Models of visual emotional perception suggest a reentrant organization of the ventral visual system with the amygdala. Using focused functional magnetic resonance imaging in humans with a sampling rate of 100 ms, here we determine the relative timing of emotional discrimination in amygdala and ventral visual cortical structures during emotional perception. Results show that amygdala and inferotemporal visual cortex differentiate emotional from nonemotional scenes ~1 s before extrastriate occipital cortex, whereas primary occipital cortex shows consistent activity across all scenes. This pattern of discrimination is consistent with a reentrant organization of emotional perception in visual processing, in which transaction between rostral ventral visual cortex and amygdala originates the identification of emotional relevance.

Introduction

Visual perception of emotionally arousing, relative to nonemotional, stimuli is associated with greater blood oxygen level-dependent (BOLD) signal across widespread regions of the ventral visual system, including inferotemporal (IT) and extrastriate occipital cortex (Bradley et al., 2003; Norris et al., 2004; Sabatinelli et al., 2005, 2007; Britton et al., 2006). Although considerable evidence suggests a relationship between activity in amygdala and inferotemporal cortex during emotional perception (Morris et al., 1998; Armony and Dolan, 2002; Vuilleumier et al., 2004; Sabatinelli et al., 2005), the means by which ventral visual system differentiates emotional from nonemotional scenes is not well defined.

Visual emotional discrimination in nonhuman primates is hypothesized to result from reentrant feedback from the amygdala to ventral visual cortex (Spiegler and Mishkin, 1981; Amaral and Price, 1984; Iwai and Yukie, 1987). Specifically, Amaral and colleagues identify dense amygdala innervation into rostral inferotemporal cortex, with more sparse innervation in caudal occipital areas (Amaral et al., 1992, 2003; Freese and Amaral, 2005). This dense interconnectivity, the high correlation between amygdala and inferotemporal activity (Morris et al., 1998; Armony and Dolan, 2002; Vuilleumier et al., 2004; Sabatinelli et al., 2005), and the role of the inferotemporal cortex in high-level visual perception (DeYoe and Van Essen, 1988; Chao et al., 1999; Grill-Spector and Malach, 2004) suggests that the process of emotional discrimination in human visual perception may originate during the interaction between amygdala and rostral inferotemporal cortex and develops later in caudal extrastriate cortex. If emotional discrimination in amygdala and inferotemporal cortex precedes emotional discrimination in occipital cortex, a reentrant model of emotional perception would be supported. If no difference in the timing of emotional discrimination is evident across extrastriate occipital cortex, IT cortex, and amygdala, the reentrant model may be insufficiently described, the recording methodology used here too insensitive, or alternative models relevant to emotional perception exclusive of amygdala feedback may be more appropriate (Heller, 1990; Posner and Petersen, 1990).

It is possible to discriminate cortical sources in ventral visual cortex with surface-based electroencephalography (Sabatinelli et al., 2007; Keil et al., 2009), yet spatial resolution is considerably reduced in rostral inferior temporal areas, in which the distance from the scalp is greatest (Russell et al., 1998; Schoffelen and Gross, 2009). Conversely, differentiating the timing of activation in visual cortex is beyond the temporal resolution of whole-brain functional magnetic resonance imaging (fMRI) acquisition techniques. However, although BOLD contrast is inherently delayed relative to neural activity, the timing of signal change within active clusters is highly reliable (Kim et al., 1997; Menon and Kim 1999; Miezin et al., 2000). By comparing the time course of BOLD signal within regions of interest (ROI) across experimental conditions, the effective temporal resolution is limited only by the sampling rate at which BOLD signal can be recorded and the reliability of the signal. Because we are not concerned with the relative timing of BOLD activation across regions, potential confounds regarding the variations in vascular anatomy, as well as individual differences in BOLD timing (Aguirre et al., 1998; Buxton et al., 1998), are avoided.

Here we record BOLD signal in amygdala and inferotemporal, extrastriate, and striate occipital cortices in a single slice, as rapidly as signal quality will allow, while participants view an event-related series of emotionally arousing and neutral pictures. If emotional discrimination in visual perception originates via feedback between amygdala and rostral inferotemporal cortex, emotional discrimination in these regions (differential BOLD
Three additional blocks, for a total of 96 trials over a stimulus category. The picture series was repeated (in unique orders) pseudorandomized, allowing no more than two successive presentations presented for 3 s, followed by a dot at the center of the screen, as described below.  

**Stimuli and procedure.** Participants were asked to maintain fixation on a dot at the center of a 7 inch liquid crystal display screen mounted directly behind the head, visible via a coil-mounted mirror (IFIS MR-compatible hardware; Intermagnetics). After three acclimation trials in which checkerboard stimuli were presented, a series of 24 picture stimuli were presented (25° visual angle) in an event-related design. The picture stimuli were chosen from the International Affective Picture System (IAPS) (Lang et al., 2008) (http://cepa.hphp.ufl.edu/Media.html) and all depicted people, including eight exemplars each of highly arousing erotic couples (pleasant: 4611, 4658, 4659, 4669, 4676, 4680, 4690, and 4694), moderately arousing neutral people (neutral: 2037, 2102, 2305, 2383, 2393, 2396, 2513, and 2595), and highly arousing mutilations (unpleasant: 3000, 3030, 3060, 3068, 3069, 3100, 3102, and 3225). The pleasant and unpleasant pictures were selected to be equivalent in normative ratings of emotional arousal (Table 1). All picture stimuli were converted to grayscale and matched for luminance and 90% quality JPEG file size by category using Adobe Photoshop 7 (Adobe Systems). Each picture was presented for 3 s, followed by a 9 s fixation-only period. Picture order was pseudorandomized, allowing no more than two successive presentations of a stimulus category. The picture series was repeated (in unique orders) in three additional blocks, for a total of 96 trials over 23 min.

**Scanning parameters.** Once participants were comfortable inside the bore, an 8 min T1-weighted three-dimensional structural volume was acquired. The prescription specified 160 sagittal slices, with 1 mm isotropic voxels in a 256 mm field of view. In addition, a single T1-weighted slice was acquired at the location of the single functional slice acquisition, described below. After structural acquisitions, the 5 mm slice prescription (gradient echo; echoplanar 64 × 64; 180 mm field of view; 25° flip angle; 30 ms echo time; 100 ms repetition time) was oriented in an oblique axial plane such that sampling of amygdala, inferotemporal cortex, and middle occipital gyrus could be obtained. The placement was tailored for each participant, originating with coverage of amygdala, and tilted for optimum coverage of visual areas of interest, and if possible, to exclude sinuses cavity coverage, as a means of limiting susceptibility artifact. This led to substantial sampling of the calcarine fissure. The single-slice prescription allows 40 ml voxels to be repetitively sampled at a temporal resolution of 100 ms.

**Data reduction.** Each participant’s 96-trial functional time series was linearly detrended, temporally smoothed with a 1 s Gaussian filter, and spatially smoothed across 2 voxels (5.625 mm full-width at half-maximum) using BrainVoyager QX 1.8 (Brain Innovation). Temporal smoothing was necessary to reduce the effects of physiological noise present in the time series at a 10 Hz sampling rate. Trials with residual head motion were removed manually, by identifying large (greater than four times the background variation) and brief spikes located by examining the average time series intensity across a majority of the voxels in the slice (a rectangular region of greater than half the voxels within the brain). This procedure resulted in the removal of <2% of total trials and no more than four trials from any subject. Average signal-to-noise [(signal − noise)/SD noise] ratios calculated from the image data were 80.5 in calcarine fissure, 81.0 in middle occipital gyrus, 97.8 in inferotemporal cortex, and 79.4 in amygdala.

The processed image series were entered into single-subject ANOVAs, identifying BOLD signal change evoked by the three picture contents (erotic, neutral people, and mutilations), using a standard two-gamma hemodynamic response function (Boynton et al., 1996). A false discovery rate (Genovese et al., 2002) of p < 0.01 was used to threshold each participant’s data. From these functional maps, four regions were sampled, including bilateral amygdala, inferotemporal cortex, middle occipital gyrus, and midline calcarine fissure. Each region was sampled across 9 voxels (356 ml) using axial neuroanatomical atlases (Talairach and Tournoux, 1988; Haines, 1995) as guides (Fig. 1). The variability of placement across subjects for all ROIs was minimized as much as possible, balancing the need for consistency across subjects with sensitivity to the central location of significant clusters within subject.

**Results**

The effect of picture content on BOLD signal at the peak of the response (4–8 s after picture onset) was greater during pleasant and unpleasant, relative to neutral, picture presentations in middle occipital gyrus (content F(2,15) = 49.93, p < 0.001; quadratic F(1,16) = 64.34, p < 0.001), inferotemporal cortex and middle occipital gyrus could be obtained. The placement was tailored for each participant, originating with coverage of amygdala, and tilted for optimum coverage of visual areas of interest, and if possible, to exclude sinuses cavity coverage, as a means of limiting susceptibility artifact. This led to substantial sampling of the calcarine fissure. The single-slice prescription allows 40 ml voxels to be repetitively sampled at a temporal resolution of 100 ms.

**Table 1. Normative ratings of the IAPS picture stimuli used in the current study**

<table>
<thead>
<tr>
<th>Valence</th>
<th>Arousal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pleasant</td>
<td>6.71 (0.36)</td>
</tr>
<tr>
<td>Neutral</td>
<td>5.27 (0.58)</td>
</tr>
<tr>
<td>Unpleasant</td>
<td>1.68 (0.18)</td>
</tr>
</tbody>
</table>

Valence and arousal are rated on a scale of 1–9, with SD in parentheses.

![Figure 1. Example of statistical overlay of BOLD signal for a single subject. Regions of interest (blue boxes) were derived from significant activity resulting from picture presentation in bilateral amygdala (am), IT, middle occipital gyrus (mOG), and calcarine fissure (calc).](image-url)
emotional arousal in inferotemporal regions relative to middle occipital gyrus.

Discussion

These data show that emotion-related increases in BOLD signal change are reliable in amygdala and inferotemporal visual cortex ~1 s before extrastriate occipital cortex. Because the sequence of basic visual processing stages places extrastriate cortex (V2) ahead of inferotemporal areas (DeYoe and Van Essen, 1988; Desimone and Ungerleider, 1989), evidence of later discrimination supports the perspective (Amaral et al., 1992; Lang et al., 1997; Shi and Davis, 2001; Vuilleumier, 2005) that emotional significance is identified by some means in the amygdala, through which feedback to rostral IT, and eventually caudal occipital areas, leads to enhanced perceptual processing and “motivated attention.”

The integration of emotional significance into visual perception as a result of amygdala — rostrocaudal recurrent processing fits well with conceptions of complex scene processing as an iterative, non-hierarchical mechanism (Lamme and Roelfsema, 2000; Grill-Spector and Malach, 2004; Hegdé and Felleman, 2007). In the primate, the initial inferotemporal cortical response to a picture of a conspecific is thought to reflect global categorization of the percept and is followed by a more sustained response that it associated with detail factors such as identity and facial expression (Nakamura et al., 1994; Sugase et al., 1999; Nishijo et al., 2008). The timing of this later stage of detail processing is consistent with estimates of categorization latency in human research (Junghöfer et al., 2001; VanRullen and Thorpe, 2001; Codispoti et al., 2006; Tsuchiya et al., 2008). Two intracranial studies have shown human amygdala differentiation of aversive from neutral pictures (Oya et al., 2002) and facial expressions (Krolak-Salmon et al., 2004) beginning 150 –200 ms after stimulus onset. We speculate that, in the current dataset, it is this later processing stage that may underlie the increased signal present in amygdala and IT cortex, necessarily delayed and smoothed through hemodynamic BOLD contrast.

The statistical thresholds resulting from the permutation resampling procedure enabled us to identify the time point at which BOLD signal in our regions of interest emotionally arousing trials differed from nonarousing trials. This analysis shows that the amygdala and inferotemporal cortex discriminated picture emotionality before middle occipital gyrus. However, the inferotemporal cortex shows statistically reliable differentiation before amygdala (2.5 vs 2.9 s after picture onset) (Fig. 2). However, a two-factor test of amygdala and inferotemporal arousal discrimination yields no interaction, and thus the difference in discrimination onset is not reliable. Perhaps more importantly, a demonstration that the inferotemporal cortex does not differentiate emotional stimuli in the absence of amygdala input (Vuilleumier et al., 2004) suggests that the discrimination originates in the amygdala and is brought about in inferotemporal areas soon thereafter.

Other means to assess the timing of BOLD signal can approximate high temporal resolution of neural activity, such as image acquisition jitter and modeling of undersampled responses (Lee et al., 2006; Duff et al., 2007; Fuhrmann Alpert et al., 2007). In a study of fear-relevant picture processing (Larson et al., 2006), amygdala activity was recorded in spider phobics and controls using a jittered image acquisition, with an effective sampling rate of 300 ms. Modeling of the BOLD signal response yielded a reduced latency of >1 s in phobics relative to controls in response to spider pictures yet no difference in signal amplitude. Modeling

**Figure 2.** Average time courses (n = 17) of picture-driven BOLD signal in amygdala, inferotemporal cortex, middle occipital gyrus, and calcarine fissure. Blue and red lines represent pleasant and unpleasant pictures, and green represents neutral pictures. The gray bar on the abscissa signifies the picture presentation period. Arrows indicate the point at which BOLD signal during an arousing picture presentation is significantly greater activity than BOLD signal before amygdala (2.5 vs 2.9 s after picture onset) (Fig. 2). However, a two-factor test of amygdala and inferotemporal arousal discrimination yields no interaction, and thus the difference in discrimination onset is not reliable. Perhaps more importantly, a demonstration that the inferotemporal cortex does not differentiate emotional stimuli in the absence of amygdala input (Vuilleumier et al., 2004) suggests that the discrimination originates in the amygdala and is brought about in inferotemporal areas soon thereafter.

Other means to assess the timing of BOLD signal can approximate high temporal resolution of neural activity, such as image acquisition jitter and modeling of undersampled responses (Lee et al., 2006; Duff et al., 2007; Fuhrmann Alpert et al., 2007). In a study of fear-relevant picture processing (Larson et al., 2006), amygdala activity was recorded in spider phobics and controls using a jittered image acquisition, with an effective sampling rate of 300 ms. Modeling of the BOLD signal response yielded a reduced latency of >1 s in phobics relative to controls in response to spider pictures yet no difference in signal amplitude. Modeling
of BOLD signal is undoubtedly a powerful means of assessing the time course of BOLD signal yet is dependent on the accurate specification of many underlying factors. Here we intended to exploit the capability of data collection to the fullest extent and thus reduce the chance of mischaracterization of the data.

Although the timing of emotional discrimination in amygdala and inferotemporal cortex relative to occipital cortex is consistent with a reentrant model of emotional perception, the support is inferential. Beyond the simple timing of emotion differentiation across structures, predictive time series techniques such as Granger causality analyses could provide support for the superordinate role of amygdala differentiation in emotional perception. In the framework of Granger causality, one measured process is said to be causal to a second if the predictability of the second process at a given time point is improved by including measures from the history of the first process. The current dataset did not lend itself to such analyses, because the brief picture periods and event-related design precluded stationary neural processing states of sufficient duration. Future work using fast-sampled fMRI and experimental designs allowing extended, stable periods of emotional processing may be more suitable.

The early difference (1–3 s) in amygdala signal across picture contents can be attributed to both an early increase during arousing pictures and a transient decrease at the onset of neutral pictures. This decrease in amygdala signal in response to nonarousing conditions in studies of emotional processing has been reported in several fMRI studies (Wright et al., 2001; Armony and Dolan, 2002; Morris et al., 2002), but the mechanism is as yet unknown. Future work explicitly controlling the predictability of stimulus conditions and timing may shed light on this possibility.

A tradeoff of focused image acquisition is reduced coverage. This study is intended to test the hypothesis that specific regions of the visual system differed from another in the timing of BOLD change across conditions, which were capable of being sampled in a single plane with the amygdala. Of course there may be other areas of the brain that show emotional differentiation at earlier or later points, and these possibilities can be addressed in additional studies.

In summary, these data show in a human sample the relative timing of emotional discrimination across amygdala and ventral visual cortical structures during emotional perception. In this analysis, amygdala and inferotemporal cortex differentiate emotional from nonemotional scenes ~1 s before secondary occipital cortex, whereas primary occipital cortex shows consistent activity across all scenes. This pattern of discrimination is consistent with a reentrant organization of emotional perception in visual processing, in which transactions between rostral ventral visual cortex and amygdala originate the identification of emotional relevance.

References