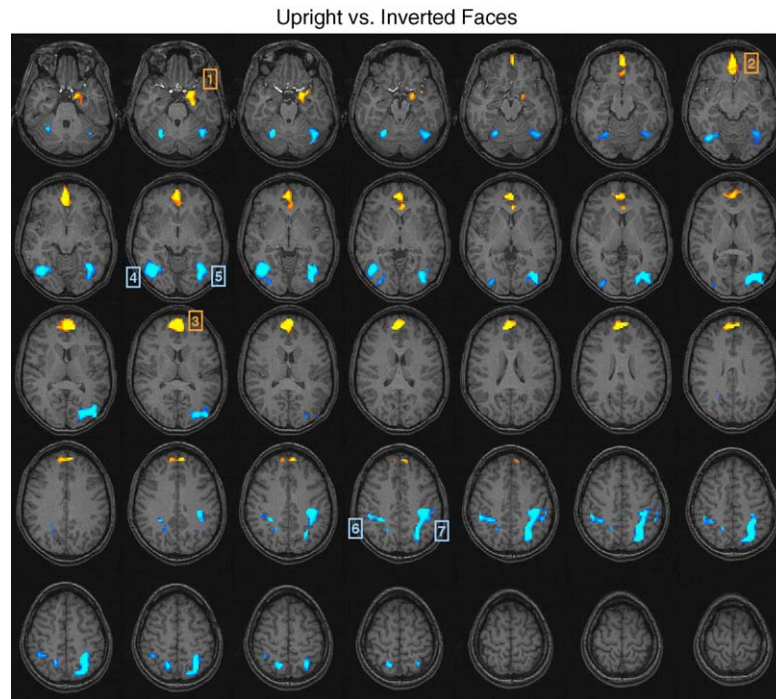


Fig. 5. Results of a whole-brain group analysis comparing response to upright vs. inverted faces. Regions responding significantly more to upright vs. inverted faces (orange) or inverted vs. upright faces (blue) at a  $p < 0.05$  corrected significance level are plotted on axial slices from a single subject's structural scan in standard space. Right hemisphere is on the right. Numbers refer to the following clusters: (1) right amygdala, (2) orbital frontal cortex, (3) medial frontal cortex, (4) right lateral occipital, (5) left lateral occipital, (6) left posterior parietal, (7) right posterior parietal. Only regions that exhibited a face-specific inversion effect (i.e. significant upright vs. inverted  $\times$  face vs. object interaction) are plotted; addition regions showing differential response to upright vs. inverted faces are listed in Table 2. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

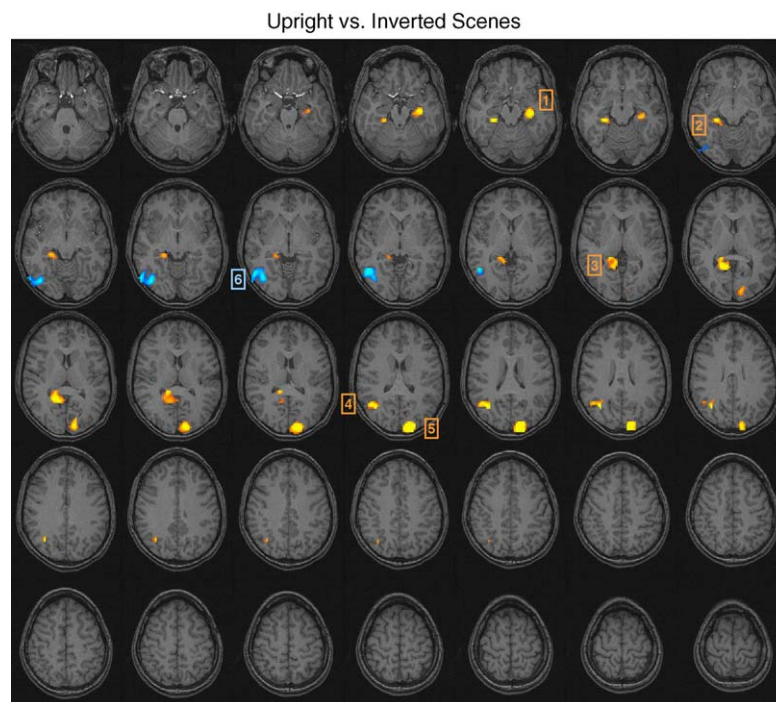


Fig. 6. Results of a whole-brain group analysis comparing response to upright vs. inverted scenes. Regions responding significantly more to upright vs. inverted scenes (orange) or inverted vs. upright scenes (blue) at a  $p < 0.05$  corrected significance level are plotted on axial slices from a single subject's structural scan in standard space. Right hemisphere is on the right. Numbers refer to the following regions: (1) right hippocampus, (2) left hippocampus/parahippocampal cortex, (3) retrosplenial cortex, (4) left superior temporal sulcus (5) right superior occipital, (6) left lateral occipital. Only regions that exhibited a scene-specific inversion effect (i.e. significant upright vs. inverted  $\times$  scene vs. object interaction) are plotted; addition regions showing differential response to upright vs. inverted scenes are listed in Table 2. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)



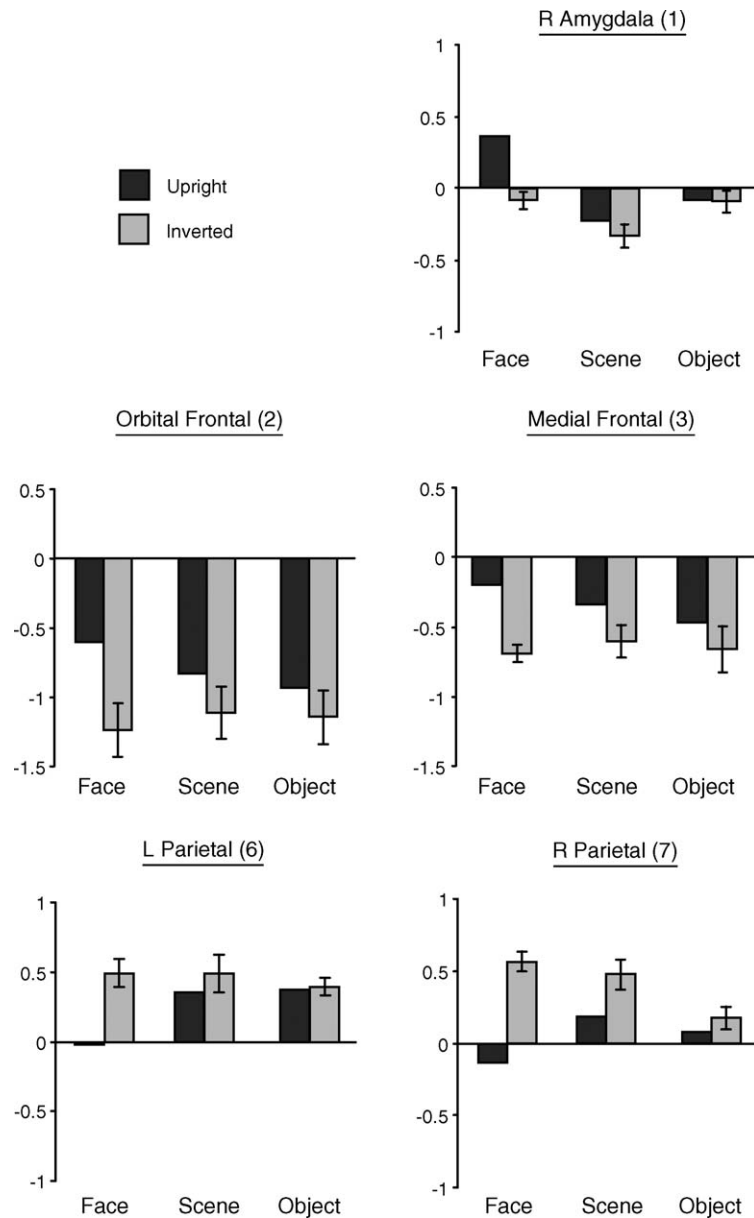


Fig. 7. Response within five regions that responded more strongly to upright vs. inverted faces (amygdala, orbital frontal cortex, or medial frontal cortex) or more strongly to inverted vs. upright faces in the whole-brain analysis. Numbers refer to labeled activations in Fig. 5. Error bars represent  $\pm 1$  S.E.M. for the upright-inverted difference for each stimulus category. Units on y-axis are percent signal change. Note that response to inverted faces in the amygdala is no greater than response to nonface objects.

attention towards different face parts during performance of the matching task.

The reverse comparison revealed three loci that responded more strongly to upright than to inverted faces: the amygdala, medial frontal cortex (BA10), and orbital frontal cortex (BA10/11). These regions are notable insofar as they have been previously implicated in the processing of social and emotional information, conveyed by both faces and nonface stimuli (see Ochsner, 2004 for review). For example, the response of the amygdala to faces is modulated by their emotional valence (Breiter et al., 1996; Morris et al., 1996; Whalen et al., 1998) and its response to names is modulated by the “goodness” or “badness” of the named person (Cunningham, Johnson, Gatenby,

Gore, & Banaji, 2003). Similarly, the medial orbitalfrontal cortex engages when judging the attractiveness of a face (O’Doherty et al., 2003) and it also responds more strongly when reading stories or viewing cartoons that include a theory-of-mind component than when reading or viewing nontheory of mind control stimuli (Fletcher et al., 1995; Gallagher et al., 2000; see also Schultz et al., 2003 for similar results in the amygdala). Thus, the greater response in these regions for upright compared to inverted faces may reflect the operation of social processing mechanisms that engage to upright faces but not to inverted faces. Note that both of the frontal lobe regions identified here responded more strongly during fixation baseline than during the matching task, consistent with earlier studies suggesting that the

baseline state, insofar as it is less attention demanding, allows for more self-monitoring (Gusnard, Akbudak, Shulman, & Raichle, 2001) or more free rein for unconstrained socially- or emotionally relevant thoughts (Christoff, Ream, & Gabrieli, 2004).

A qualitatively similar pattern of results was observed for scenes. Greater response to inverted than to upright scenes was observed in the left lateral occipital complex, while greater response to upright than to inverted scenes was observed in the right hippocampus, left parahippocampal cortex/posterior hippocampus (slightly anterior to the PPA), retrosplenial cortex, left superior temporal sulcus, and a posterior occipital region whose coordinates are consonant with visual area V3a (Figs. 6 and 8). Previous studies have implicated several of these regions in navigational processing. For example, the right hippocampus has been shown to be active when subjects successfully navigate through a complex virtual-reality environment (Hartley, Maguire, Spiers, & Burgess, 2003; Maguire et al., 1998), while retrosplenial cortex is activated during mental recapitulation of routes through familiar environments (Ino et al., 2002; Rosenbaum, Ziegler, Wincour, Grady, & Moscovitch, 2004), retrieval of spatial information (Burgess, Maguire, Spiers, & O'Keefe, 2001), and passive viewing of environmental scenes

(O'Craven & Kanwisher, 2000). In contrast, the greater response to upright than to inverted scenes in V3a might relate to physical differences between the stimuli such as the fact that upright scenes contain a prominent ground plane in the lower visual field while inverted scenes do not. The de-emphasis of the ground plane after scene inversion might lead to reduced depth perception and consequent reduced engagement of V3a, which is known to respond strongly to depth-related cues (Backus, Fleet, Parker, & Heeger, 2001).

In sum, the whole-brain analysis revealed qualitatively similar effects of inversion for both faces and scenes. For both stimulus classes, inverted stimuli activated the lateral occipital complex more strongly than upright stimuli. A similar pattern was observed in posterior parietal cortex, although here the upright versus inverted difference only passed the corrected significance level for faces. Furthermore, for both stimulus classes, upright stimuli more strongly activated regions that have been previously implicated in post-recognition processes that can only be effectively performed on the upright stimulus.

Interestingly, the response to the inverted version of the preferred stimulus was no greater than the response to non-preferred objects in many of these putative post-recognition

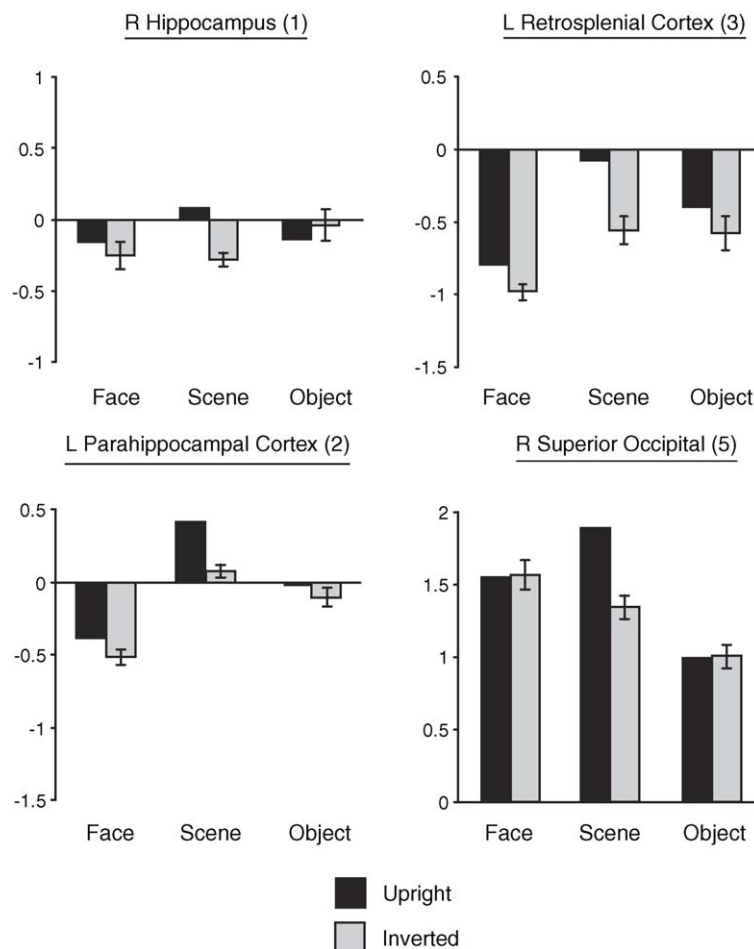


Fig. 8. Response within four regions that responded more strongly to upright vs. inverted scenes (hippocampus, parahippocampal cortex, retrosplenial cortex, superior occipital gyrus). Numbers refer to labeled activations in Fig. 6. Error bars represent  $\pm 1$  S.E.M. for the upright-inverted difference for each stimulus category. Units on y-axis are percent signal change. Note that response to inverted scenes in the hippocampus, parahippocampal cortex, and retrosplenial cortex is no greater than response to nonface objects.

regions. Figs. 7 and 8 illustrate this pattern in the amygdala, hippocampus, parahippocampal cortex, and retrosplenial cortex. It is possible that this result is a trivial consequence of the fact that these regions were defined on the basis of the upright > inverted contrast (and thus would be expected to include voxels with particularly low responses to the inverted stimulus). However, the same pattern was observed in the amygdala and retrosplenial cortex when these regions were defined based on different contrasts performed on the independently acquired localizer dataset. Specifically, within an amygdala ROI defined by a face > object contrast, the response to inverted faces was significantly reduced compared to the response to upright faces [left amygdala  $t(11) = 2.6$ ,  $p < 0.05$ ; right amygdala  $t(11) = 5.9$ ,  $p < 0.001$ ] but did not differ from the response to objects [both  $t$ 's < 1.1, n.s.]. Similarly, within a retrosplenial ROI defined by a scene > object contrast, the response to inverted scenes was significantly reduced compared to the response to upright scenes [left  $t(11) = 4.2$ ,  $p < 0.01$ ; right  $t(11) = 3.6$ ,  $p < 0.05$ ] but did not differ from the response to objects [both  $t$ 's < 1, n.s.]. Thus, these regions do not appear to differentiate between inverted versions of the preferred stimulus and nonpreferred objects. As such, they make an interesting contrast to regions such as the PPA and FFA in which strong category specificity is preserved after inversion.

#### 4. Discussion

We examined the behavioral and neural consequences of face and scene inversion with the aim of identifying similarities and differences between the effects induced by these two manipulations. In particular, we addressed four questions. First, does face inversion lead to greater engagement of regions involved in generic object processing? Second, does scene inversion lead to a behavioral penalty and a similar shift towards greater engagement of generic object-processing regions? Third, can we link changes in the fMRI response engendered by face and scene inversion to the corresponding behavioral effects? Fourth, to what extent do face and scene inversion affect processing in regions outside of the occipitotemporal territory that was the focus of earlier studies? We found that (1) generic object processing regions responded more strongly to inverted faces than to upright faces, (2) behavioral performance was significantly impaired after scene inversion – but not as reliably as after face inversion – and generic object processing regions responded more strongly to inverted scenes than to upright scenes, (3) behavioral and fMRI inversion effects were not strongly correlated in occipitotemporal regions, but behavioral and fMRI scene inversion effects were correlated in medial frontal cortex, (4) several cortical regions outside of occipitotemporal cortex responded more strongly to upright faces than to inverted faces, while others responded more strongly to upright scenes than to inverted scenes. We now consider these four main results in turn.

Previous studies observed greater response in “object-processing” regions when subjects viewed inverted faces than when they viewed upright faces (Aguirre et al., 1999; Haxby et al., 1999). The current study confirmed and extended these results. In particular, the earlier studies classified occipitotemporal regions as either face-selective or object-selective depending

on whether they responded more strongly to faces or to the nonface control category (objects in one study, houses in the other). These binary classification schemes did not distinguish between voxels that respond preferentially to nonface/nonscene objects and voxels that respond preferentially to scenes. In the current study, we distinguished between face-responsive territory in the lateral fusiform gyrus and superior temporal sulcus (FFA and STS), scene-responsive territory in the collateral sulcus (PPA), and generic object-responsive territory in the middle fusiform gyrus and lateral occipital region (MF and LO). Significant face inversion-specific increases were found in the lateral occipital and right middle fusiform object regions. Although the right PPA and STS also responded more strongly to inverted than to upright faces, these effects were not specific to faces insofar as they were not significantly stronger than the similar increases seen after object inversion. Thus, our results suggest that face inversion leads to greater engagement of generic-object-recognition mechanisms in the lateral occipital complex (and perhaps also the right middle fusiform gyrus). In contrast, our results do not strongly support the hypothesis that scene and landmark recognition mechanisms in the PPA are involved in recognizing inverted faces.

This inversion-related shift towards greater response in the lateral occipital complex was not accompanied by reduction of response in the fusiform face area. Other studies of face inversion in the FFA have obtained mixed results: FFA response to inverted faces during performance of a matching task is either equal (Aguirre et al., 1999), or slightly but significantly smaller (Kanwisher et al., 1998; Haxby et al., 1999; Yovel & Kanwisher, 2004) than response to upright faces, and much greater than response to nonface objects (Aguirre et al., 1999; Haxby et al., 1999; Kanwisher et al., 1998; Yovel & Kanwisher, 2004). Furthermore, one study that directly compared the effects of face and house inversion in the FFA found no significant difference in the size of these effects (Haxby et al., 1999). If we assume that the FFA supports processing mechanisms that only operate efficiently on upright faces, then these results suggest that the FFA may respond automatically to faces even when its engagement is not essential for performance of the task at hand (Kanwisher et al., 1998). Results from at least three other recent studies support the claim that there can be decoupling between face-selective activity in the FFA and the successful operation of face-specific recognition mechanisms. First, Grill-Spector, Knouf, and Kanwisher (2004) found strong FFA response to briefly-presented and masked faces whose presence was detected but whose identity could not be ascertained. Second, Cox, Meyers, and Sinha (2004) observed strong FFA response when contextual cues such as the presence of a body on the screen implied the existence of a face but no actual face appeared in the stimulus. Third, Hasson, Avidan, Deouell, Bentin, and Malach (2003) reported face-selective activity in the FFA in a congenital prosopagnosic despite poor face recognition ability in this subject. In all three experiments, the FFA responds strongly in situations where face recognition does not occur.

We also observed a behavioral inversion penalty for scenes. Specifically, reaction times were longer when subjects made same/different judgments on inverted scenes than when they

made these judgments on upright scenes. This scene inversion effect was of comparable magnitude to the face inversion effect. In contrast, earlier studies found either no penalty for scene inversion (Wright & Roberts, 1996) or penalties that were significantly smaller than that observed for face inversion (Scapinello & Yarmey, 1970; Yin, 1969). Procedural differences between studies could possibly account for these discrepant results. In the earlier studies subjects inspected each scene for several seconds while in the current experiment subjects inspected each scene for only 1 s. Thus, subjects in the previous experiments might have had more time to separately encode the component features or objects within each scene, a strategy that might have led to a relatively small performance difference between upright and inverted scenes. Interestingly, in the current experiment the scene inversion effect was much stronger during the first experimental run than during subsequent runs—in contrast to the face inversion effect, which was of comparable size over all four runs. This pattern suggests that subjects were able to learn strategies for compensating for scene inversion over the course of the experiment. The face inversion effect, on the other hand, appears to be more robust. Scene inversion was accompanied by greater response in the lateral occipital object region – paralleling this regions' greater response to inverted compared to upright faces – and reduced response in the PPA.

Our results indicate that the neural effects of face and scene inversion are similar but not entirely parallel. On the neural level, both face and scene inversion lead to greater engagement of generic object-processing mechanisms in LO. Prior studies have suggested that the lateral occipital complex may be involved in the extraction of visual features prior to their combination into representations of whole objects, faces, or scenes in the ventral occipital temporal (VOT) region encompassing the FFA, MF, and PPA (Lerner, Hendler, Ben-Bashat, Harel, & Malach, 2001; Malach, Levy, & Hasson, 2002). Thus, this increased response in LO after face and scene inversion might reflect more intense processing of face parts and scene components when matching inverted stimuli. At the level of the VOT, scene inversion led to decreased response in the PPA while face inversion led to increased response in MF but no change in the FFA. These results are suggestive of continued engagement of face-processing mechanisms in the FFA to inverted faces, but reduced engagement of scene-processing mechanisms in the PPA to inverted scenes.

How do the differences in the neural effects of face and scene inversion illuminate the differences in the behavioral effects? In the case of scene inversion, we hypothesize that PPA scene representations might be optimized to support spatial orientation and navigation in the real world rather than matching scenes as in the present task. As such, the optimal strategy for matching inverted scenes might be to focus on low-level visual features processed by LO rather than on the whole-scene PPA representations that are sensitive to inversion. The fact that PPA response to scenes is reduced after inversion suggests that the visual system can learn to partially bypass these high-level representations to make match/nonmatch judgments. Once this strategy is adopted in the course of the experiment, subjects might be no more impaired on inverted scenes than on upright scenes

and the behavioral scene inversion penalty would be minimal, as we observe. In contrast, FFA representations are likely optimized for recognition tasks that are very similar to the matching task. Inversion would impair the ability to fully utilize these whole-face representations, leading to a decrement in performance that could not be easily eliminated. Indeed, the pattern of continued strong engagement of the FFA to inverted faces accompanied by increased engagement of LO and MF may indicate that inverted faces are processed in parallel by generic object-processing mechanisms in LO and face-specific mechanisms in the FFA. As such, the reaction time penalty for inverted faces might reflect the need for additional processing time to mediate between these two suboptimal representations.

Under this interpretation, increased response in LO is the primary neural correlates of the behavioral face and scene inversion effects. However, increased engagement of LO merely sets the stage—it leads to a robust behavioral inversion effect in the case of faces but not in the case of scenes. Behavioral inversion effects might depend as much on how face and scene representations are used flexibly or automatically by later processing stages as they do on initial encoding differences in occipitotemporal visual cortex. Indeed, we failed to find any correlation between the size of the behavioral and fMRI face and scene inversion effects in occipitotemporal regions. Although this null result may simply reflect a lack of power due to the small number of subjects in the experiment, it stands in contrast to the significant correlation between the size of the behavioral and fMRI scene inversion effects (but not face inversion effects) observed in medial frontal cortex. Activity in this region is known to decrease relative to baseline when subjects perform difficult cognitive tasks (Gusnard et al., 2001). We hypothesize that some subjects were able to adopt more efficient recognition strategies for inverted scenes that lead to both faster reaction times and less deactivation in the medial frontal lobe. In contrast, recognition strategies might have been less variable for inverted faces. Previous studies have suggested that there is a high degree of automaticity in face processing and body processing (Downing, Bray, Rogers, & Childs, 2004; Mack & Rock, 1998; Ro, Russell, & Lavie, 2001), both stimulus categories for which inversion effects have been established. The flip side of this automaticity may be a lack of flexibility during performance of some cognitive tasks.

Finally, a whole-brain analysis revealed many regions that responded more strongly to upright versus inverted scenes and upright versus inverted faces. Interestingly, some of these regions responded no more strongly to the inverted stimulus than to generic objects. We hypothesize that many of these regions may be involved in post-recognition processes that can only be performed on the upright stimulus. For example, some of the regions found to be more active for upright versus inverted faces have been previously implicated in processing of social information such as expression, emotion, or intent, while some of the regions found to be more active for upright versus inverted scenes have been previously implicated in navigational processing. These results underline the fact that faces and scenes are stimuli of particular interest to human beings and as such they activate specialized neural systems both in occipitotemporal cortex and beyond.

In general, our results suggest that objects, faces, and scenes go through at least three levels of processing that can be affected by stimulus inversion. First, the lateral occipital complex represents visual features prior to their combination into whole objects (Lerner et al., 2001). This region responds more strongly to inverted stimuli irrespective of whether the stimulus is a scene or a face, perhaps because attention is drawn more closely to individual visual features after inversion. LO may also be particularly important for distinguishing between objects that do not share the same parts, such as the objects in this experiment. Second, the ventral occipital temporal region supports distinct (albeit possibly interdigitated—Haxby et al., 2001) mechanisms for representing stimuli as holistic units (FFA), conglomerations of connected parts (MF), or spatially distributed sets of bounding surfaces (PPA). These mechanisms are differentially affected by face and scene inversion. Finally, once the shape of the stimulus is identified and the stimulus is recognized, post-recognition mechanisms operate to determine the social qualities of the individual or the location of the place. These mechanisms engage more readily to upright than to inverted stimuli. Future studies should test these ideas by more closely linking activity in specific cortical regions with performance on behavioral tasks (e.g. Grill-Spector et al., 2004).

### Acknowledgements

We thank Elaina Lord, Justin Lacasse, Nick Tang, Paul Hirschorn, and Mark Brack for their assistance with this experiment, and Paul Downing and Martha Farah for comments on the manuscript. This research was supported in part by NIMH/NINDS grant NS045839 to the Center for Functional Neuroimaging at the University of Pennsylvania and funds provided by the School of Arts and Sciences.

### References

- Aguirre, G. K., & D'Esposito, M. (1999). Topographical disorientation: A synthesis and taxonomy. *Brain*, *122*, 1613–1628.
- Aguirre, G. K., Singh, R., & D'Esposito, M. (1999). Stimulus inversion and the responses of face and object-sensitive cortical areas. *Neuroreport*, *10*(1), 189–194.
- Aguirre, G. K., Zarahn, E., & D'Esposito, M. (1997). Empirical analyses of BOLD fMRI statistics. II. Spatially smoothed data collected under null-hypothesis and experimental conditions. *Neuroimage*, *5*(3), 199–212.
- Aguirre, G. K., Zarahn, E., & D'Esposito, M. (1998). An area within human ventral cortex sensitive to “building” stimuli: Evidence and implications. *Neuron*, *21*(2), 373–383.
- Backus, B. T., Fleet, D. J., Parker, A. J., & Heeger, D. J. (2001). Human cortical activity correlates with stereoscopic depth perception. *Journal of Neurophysiology*, *86*(4), 2054–2068.
- Bartlett, J. C., & Searcy, J. (1993). Inversion and configuration of faces. *Cognitive Psychology*, *25*(3), 281–316.
- Breiter, H. C., Etcoff, N. L., Whalen, P. J., Kennedy, W. A., Rauch, S. L., Buckner, R. L., et al. (1996). Response and habituation of the human amygdala during visual processing of facial expression. *Neuron*, *17*(5), 875–887.
- Burgess, N., Maguire, E. A., Spiers, H. J., & O'Keefe, J. (2001). A temporoparietal and prefrontal network for retrieving the spatial context of lifelike events. *Neuroimage*, *14*(2), 439–453.
- Cao, J. (1999). The size of the connected components of excursion sets of  $\chi^2$ ,  $t$  and  $F$  fields. *Advances in Applied Probability*, *31*(3), 579–595.
- Carey, S., & Diamond, R. (1977). From piecemeal to configurational representation of faces. *Science*, *195*(4275), 312–314.
- Christoff, K., Ream, J. M., & Gabrieli, J. D. (2004). Neural basis of spontaneous thought processes. *Cortex*, *40*(4–5), 623–630.
- Cox, D., Meyers, E., & Sinha, P. (2004). Contextually evoked object-specific responses in human visual cortex. *Science*, *304*(5667), 115–117.
- Cunningham, W. A., Johnson, M. K., Gatenby, J. C., Gore, J. C., & Banaji, M. R. (2003). Neural components of social evaluation. *Journal of Personality & Social Psychology*, *85*(4), 639–649.
- Diamond, R., & Carey, S. (1986). Why faces are and are not special: An effect of expertise. *Journal of Experimental Psychology: General*, *115*(2), 107–117.
- Downing, P. E., Bray, D., Rogers, J., & Childs, C. (2004). Bodies capture attention when nothing is expected. *Cognition*, *93*(1), B27–B38.
- Downing, P. E., Jiang, Y. H., Shuman, M., & Kanwisher, N. (2001). A cortical area selective for visual processing of the human body. *Science*, *293*(5539), 2470–2473.
- Epstein, R. A. (2005). The cortical basis of visual scene processing. *Visual Cognition*, *12*(6), 954–978.
- Epstein, R., DeYoe, E. A., Press, D. Z., Rosen, A. C., & Kanwisher, N. (2001). Neuropsychological evidence for a topographical learning mechanism in parahippocampal cortex. *Cognitive Neuropsychology*, *18*(6), 481–508.
- Epstein, R., & Kanwisher, N. (1998). A cortical representation of the local visual environment. *Nature*, *392*(6676), 598–601.
- Farah, M. J., Tanaka, J. W., & Drain, H. M. (1995). What causes the face inversion effect? *Journal of Experimental Psychology: Human Perception and Performance*, *21*(3), 628–634.
- Farah, M. J., Wilson, K. D., Drain, M., & Tanaka, J. N. (1998). What is “special” about face perception? *Psychology Review*, *105*(3), 482–498.
- Fletcher, P. C., Happe, F., Frith, U., Baker, S. C., Dolan, R. J., Frackowiak, R. S., et al. (1995). Other minds in the brain: A functional imaging study of “theory of mind” in story comprehension. *Cognition*, *57*(2), 109–128.
- Freire, A., Lee, K., & Symons, L. A. (2000). The face-inversion effect as a deficit in the encoding of configural information: Direct evidence. *Perception*, *29*(2), 159–170.
- Friston, K. J., Worsley, K. J., Frackowiak, R. S. J., Mazziotta, J. C., & Evans, A. C. (1994). Assessing the significance of focal activations using their spatial extent. *Human Brain Mapping*, *1*, 210–220.
- Gallagher, H. L., Happe, F., Brunswick, N., Fletcher, P. C., Frith, U., & Frith, C. D. (2000). Reading the mind in cartoons and stories: An fMRI study of ‘theory of mind’ in verbal and nonverbal tasks. *Neuropsychologia*, *38*(1), 11–21.
- Gauthier, I., & Tarr, M. J. (2002). Unraveling mechanisms for expert object recognition: Bridging brain activity and behavior. *Journal of Experimental Psychology: Human Perception and Performance*, *28*(2), 431–446.
- Gauthier, I., Tarr, M. J., Anderson, A. W., Skudlarski, P., & Gore, J. C. (1999). Activation of the middle fusiform ‘face area’ increases with expertise in recognizing novel objects. *Nature Neuroscience*, *2*(6), 568–573.
- Grill-Spector, K. (2003). The neural basis of object perception. *Current Opinion in Neurobiology*, *13*(2), 159–166.
- Grill-Spector, K., Knouf, N., & Kanwisher, N. (2004). The fusiform face area subserves face perception, not generic within-category identification. *Nature Neuroscience*, *7*(5), 555–562.
- Grill-Spector, K., Kourtzi, Z., & Kanwisher, N. (2001). The lateral occipital complex and its role in object recognition. *Vision Research*, *41*(10–11), 1409–1422.
- Gusnard, D. A., Akbudak, E., Shulman, G. L., & Raichle, M. E. (2001). Medial prefrontal cortex and self-referential mental activity: Relation to a default mode of brain function. *Proceedings of the National Academy of Sciences of the United States of America*, *98*(7), 4259–4264.
- Hartley, T., Maguire, E. A., Spiers, H. J., & Burgess, N. (2003). The well-worn route and the path less traveled: Distinct neural bases of route following and wayfinding in humans. *Neuron*, *37*(5), 877–888.
- Hasson, U., Avidan, G., Deouell, L. Y., Bentin, S., & Malach, R. (2003). Face-selective activation in a congenital prosopagnosic subject. *Journal of Cognitive Neuroscience*, *15*(3), 419–431.

- Hasson, U., Harel, M., Levy, I., & Malach, R. (2003). Large-scale mirror-symmetry organization of human occipito-temporal object areas. *Neuron*, 37(6), 1027–1041.
- Haxby, J. V., Gobbini, M. I., Furey, M. L., Ishai, A., Schouten, J. L., & Pietrini, P. (2001). Distributed and overlapping representations of faces and objects in ventral temporal cortex. *Science*, 293(5539), 2425–2430.
- Haxby, J. V., Ungerleider, L. G., Clark, V. P., Schouten, J. L., Hoffman, E. A., & Martin, A. (1999). The effect of face inversion on activity in human neural systems for face and object perception. *Neuron*, 22(1), 189–199.
- Hoffman, E. A., & Haxby, J. V. (2000). Distinct representations of eye gaze and identity in the distributed human neural system for face perception. *Nature Neuroscience*, 3(1), 80–84.
- Ino, T., Inoue, Y., Kage, M., Hirose, S., Kimura, T., & Fukuyama, H. (2002). Mental navigation in humans is processed in the anterior bank of the parieto-occipital sulcus. *Neuroscience Letters*, 322(3), 182–186.
- Kanwisher, N. (2004). The ventral visual object pathway in humans: Evidence from fMRI. In L. M. Chalupa & J. S. Werner (Eds.), *The visual neurosciences: vol. 2* (pp. 1179–1189). Cambridge MA: MIT Press.
- Kanwisher, N., McDermott, J., & Chun, M. M. (1997). The fusiform face area: A module in human extrastriate cortex specialized for face perception. *Journal of Neuroscience*, 17(11), 4302–4311.
- Kanwisher, N., Tong, F., & Nakayama, K. (1998). The effect of face inversion on the human fusiform face area. *Cognition*, 68(1), B1–B11.
- Kastner, S., Pinsk, M. A., De Weerd, P., Desimone, R., & Ungerleider, L. G. (1999). Increased activity in human visual cortex during directed attention in the absence of visual stimulation. *Neuron*, 22(4), 751–761.
- Lerner, Y., Hendler, T., Ben-Bashat, D., Harel, M., & Malach, R. (2001). A hierarchical axis of object processing stages in the human visual cortex. *Cerebral Cortex*, 11(4), 287–297.
- Leube, D. T., Yoon, H. W., Rapp, A., Erb, M., Grodd, W., Bartels, M., et al. (2003). Brain regions sensitive to the face inversion effect: A functional magnetic resonance imaging study in humans. *Neuroscience Letters*, 342(3), 143–146.
- Mack, A., & Rock, I. (1998). *Inattention blindness*. Cambridge MA: MIT Press.
- Maguire, E. A., Burgess, N., Donnett, J. G., Frackowiak, R. S. J., Frith, C. D., & O'Keefe, J. (1998). Knowing where and getting there: A human navigation network. *Science*, 280(5365), 921–924.
- Malach, R., Levy, I., & Hasson, U. (2002). The topography of high-order human object areas. *Trends in Cognitive Sciences*, 6(4), 176–184.
- Malach, R., Reppas, J. B., Benson, R. R., Kwong, K. K., Jiang, H., Kennedy, W. A., et al. (1995). Object-related activity revealed by functional magnetic-resonance-imaging in human occipital cortex. *Proceedings of the National Academy of Sciences of the United States of America*, 92(18), 8135–8139.
- Maurer, D., Grand, R. L., & Mondloch, C. J. (2002). The many faces of configural processing. *Trends in Cognitive Sciences*, 6(6), 255–260.
- McCarthy, G., Puce, A., Gore, J. C., & Allison, T. (1997). Face-specific processing in the human fusiform gyrus. *Journal of Cognitive Neuroscience*, 9(5), 605–610.
- Mendez, M. F., & Chierri, M. M. (2003). Agnosia for scenes in topographagnosia. *Neuropsychologia*, 41(10), 1387–1395.
- Morris, J. S., Frith, C. D., Perrett, D. I., Rowland, D., Young, A. W., Calder, A. J., et al. (1996). A differential neural response in the human amygdala to fearful and happy facial expressions. *Nature*, 383(6603), 812–815.
- Ochsner, K. N. (2004). Current directions in social cognitive neuroscience. *Current Opinion in Neurobiology*, 14(2), 254–258.
- O'Craven, K. M., & Kanwisher, N. (2000). Mental imagery of faces and places activates corresponding stimulus-specific brain regions. *Journal of Cognitive Neuroscience*, 12(6), 1013–1023.
- O'Doherty, J., Winston, J., Critchley, H., Perrett, D., Burt, D. M., & Dolan, R. J. (2003). Beauty in a smile: The role of medial orbitofrontal cortex in facial attractiveness. *Neuropsychologia*, 41(2), 147–155.
- Puce, A., Allison, T., Asgari, M., Gore, J. C., & McCarthy, G. (1996). Differential sensitivity of human visual cortex to faces, letterstrings, and textures: A functional magnetic resonance imaging study. *Journal of Neuroscience*, 16(16), 5205–5215.
- Puce, A., Allison, T., Bentin, S., Gore, J. C., & McCarthy, G. (1998). Temporal cortex activation in humans viewing eye and mouth movements. *Journal of Neuroscience*, 18(6), 2188–2199.
- Reed, C. L., Stone, V. E., Bozova, S., & Tanaka, J. (2003). The body-inversion effect. *Psychological Science*, 14(4), 302–308.
- Rhodes, G., Brake, S., & Atkinson, A. P. (1993). What's lost in inverted faces? *Cognition*, 47(1), 25–57.
- Rhodes, G., Byatt, G., Michie, P. T., & Puce, A. (2004). Is the fusiform face area specialized for faces, individuation, or expert individuation? *Journal of Cognitive Neuroscience*, 16(2), 189–203.
- Ro, T., Russell, C., & Lavie, N. (2001). Changing faces: A detection advantage in the flicker paradigm. *Psychological Science*, 12(1), 94–99.
- Rosenbaum, R. S., Ziegler, M., Wincour, G., Grady, C. L., & Moscovitch, M. (2004). "I have often walked down this street before": fMRI studies on the hippocampus and other structures during mental navigation of an old environment. *Hippocampus*, 14(7), 826–835.
- Rossion, B., & Gauthier, I. (2002). How does the brain process upright and inverted faces? *Behavioral and Cognitive Neuroscience Reviews*, 1(1), 62–74.
- Scapinello, K. F., & Yarmey, A. D. (1970). Role of familiarity and orientation in immediate and delayed recognition of pictorial stimuli. *Psychonomic Science*, 21(6), 329–331.
- Schultz, R. T., Grelotti, D. J., Klin, A., Kleinman, J., Van der Gaag, C., Marois, R., et al. (2003). The role of the fusiform face area in social cognition: Implications for the pathobiology of autism. *Philosophical Transactions of the Royal Society of London-Series B: Biological Sciences*, 358(1430), 415–427.
- Searcy, J. H., & Bartlett, J. C. (1996). Inversion and processing of component and spatial-relational information in faces. *Journal of Experimental Psychology: Human Perception and Performance*, 22(4), 904–915.
- Valentine, T. (1988). Upside-down faces: A review of the effect of inversion upon face recognition. *British Journal of Psychology*, 79(Pt 4), 471–491.
- Whalen, P. J., Rauch, S. L., Etcoff, N. L., McInerney, S. C., Lee, M. B., & Jenike, M. A. (1998). Masked presentations of emotional facial expressions modulate amygdala activity without explicit knowledge. *Journal of Neuroscience*, 18(1), 411–418.
- Wojciulik, E., & Kanwisher, N. (1998). Implicit but not explicit feature binding in a Balint's patient. *Visual Cognition*, 5(1–2), 157–181.
- Wright, A. A., & Roberts, W. A. (1996). Monkey and human face perception: Inversion effects for human faces but not for monkey faces or scenes. *Journal of Cognitive Neuroscience*, 8(3), 278–290.
- Yin, R. K. (1969). Looking at upside-down faces. *Journal of Experimental Psychology*, 81(1), 141–145.
- Yovel, G., & Kanwisher, N. (2004). Face perception: Domain specific, not process specific. *Neuron*, 44(5), 889–898.
- Zarahn, E., Aguirre, G. K., & D'Esposito, M. (1997). Empirical analyses of BOLD fMRI statistics. I. Spatially unsmoothed data collected under null-hypothesis conditions. *Neuroimage*, 5(3), 179–197.